The Annals of Applied Probability
2021, Vol. 31, No. 1, 1–38
https://doi.org/10.1214/20-AAP1581

NONPARAMETRIC ESTIMATION FOR LINEAR SPDES FROM LOCAL MEASUREMENTS

BY RANDOLF ALTMEYER[*] AND MARKUS REISS[†]

Institut für Mathematik, Humboldt-Universität zu Berlin, [*]*altmeyrx@math.hu-berlin.de;* [†]*mreiss@math.hu-berlin.de*

The coefficient function of the leading differential operator is estimated from observations of a linear stochastic partial differential equation (SPDE). The estimation is based on continuous time observations which are localised in space. For the asymptotic regime with fixed time horizon and with the spatial resolution of the observations tending to zero, we provide rate-optimal estimators and establish scaling limits of the deterministic PDE and of the SPDE on growing domains. The estimators are robust to lower order perturbations of the underlying differential operator and achieve the parametric rate even in the nonparametric setup with a spatially varying coefficient. A numerical example illustrates the main results.

1. Introduction. While there is a large amount of work on probabilistic, analytical and recently also computational aspects of stochastic partial differential equations (SPDEs), many natural statistical questions are open. With this work we want to enlarge the scope of statistical methodology in two major directions. First, we consider observations of a solution path that are local in space and we ask whether the underlying differential operator or rather its local characteristics can be estimated from this local information only. Second, we allow the coefficients in the differential operator to vary in space and we pursue nonparametric estimation of the coefficient functions, as opposed to parametric estimation approaches for finite-dimensional global parameters in the coefficients. Naturally, both directions are intimately connected.

As a concrete model we consider the parabolic SPDE

$$dX(t) = A_\vartheta X(t)\,dt + B\,dW(t), \qquad t \in [0, T],$$

with the second-order differential operator $A_\vartheta z := \operatorname{div}(\vartheta\nabla z) + \langle a, \nabla z\rangle + bz$ on some bounded domain $\Lambda \subseteq \mathbb{R}^d$ with Dirichlet boundary conditions, see Section 2 for formal details. The coefficient functions ϑ, a, b are unknown on Λ and we aim at estimating $\vartheta : \Lambda \to \mathbb{R}^+$, which models the diffusivity in a stochastic heat equation. The functions a, b as well as the operator B in front of the driving space–time white noise process dW form an unknown nuisance part. Linear SPDEs of this form appear in many applications, including neuroscience (Walsh [37]), oceanography (Frankignoul [13]), geostatistics (Sigrist, Künsch and Stahel [35]), surface growth (Edwards and Wilkinson [38]) and finance (Cont [8]).

Measurements of a solution process X necessarily have a minimal spatial resolution $\delta > 0$ and we dispose of the observations $\langle X(t), K_{\delta,x_0}\rangle$, where the solution is integrated in the spatial domain against a kernel function K_{δ,x_0} with support of diameter δ around some $x_0 \in \Lambda$. We keep the time span T fixed and construct an estimator, called *proxy MLE*, which for the resolution asymptotics $\delta \to 0$ converges at rate δ to $\vartheta(x_0)$ and satisfies a CLT, which we derive in the case of a local multiplication covariance operator B in the SPDE. Another estimator, the so-called *augmented MLE*, will even converge under far more general conditions

Received April 2019; revised November 2019.
MSC2020 subject classifications. Primary 60H15, 60F05; secondary 62G05, 35J15.
Key words and phrases. Nonparametric estimation, fourth moment theorem, Gaussian process, Feynman–Kac, localised scaling limits.

and exhibit a smaller asymptotic variance, but requires a second local observation process $\langle X(t), \Delta K_{\delta,x_0}\rangle$ in terms of the Laplace operator Δ. Clearly, if we have access to these observations around all $x_0 \in \Lambda$, then both estimators can be used to estimate the diffusivity function ϑ nonparametrically on all of Λ.

These results are statistically remarkable. First of all, even for the parametric case that ϑ is a constant, it is not immediately clear that ϑ is identified (i.e., exactly recovered) from local observations in a shrinking neighbourhood around some $x_0 \in \Lambda$ only. Probabilistically, this means that the local observation laws are mutually singular for different values of ϑ. What is more, the bias-variance trade-off paradigm in nonparametric statistics does not apply: asymptotic bias and standard deviation are both of order δ and the CLT provides us even with a simple pointwise confidence interval for ϑ. The robustness of the estimators to lower order parts in the differential operator and unknown B is very attractive for applications. The rate δ is shown to be the best achievable rate in a minimax sense even for constant ϑ without nuisance parts.

The fundamental probabilistic structure behind these results is a universal scaling limit of the observation process for $\delta \to 0$. At a highly localised level, the differential operator A_ϑ behaves like $\vartheta(x_0)\Delta$, as expressed in Corollary 3.6 below, and the construction of the estimators shows a certain scaling invariance with respect to B. To study these scaling limits, we need to consider the deterministic PDE on growing domains via the stochastic Feynman–Kac approach and to deduce tight asymptotics for the action of the semigroup and the heat kernels. Further tools like the fourth moment theorem or the Feldman–Hajek Theorem rely on the underlying Gaussian structure, but extensions to semi-linear SPDEs seem possible.

Let us compare our localisation approach to the spectral approach, introduced by Huebner, Khasminskii and Rozovskii [17] and then in Huebner and Rozovskii [19] for parametric estimation. In the simplest case $A_\vartheta = \vartheta\Delta$ for some $\vartheta > 0$ and B commuting with A_ϑ, the SPDE solution can be expressed in the eigenbasis of the Laplace operator Δ. If the first N coefficient processes (Fourier modes of X) are observed, then a maximum-likelihood estimator for ϑ is asymptotically efficient as $N \to \infty$. This approach has turned out to be very versatile, allowing also for estimating time-dependent $\vartheta(t)$ nonparametrically (Huebner and Lototsky [18]) or to cover nonlinear SPDEs (Cialenco and Glatt-Holtz [6], Pasemann and Stannat [31]). In particular, it helps to understand that the coefficient in the leading order of the differential operator can be estimated with better rates than lower order coefficients. The methodology, however, is intrinsically bound to observations in the spectral domain and to operators A_ϑ whose eigenfunctions, at least in the leading order, are independent of ϑ. In contrast, we work with local observations in space and the unknown spectrum of the operators A_ϑ does not harm us. More conceptually, we rely on the local action of the differential operator A_ϑ, while the spectral approach also applies to an abstract operator in a Hilbert space setting.

Our case of spatially varying coefficients has been considered first by Aihara and Sunahara [3] (with $a = b = 0$) in a filtering problem. The corresponding nonparametric estimation problem is then addressed by Aihara and Bagchi [2] with a sieve least squares estimator, but they achieve consistency only for global observations with a growing time horizon $T \to \infty$. In a stationary one-dimensional setting Bibinger and Trabs [4] ask whether the parameter $\vartheta > 0$ can be estimated when observing the solution only at x_0 over a fixed time interval $[0, T]$. Interestingly, in the case $B = \sigma^2 I$ the parameter ϑ cannot be recovered if the level σ of the space–time white noise is unknown (see also lower bounds of Hildebrandt and Trabs [16]). For a recent and exhaustive survey on statistics for SPDEs we refer to Cialenco [5].

In Section 2 the SPDE and the observation model are introduced and in Section 3 the scaling properties along with the resolution level δ are discussed. Section 4 derives our estimators via a least-squares and a likelihood approach and provides basic insight into their error analysis. The main convergence results as well as a minimax lower bound are presented in Section 5. The findings are illustrated by a numerical example in Section 6. While the main steps in the proofs are presented together with the results, all more technical arguments are delegated to the Appendix.

2. The model.

2.1. *Notation.* Let Λ be a bounded open set in $\mathbb{R}^d$ with C^2-boundary $\partial\Lambda$ and consider $L^2(\Lambda)$ with the usual L^2-norm $\|\cdot\| := \|\cdot\|_{L^2(\Lambda)}$. For any open set U in $\mathbb{R}^d$ and any linear operator $A : L^2(U) \to L^2(U)$ let $\|A\|_{L^2(U)} := \|A\|_{L^2(U)\to L^2(U)}$ denote the operator norm, and let $H^k(U)$ for $k \in \mathbb{N}$ be the L^2-Sobolev spaces. Define $H_0^1(\Lambda)$ as the closure of $C_c^\infty(\overline{\Lambda})$ in $H^1(\Lambda)$. We write $\langle\cdot,\cdot\rangle_{\mathbb{R}^d}$ for the Euclidean inner product and $|\cdot|$ for the norm. Let us define a second order elliptic operator with Dirichlet boundary conditions

$$A_\vartheta = \Delta_\vartheta + A_0, \qquad \mathcal{D}(A_\vartheta) = H_0^1(\Lambda) \cap H^2(\Lambda),$$

where $\Delta_\vartheta z = \operatorname{div}(\vartheta\nabla z) = \sum_{i=1}^d \partial_i(\vartheta\partial_i z)$ is the weighted Laplace operator with spatially varying diffusivity $\vartheta \in C^{1+\alpha}(\overline{\Lambda})$ for $\alpha > 0$, $\min_{x\in\overline{\Lambda}} \vartheta(x) > 0$, and where $A_0 z = \langle a, \nabla z\rangle_{\mathbb{R}^d} + bz$ with functions $a \in C^{1+\alpha}(\overline{\Lambda}; \mathbb{R}^d)$, $b \in C^\alpha(\overline{\Lambda})$. The regularity conditions on ϑ, a, b are such that the deterministic PDE $\frac{d}{dt}u(t) = A_\vartheta^* u(t)$ with initial value $z \in L^2(\Lambda)$ has a sufficiently smooth solution (see proof of Proposition 3.5 below). Let $(S_\vartheta(t))_{t\geq 0}$ denote the analytic semigroup on $L^2(\Lambda)$ generated by A_ϑ (cf. Theorem 3.1.3 of Lunardi [27]), while $(e^{t\Delta})_{t\geq 0}$ is the heat semigroup on $L^2(\mathbb{R}^d)$ generated by $\Delta = \Delta_1$ with domain $H^2(\mathbb{R}^d)$.

2.2. *The SPDE model.* Throughout this work $T < \infty$ is fixed. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$ be a filtered probability space with a cylindrical Brownian motion W on $L^2(\Lambda)$ (dW is also referred to as *space–time* white noise), and let $B : L^2(\Lambda) \to L^2(\Lambda)$ be a bounded linear operator, which is not assumed to be trace class. We study the linear stochastic partial differential equation

$$\begin{cases} dX(t) = A_\vartheta X(t)\,dt + B\,dW(t), & 0 < t \leq T, \\ X(0) = X_0, & \\ X(t)|_{\partial\Lambda} = 0, & 0 < t \leq T, \end{cases} \tag{2.1}$$

with deterministic initial value $X_0 \in L^2(\Lambda)$.

Our statistical analysis below relies on linear functionals of $X(t)$ rather than on $X(t)$ itself. We therefore use the weak solution concept of Da Prato and Zabczyk [9]. If $\int_0^T \|S_\vartheta(t)B\|_{HS(L^2(\Lambda))}^2\,dt < \infty$ with Hilbert–Schmidt norm $\|\cdot\|_{HS(L^2(\Lambda))}$, then the unique weak solution $(X(t))_{0\leq t\leq T}$ of the SPDE (2.1) is given by the variation of constants formula (cf. Theorem 5.4 of Da Prato and Zabczyk [9])

$$X(t) = S_\vartheta(t)X_0 + \int_0^t S_\vartheta(t-s)B\,dW(s). \tag{2.2}$$

It takes values in $L^2(\Lambda)$ and satisfies for $z \in H_0^1(\Lambda) \cap H^2(\Lambda)$

$$d\langle X(t), z\rangle = \langle X(t), A_\vartheta^* z\rangle\,dt + d\langle BW(t), z\rangle. \tag{2.3}$$

Clearly, for $z \in L^2(\Lambda)$

$$\langle X(t), z\rangle = \langle S_\vartheta(t)X_0, z\rangle + \int_0^t \langle S_\vartheta^*(t-s)z, B\,dW(s)\rangle. \tag{2.4}$$

If $\int_0^T \|S_\vartheta(t)B\|^2_{HS(L^2(\Lambda))}\,dt = \infty$, then the stochastic integral in (2.2) is well defined only in a space of distributions. For example, if $H^{-s}(\Lambda)$ is a fractional Sobolev space of negative order with $s > d/4$, then the natural embedding $\iota : L^2(\Lambda) \to H^{-s}(\Lambda)$ is a Hilbert–Schmidt operator such that $\int_0^T \|\iota S_\vartheta(t)B\|^2_{HS(L^2(\Lambda),H^{-s}(\Lambda))}\,dt < \infty$, and $X(t)$ takes values in $H^{-s}(\Lambda)$ (cf. Remark 5.6 of Hairer [15]). Still, (2.3) and (2.4) remain valid, if $\langle X(t), z\rangle$ and $\langle X(t), A_\vartheta^* z\rangle$ are interpreted as dual pairings between $H^{-s}(\Lambda)$ and its dual space for $z \in C_c^\infty(\overline{\Lambda})$.

On the other hand, denote the right hand side of the equation in (2.4) by $\ell(t, z)$ and observe that it is always well-defined for any $z \in L^2(\Lambda)$, independent of the space in which (2.2) makes sense; cf. Lemma 2.4.2 of Liu and Röckner [25]. The resulting process $\ell := (\ell(t, z))_{0\le t\le T, z\in L^2(\Lambda)}$ thus extends the linear forms $z \mapsto \langle X(t), z\rangle$ from $C_c^\infty(\overline{\Lambda})$ to $L^2(\Lambda)$. It has the following properties.

PROPOSITION 2.1. *ℓ is a Gaussian process with mean function $(t, z) \mapsto \langle S_\vartheta(t)X_0, z\rangle$ and covariance function at $0 \le t, t' \le T$, $z, z' \in L^2(\Lambda)$ given by*

$$\operatorname{Cov}\big(\ell(t, z), \ell(t', z')\big) = \int_0^{t\wedge t'} \big\langle B^* S_\vartheta^*(t - s)z, B^* S_\vartheta^*(t' - s)z'\big\rangle\, ds. \tag{2.5}$$

Moreover, ℓ satisfies (2.3) for $z \in H_0^1(\Lambda) \cap H^2(\Lambda)$, if $\langle X(t), z\rangle$ and $\langle X(t), A_\vartheta^* z\rangle$ are replaced by $\ell(t, z)$ and $\ell(t, A_\vartheta^* z)$.

PROOF. By (2.4), $\ell(t, z)$ for $z \in L^2(\Lambda)$ is Gaussian with mean $\langle S_\vartheta(t)X_0, z\rangle$. Itô's isometry (Proposition 4.28 of Da Prato and Zabczyk [9]) proves (2.5). If $z \in C_c^\infty(\overline{\Lambda})$, then $\ell(t, z) = \langle X(t), z\rangle$ satisfies $d\ell(t, z) = \ell(t, A_\vartheta^* z)\,dt + d\langle BW(t), z\rangle$. This extends to $z \in H_0^1(\Lambda) \cap H^2(\Lambda)$ by approximation and continuity of $\ell : [0, T] \times L^2(\Lambda) \mapsto L^2(\mathbb{P})$ from (2.5). $\square$

In the following, justified by this proposition, we write $\langle X(t), z\rangle$ for $0 \le t \le T$ and $z \in L^2(\Lambda)$ instead of $\ell(t, z)$.

2.3. *Local observations.* Throughout this work let $x_0 \in \Lambda$ be fixed. The following rescaling will be useful in the sequel: for $z \in L^2(\mathbb{R}^d)$ and $\delta > 0$ set

$$\Lambda_{\delta,x_0} := \delta^{-1}(\Lambda - x_0) = \{\delta^{-1}(x - x_0) : x \in \Lambda\} \quad \text{and} \quad \Lambda_{0,x_0} := \mathbb{R}^d,$$
$$z_{\delta,x_0}(x) := \delta^{-d/2} z\big(\delta^{-1}(x - x_0)\big), \quad x \in \mathbb{R}^d.$$

Fix a function $K \in H^2(\mathbb{R}^d)$, called kernel, with compact support in Λ_{δ,x_0}. The compact support ensures that K_{δ,x_0} is localized around x_0 and $K_{\delta,x_0} \in H_0^1(\Lambda) \cap H^2(\Lambda)$, $\|K_{\delta,x_0}\| = \|K\|_{L^2(\mathbb{R}^d)}$. Local measurements of X at x_0 with resolution level δ until time T are described by the real-valued processes $X_{\delta,x_0} = (X_{\delta,x_0}(t))_{0\le t\le T}$, $X^\Delta_{\delta,x_0} = (X^\Delta_{\delta,x_0}(t))_{0\le t\le T}$,

$$X_{\delta,x_0}(t) = \big\langle X(t), K_{\delta,x_0}\big\rangle, \tag{2.6}$$

$$X^\Delta_{\delta,x_0}(t) = \big\langle X(t), \Delta K_{\delta,x_0}\big\rangle. \tag{2.7}$$

Note that it is sufficient to observe $X_{\delta,x}(t)$ for x in a neighbourhood of x_0 in order to provide us with $X^\Delta_{\delta,x_0}(t) = \Delta X_{\delta,\cdot}(t)|_{x=x_0}$. Examples for K can be found in Section 6.

The process X_{δ,x_0} satisfies $X_{\delta,x_0}(0) = \langle X_0, K_{\delta,x_0}\rangle$ and

$$dX_{\delta,x_0}(t) = \big\langle X(t), A_\vartheta^* K_{\delta,x_0}\big\rangle\,dt + \big\|B^* K_{\delta,x_0}\big\|\,d\overline{W}(t) \tag{2.8}$$

with the scalar Brownian motion $\overline{W}(t) = \langle BW(t), K_{\delta,x_0}\rangle / \|B^* K_{\delta,x_0}\|$, whenever $\|B^* K_{\delta,x_0}\| > 0$.

3. Scaling assumptions.

3.1. *Rescaled operators and semigroups.* Let us study how A_ϑ^* and $S_\vartheta^*(t)$ act on localized functions z_{δ,x_0}. For this note first that $A_\vartheta^* = \Delta_\vartheta + A_0^*$ with $A_0^* z = -\operatorname{div}(az) + bz$ has domain $\mathcal{D}(A_\vartheta^*) = H_0^1(\Lambda) \cap H^2(\Lambda)$. For $\delta > 0$ define similarly the operator $A_{\vartheta,\delta,x_0}^* = \Delta_{\vartheta(x_0+\delta\cdot)} + A_{0,\delta,x_0}^*$ with domain $\mathcal{D}(A_{\vartheta,\delta,x_0}^*) = H_0^1(\Lambda_{\delta,x_0}) \cap H^2(\Lambda_{\delta,x_0})$, where for $z \in C_c^\infty(\Lambda_{\delta,x_0})$

$$A_{0,\delta,x_0}^* z = -\delta \operatorname{div}\big(a(x_0+\delta\cdot)z\big) + \delta^2 b(x_0+\delta\cdot)z. \tag{3.1}$$

The operator $A_{\vartheta,\delta,x_0}^*$ generates again an analytic semigroup $(S_{\vartheta,\delta,x_0}^*(t))_{t\geq 0}$ on $L^2(\Lambda_{\delta,x_0})$ (Lemma 7.3.4 of Pazy [32]). The following scaling properties are fundamental for our analysis:

LEMMA 3.1. *For $\delta > 0$:*

(i) *If $z \in H_0^1(\Lambda_{\delta,x_0}) \cap H^2(\Lambda_{\delta,x_0})$, then $A_\vartheta^* z_{\delta,x_0} = \delta^{-2}(A_{\vartheta,\delta,x_0}^* z)_{\delta,x_0}$.*
(ii) *If $z \in L^2(\Lambda_{\delta,x_0})$, then $S_\vartheta^*(t) z_{\delta,x_0} = (S_{\vartheta,\delta,x_0}^*(t\delta^{-2})z)_{\delta,x_0}$, $t \geq 0$.*

PROOF. It suffices to prove the result for $z \in C_c^\infty(\overline{\Lambda}_{\delta,x_0})$. In this case, (i) follows immediately, noting that $z_{\delta,x_0} \in C_c^\infty(\overline{\Lambda})$. For (ii) set $w(t) = (S_{\vartheta,\delta,x_0}^*(t\delta^{-2})z)_{\delta,x_0} \in L^2(\Lambda)$. As $(S_{\vartheta,\delta,x_0}^*(t))_{t\geq 0}$ is an analytic semigroup, we have $S_{\vartheta,\delta,x_0}^*(t)z \in \mathcal{D}(A_{\vartheta,\delta,x_0}^*) = H_0^1(\Lambda_{\delta,x_0}) \cap H^2(\Lambda_{\delta,x_0})$ and so by (i)

$$\frac{d}{dt} w(t) = \delta^{-2}\big(A_{\vartheta,\delta,x_0}^* S_{\vartheta,\delta,x_0}^*\big(t\delta^{-2}\big)z\big)_{\delta,x_0} = A_\vartheta^* w(t).$$

Since $w(0) = z_{\delta,x_0}$, we conclude that $w(t) = S_\vartheta^*(t) z_{\delta,x_0}$ from $u(t) = S_\vartheta^*(t) z_{\delta,x_0}$ being the unique solution in $C([0,\infty); H_0^1(\Lambda) \cap H^2(\Lambda)) \cap C^1([0,\infty); L^2(\Lambda))$ of

$$\frac{d}{dt} u(t) = A_\vartheta^* u(t), \quad t \geq 0, u(0) = z_{\delta,x_0}. \qquad \square$$

Applying $S_\vartheta^*(t)$ to a localized function z_{δ,x_0} is therefore equivalent to applying a different semigroup, rescaled in time and space, to the fixed function z.

3.2. *Scaling of B.* Just as with A_ϑ^* we also need that B^* behaves nicely when applied to localized functions. For this we shall assume a scaling limit for B^*, which does not degenerate in combination with K.

ASSUMPTION 3.2. There are bounded linear operators $B_{\delta,x_0}, B_{0,x_0} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ such that $B^*(z_{\delta,x_0}) = (B_{\delta,x_0}^* z)_{\delta,x_0}$ for $z \in L^2(\mathbb{R}^d)$ with support in Λ_{δ,x_0} and $B_{\delta,x_0}^* z \to B_{0,x_0}^* z$ for $z \in L^2(\mathbb{R}^d)$ and $\delta \to 0$. Introducing

$$\Psi(z, z') := \int_0^\infty \langle B_{0,x_0}^* e^{s\Delta} z, B_{0,x_0}^* e^{s\Delta} z' \rangle_{L^2(\mathbb{R}^d)} \, ds, \quad z, z' \in L^2(\mathbb{R}^d), \tag{3.2}$$

assume the *nondegeneracy conditions* $\|B_{0,x_0}^* K\|_{L^2(\mathbb{R}^d)} > 0$, $\Psi(\Delta K, \Delta K) > 0$.

REMARK 3.3. We shall see that after an appropriate rescaling $\vartheta(x_0)^{-1}\Psi(z, z')$ becomes the limiting covariance in (2.5) (cf. Proposition A.8 below). $\Psi(\Delta K, \Delta K)$ is always nonnegative and finite because

$$\Psi(\Delta K, \Delta K) \leq \|B_{0,x_0}^*\|_{L^2(\mathbb{R}^d)}^2 \int_0^\infty \|e^{s\Delta}\Delta K\|_{L^2(\mathbb{R}^d)}^2 \, ds \leq \frac{\|B_{0,x_0}^*\|_{L^2(\mathbb{R}^d)}^2}{2} \|\nabla K\|_{L^2(\mathbb{R}^d)}^2,$$

using $\|e^{s\Delta}\Delta K\|_{L^2(\mathbb{R}^d)}^2 = \langle e^{2s\Delta}\Delta K, \Delta K\rangle_{L^2(\mathbb{R}^d)}$ and $\int_0^\infty e^{2s\Delta}\Delta K \, ds = -\frac{1}{2}K$.

EXAMPLES 3.4.

(a) For a bounded continuous function $\sigma : \mathbb{R}^d \to (0, \infty)$ define the multiplication operator $M_\sigma : L^2(\Lambda) \to L^2(\Lambda)$, $M_\sigma z(x) := (\sigma z)(x) = \sigma(x)z(x)$. With $B = B^* = M_\sigma$ the SPDE in (2.1) can be written informally as

$$\dot{X}(t, x) = A_\vartheta X(t, x) + \sigma(x)\dot{W}(t, x), \quad 0 < t \leq T, x \in \Lambda.$$

Note that B^* commutes with A_ϑ only if σ is constant. For $z \in L^2(\Lambda_{\delta,x_0})$ we find that $B^* z_{\delta,x_0} = (M_{\sigma(\delta\cdot+x_0)}z)_{\delta,x_0}$ and so $B_{\delta,x_0} = M_{\sigma(\delta\cdot+x_0)}$. Then $\|B^*_{\delta,x_0}z - \sigma(x_0)z\|_{L^2(\mathbb{R}^d)} \to 0$ for $z \in L^2(\mathbb{R}^d)$, $\delta \to 0$, and thus $B^*_{0,x_0} = M_{\sigma(x_0)}$ is the multiplication operator on $L^2(\mathbb{R}^d)$ with the constant $\sigma(x_0)$. For $z \in H^2(\mathbb{R}^d)$, $z' \in L^2(\mathbb{R}^d)$ we have (cf. Remark 3.3)

$$\Psi(\Delta z, z') = -\frac{\sigma^2(x_0)}{2}\langle z, z'\rangle_{L^2(\mathbb{R}^d)} \tag{3.3}$$

and integration by parts shows $\Psi(\Delta K, \Delta K) = \frac{\sigma^2(x_0)}{2}\|\nabla K\|^2_{L^2(\mathbb{R}^d)}$. The nondegeneracy conditions are clearly satisfied.

(b) Let σ be as in (a) and consider with bounded $\eta \in C^2(\mathbb{R}^d)$, $\min_{x\in\mathbb{R}^d} \eta(x) > 0$, the perturbed multiplication operator $B = B^* = M_\sigma + (-\Delta_\eta)^{-\gamma}$, $\gamma > 0$. By functional calculus $B^* z_{\delta,x_0} = (B^*_{\delta,x_0}z)_{\delta,x_0}$ for $z \in L^2(\Lambda_{\delta,x_0})$ with $B_{\delta,x_0} = M_{\sigma(\delta\cdot+x_0)} + \delta^{2\gamma}(-\Delta_{\eta(\delta\cdot+x_0)})^{-\gamma}$ and $\|B^*_{\delta,x_0}z - \sigma(x_0)z\|_{L^2(\mathbb{R}^d)} \to 0$ for $z \in L^2(\mathbb{R}^d)$, $\delta \to 0$. B_{0,x_0} and $\Psi(\Delta K, \Delta K)$ are as in (a).

(c) Assumption 3.2 excludes $B = (-\Delta)^{-\gamma}$, $\gamma > 0$, a typical choice to obtain smooth solutions X; cf. Da Prato and Zabczyk [9], Chapter 5.5. Indeed, by (b) $B^*_{\delta,x_0} = \delta^{2\gamma}(-\Delta)^{-\gamma}$ and so $B^*_{0,x_0} = 0$, violating the nondegeneracy conditions. This problem can be solved by modifying the test function K_{δ,x_0}. For example, if $A_\vartheta = \vartheta\Delta$ for constant $\vartheta > 0$ and $X_0 \in \mathcal{D}((-\Delta)^\gamma)$, then assume we have access to $\langle X(t), (-\Delta)^\gamma K_{\delta,x_0}\rangle$, $\langle X(t), (-\Delta)^\gamma \Delta K_{\delta,x_0}\rangle$ instead of (2.6), (2.7). Since B and A_ϑ commute, $\langle X(\cdot), (-\Delta)^\gamma K_{\delta,x_0}\rangle$ has the same distribution as $\langle \tilde{X}(\cdot), K_{\delta,x_0}\rangle$, where $\tilde{X}$ corresponds to the SPDE (2.1) with $B = I$ and $\tilde{X}_0 = (-\Delta)^\gamma X_0$, and so Assumption 3.2 is satisfied.

3.3. *From bounded to unbounded domains*. Lemma 3.1 and Assumption 3.2 allow us to rewrite the covariance function of X_{δ,x_0} for $t, t' \geq 0$:

$$\begin{aligned} &\operatorname{Cov}(\delta^{-1}X_{\delta,x_0}(t\delta^2), \delta^{-1}X_{\delta,x_0}(t'\delta^2)) \\ &\quad= \int_0^{t\wedge t'} \langle B^*_{\delta,x_0}S^*_{\vartheta,\delta,x_0}(t-s)K, B^*_{\delta,x_0}S^*_{\vartheta,\delta,x_0}(t'-s)K\rangle_{L^2(\Lambda_{\delta,x_0})}\, ds. \end{aligned} \tag{3.4}$$

In order to see how this behaves when $\delta \to 0$, note that the domain Λ_{δ,x_0} grows and we find from (3.1) that $A^*_{\vartheta,\delta,x_0}K \to \vartheta(x_0)\Delta K$ in $L^2(\mathbb{R}^d)$. This motivates the following result, proved in Appendix A.2.

PROPOSITION 3.5. *For $t > 0$:*

(i) *If $\delta > 0$ and $z \in C(\overline{\Lambda}_{\delta,x_0})$, then $|(S^*_{\vartheta,\delta,x_0}(t)z)(x)| \leq c_3 e^{c_1\delta^2 t}(e^{c_2 t\Delta}|z|)(x)$ for all $x \in \Lambda_{\delta,x_0}$ with universal constants $c_1, c_2, c_3 > 0$.*
(ii) *If $z \in L^2(\mathbb{R}^d)$, then $S^*_{\vartheta,\delta,x_0}(t)(z|_{\Lambda_{\delta,x_0}}) \to e^{\vartheta(x_0)t\Delta}z$ in $L^2(\mathbb{R}^d)$ for $\delta \to 0$.*

This means that the solution of

$$\frac{d}{dt}u^{(\delta)}(t) = (A^*_{\vartheta,\delta,x_0}u^{(\delta)})(t), \quad u^{(\delta)}(0) = z,$$

on $L^2(\Lambda_{\delta,x_0})$ with *bounded domain* Λ_{δ,x_0} converges to the solution of the heat equation

$$\frac{d}{dt}u(t) = \vartheta(x_0)\Delta u(t), \quad u(0) = z,$$

on $L^2(\mathbb{R}^d)$ with *unbounded domain* $\mathbb{R}^d$. This scaling limit, which seems natural but is nevertheless nontrivial, lies at the heart of the analysis for the covariance function. Yet, the convergence in Proposition 3.5(ii) does not hold uniformly in z, which complicates the approximations in the covariance analysis.

Applying the proposition to (3.4) also implies a scaling limit for the SPDE in (2.1), where for simplicity a zero initial condition is assumed:

THEOREM 3.6. *Let $X_0 = 0$ and set $Z_\delta(t,z) := \delta^{-1}\langle X(t\delta^2), (z|_{\Lambda_{\delta,x_0}})_{\delta,x_0}\rangle$ for $t \geq 0$, $z \in L^2(\mathbb{R}^d)$. Under Assumption 3.2 the finite dimensional distributions of $(Z_\delta(t,z))_{t\geq 0, z\in L^2(\mathbb{R}^d)}$ converge to those of $(Z_0(t,z))_{t\geq 0, z\in L^2(\mathbb{R}^d)}$, $Z_0(t,z) = \langle Y(t), z\rangle_{L^2(\mathbb{R}^d)}$, solving the stochastic heat equation on $L^2(\mathbb{R}^d)$ with space–time white noise dW on $L^2(\mathbb{R}^d)$:*

$$\begin{cases} dY(t) = \vartheta(x_0)\Delta Y(t)\,dt + B_{0,x_0}\,dW(t), & t > 0, \\ Y(0) = 0. \end{cases}$$

PROOF. According to (3.4) Z_δ is a centered Gaussian process with covariance function $\operatorname{Cov}(Z_\delta(t,z), Z_\delta(t',z'))$ for $t, t' \geq 0$, $z, z' \in L^2(\mathbb{R}^d)$ equal to

$$\int_0^{t\wedge t'} \langle B^*_{\delta,x_0} S^*_{\vartheta,\delta,x_0}(t-s)(z|_{\Lambda_{\delta,x_0}}), B^*_{\delta,x_0} S^*_{\vartheta,\delta,x_0}(t'-s)(z'|_{\Lambda_{\delta,x_0}})\rangle_{L^2(\Lambda_{\delta,x_0})}\,ds.$$

It is enough to show that this converges to

$$\operatorname{Cov}(Z_0(t,z), Z_0(t',z')) = \int_0^{t\wedge t'} \langle B^*_{0,x_0} e^{\vartheta(x_0)(t-s)\Delta} z, B^*_{0,x_0} e^{\vartheta(x_0)(t'-s)\Delta} z'\rangle_{L^2(\mathbb{R}^d)}\,ds.$$

Approximating z by continuous functions, the semigroup bound in Proposition 3.5(i) gives $\sup_{0<\delta\leq 1}\sup_{s\leq t}\|S^*_{\vartheta,\delta,x_0}(s)(z|_{\Lambda_{\delta,x_0}})\|_{L^2(\Lambda_{\delta,x_0})} < \infty$, while Assumption 3.2 implies $B^*_{\delta,x_0}u^{(\delta)} \to B^*_{0,x_0}u$ for any $u^{(\delta)} \to u$, invoking the uniform boundedness principle. By Proposition 3.5(ii) we have $S^*_{\vartheta,\delta,x_0}(s)(z|_{\Lambda_{\delta,x_0}}) \to e^{\vartheta(x_0)s\Delta}z$ in $L^2(\mathbb{R}^d)$. Arguing in the same way with respect to z', the dominated convergence theorem shows the claim. □

This theorem demonstrates the strength of local measurements that at small scales the highest order differential operator dominates, together with the local coefficient $\vartheta(x_0)$ and the local operator B_{0,x_0} in the noise.

3.4. *The initial condition.* For X_0 we require the following scaling behaviour:

ASSUMPTION 3.7 ($z;\beta$). For $\beta > 0$ and $z \in H^2(\mathbb{R}^d)$ with compact support in Λ_{δ,x_0} for $\delta > 0$, the initial condition X_0 satisfies

$$\int_0^T \langle S_\vartheta(t)X_0, (\Delta z)_{\delta,x_0}\rangle^2\,dt = o(\ell_{d,2}(\delta)^{-1}\delta^\beta), \quad \delta \to 0,$$

where $\ell_{d,2}(\delta) = \log(\delta^{-1})$ for $d = 2$ and $\ell_{d,2}(\delta) = 1$ otherwise.

Under this assumption the initial condition becomes negligible in the estimation procedure. It is true under general conditions.

LEMMA 3.8. *Assumption* 3.7$(z;\beta)$ *is satisfied for all* $z\in H^2(\mathbb{R}^d)$ *with compact support in* Λ_{δ,x_0} *for* $\delta>0$ *and*:

(i) $\beta=2$ *if* $X_0\in L^p(\Lambda)$ *for some* $p>2$, *in particular if* $X_0\in C(\overline{\Lambda})$,
(ii) $\beta=3$ *if* $X_0\in\mathcal{D}(A_\vartheta)$ *and for* $d=1$ *also* $\vartheta\in C^{1+\alpha'}(\overline{\Lambda})$, $\alpha'>1/2$, *and* $\int_{\mathbb{R}}z(x)\,dx=0$.

PROOF. This follows from Lemma A.7(ii), (iii) below, noting $\gamma(d,p)>2$ in (ii) for $p>2$, $d\geq 1$. $\square$

4. The two estimation methods.

4.1. *Motivation and construction.* We give two motivations for deriving the estimators in the parametric case $A_\vartheta=\vartheta\Delta$ with constant $\vartheta>0$, $B=I$. As we shall see later, these estimators will then work quite universally for nonparametric specifications of ϑ and general A_ϑ and B.

Least squares approach. In the deterministic situation of (2.8) without driving noise (i.e., $A_\vartheta=\vartheta\Delta$ and $B=0$) we recover ϑ via $\dot X_{\delta,x_0}(t)=\vartheta X^{\Delta}_{\delta,x_0}(t)$ for all $t\in[0,T]$. A standard least-squares ansatz in the noisy situation would therefore lead to an estimator $\hat\vartheta=\operatorname{argmin}_\vartheta\int_0^T(\dot X_{\delta,x_0}(t)-\vartheta X^{\Delta}_{\delta,x_0}(t))^2\,dt$. While this itself is certainly not well defined, the corresponding normal equations yield the feasible estimator

$$\hat\vartheta_\delta^{LS}=\frac{\int_0^T X^{\Delta}_{\delta,x_0}(t)\,dX_{\delta,x_0}(t)}{\int_0^T X^{\Delta}_{\delta,x_0}(t)^2\,dt},$$

compare with the approach by Maslowski and Tudor [28] for fractional noise.

Likelihood approach. Assume that only X_{δ,x_0} in (2.8) is observed with $A_\vartheta=\vartheta\Delta$, $B=I$. Denote by $\mathbb{P}_\vartheta^{\delta,x_0}$ and $\mathbb{P}_0$ the laws of X_{δ,x_0} and $\|K\|_{L^2(\mathbb{R}^d)}\overline{W}$ on the canonical path space $(C([0,T]),\|\cdot\|_\infty)$ equipped with its Borel sigma algebra. Typically, the likelihood of $\mathbb{P}_\vartheta^{\delta,x_0}$ with respect to $\mathbb{P}_0$ is determined via Girsanov's theorem. This is not immediate from (2.8), because X^{Δ}_{δ,x_0} cannot be obtained from X_{δ,x_0} for fixed x_0. Therefore we employ Liptser and Shiryaev [24], Theorem 7.17 and write X_{δ,x_0} as the diffusion-type process

$$dX_{\delta,x_0}(t)=\vartheta m_\vartheta(t)\,dt+\|K\|_{L^2(\mathbb{R}^d)}\,d\tilde W(t),\qquad t\in[0,T],$$

with a different scalar Brownian motion $\tilde W=(\tilde W(t))_{0\leq t\leq T}$, adapted to the filtration generated by X_{δ,x_0}, and

$$m_\vartheta(t)=\mathbb{E}_\vartheta\big[X^{\Delta}_{\delta,x_0}(t)|\big(X_{\delta,x_0}(s)\big)_{0\leq s\leq t}\big].$$

Girsanov's theorem in the form of Liptser and Shiryaev [24], Theorem 7.18, applies and we find that $\mathbb{P}_\vartheta^{\delta,x_0}$ has with respect to $\mathbb{P}_0$ the likelihood

$$\mathcal{L}(\vartheta,X_{\delta,x_0})=\exp\bigg(\vartheta\int_0^T\frac{m_\vartheta(t)}{\|K\|^2_{L^2(\mathbb{R}^d)}}\,dX_{\delta,x_0}(t)-\frac{\vartheta^2}{2}\int_0^T\frac{m_\vartheta(t)^2}{\|K\|^2_{L^2(\mathbb{R}^d)}}\,dt\bigg).$$

Computing the conditional expectation $m_\vartheta(t)$ is a nonexplicit filtering problem, even in the parametric case $A_\vartheta=\vartheta\Delta$. In particular, m_ϑ depends on ϑ in a highly nonlinear way. We pursue two different modifications:

Augmented MLE. If we observe X^{Δ}_{δ,x_0} additionally, then we can just replace the conditional expectation $m_\vartheta(t)$ in the likelihood by its argument $X^{\Delta}_{\delta,x_0}(t)$, which is in particular independent of ϑ. Maximizing this modified likelihood leads to the augmented MLE

$$\hat{\vartheta}^A_\delta = \frac{\int_0^T X^{\Delta}_{\delta,x_0}(t)\,dX_{\delta,x_0}(t)}{\int_0^T X^{\Delta}_{\delta,x_0}(t)^2\,dt}. \tag{4.1}$$

We remark that $\hat{\vartheta}^A_\delta = \hat{\vartheta}^{LS}_\delta$.

Proxy MLE. If we do not dispose of additional observations, we can approximate $m_\vartheta(t)$ by the conditional expectation $\mathbb{E}_\vartheta[X^{\Delta}_{\delta,x_0}(t)|X_{\delta,x_0}(t)]$. In our simplified setup with $A_\vartheta = \vartheta\Delta$ and $B = I$ the projected finite-dimensional process $(\langle X(t), z_i\rangle)_{1\le i\le m}$ for $z_i \in L^2(\Lambda)$ admits a stationary solution $\langle X(t), z_i\rangle = \int_{-\infty}^t \langle S_\vartheta(t-s)z_i, dW(s)\rangle$, $i = 1,\dots,m$, with a two-sided cylindrical Brownian motion $(W(t))_{t\in\mathbb{R}}$, provided the variances remain finite. Note that we need not require $X_0 = \int_{-\infty}^t S^*_\vartheta(t-s)\,dW(s)$ to exist, but only that the finite-dimensional projection $(\langle X_0, z_i\rangle)_{1\le i\le m}$ follows the right law, which is always feasible. If we choose $z_1 = K_{\delta,x_0}$, $z_2 = \Delta K_{\delta,x_0}$, then the process $(X_{\delta,x_0}, X^{\Delta}_{\delta,x_0})$ is stationary with

$$\begin{aligned}
\operatorname{Var}(X_{\delta,x_0}(t)) &= \int_{-\infty}^t \langle S_\vartheta(2t-2s)K_{\delta,x_0}, K_{\delta,x_0}\rangle\,ds \\
&= \frac{1}{2\vartheta}\langle(-\Delta)^{-1}K_{\delta,x_0}, K_{\delta,x_0}\rangle, \qquad (4.2)\\
\operatorname{Cov}(X^{\Delta}_{\delta,x_0}(t), X_{\delta,x_0}(t)) &= \int_{-\infty}^t \langle S_\vartheta(2t-2s)\Delta K_{\delta,x_0}, K_{\delta,x_0}\rangle\,ds = \frac{-1}{2\vartheta}\|K\|^2_{L^2(\mathbb{R}^d)}.
\end{aligned}$$

In general, $\langle(-\Delta)^{-1}K_{\delta,x_0}, K_{\delta,x_0}\rangle$ may not exist, but if we assume the existence of $\tilde{K} \in H^4(\mathbb{R}^d)$ with $\Delta\tilde{K} = K$ and compact support in Λ_{δ,x_0}, then by the scaling in Lemma 3.1, $\operatorname{Var}(X_{\delta,x_0}(t)) = \frac{\delta^2}{2\vartheta}\|\nabla\tilde{K}\|^2_{L^2(\mathbb{R}^d)} < \infty$ follows. In this situation we therefore find that $\mathbb{E}_\vartheta[X^{\Delta}_{\delta,x_0}(t)|X_{\delta,x_0}(t)]$ equals

$$\frac{\operatorname{Cov}(X^{\Delta}_{\delta,x_0}(t), X_{\delta,x_0}(t))}{\operatorname{Var}(X_{\delta,x_0}(t))}X_{\delta,x_0}(t) = \frac{-\delta^{-2}\|K\|^2_{L^2(\mathbb{R}^d)}}{\|\nabla\tilde{K}\|^2_{L^2(\mathbb{R}^d)}}X_{\delta,x_0}(t).$$

This expression is again independent of ϑ. Using it as an approximation of $m_\vartheta(t)$ in the likelihood and neglecting the boundary terms in the identity $2\int_0^T X_{\delta,x_0}(t)\,dX_{\delta,x_0}(t) = (X^2_{\delta,x_0}(T) - X^2_{\delta,x_0}(0)) - \langle X_{\delta,x_0}\rangle_T$ with quadratic variation $\langle X_{\delta,x_0}\rangle_T$, we obtain the proxy MLE

$$\hat{\vartheta}^P_\delta = \frac{\|\nabla\tilde{K}\|^2_{L^2(\mathbb{R}^d)}}{2\|K\|^2_{L^2(\mathbb{R}^d)}}\frac{\langle X_{\delta,x_0}\rangle_T}{\delta^{-2}\int_0^T X_{\delta,x_0}(t)^2\,dt}. \tag{4.3}$$

Note that the quadratic variation $\langle X_{\delta,x_0}\rangle_T = T\|B^*K_{\delta,x_0}\|^2$ is known to us from observing X_{δ,x_0} continuously in time.

REMARK 4.1. A sufficient condition for the existence of $\tilde{K}$ is $\int_{\mathbb{R}^d} K(x)\,dx = 0$, $\int_{\mathbb{R}^d} xK(x)\,dx = 0$ by Lemma A.5(iii) below.

4.2. *Basic error decomposition.* Let us discuss the basic error analysis for the augmented MLE $\hat{\vartheta}_\delta^A$ and the proxy MLE $\hat{\vartheta}_\delta^P$ in the general nonparametric framework of Section 2. Since we only use local measurements around x_0, it might be expected that asymptotically we are lead to estimating $\vartheta(x_0)$. Let us point out that this is indeed true, but a priori not clear because all values of $\vartheta(x)$ enter into the observations X_{δ,x_0} and it must be excluded that the resulting bias spoils the estimator.

Augmented MLE. Consider $\hat{\vartheta}_\delta^A(x_0) = \hat{\vartheta}_\delta^A$ from (4.1). Then insertion of equation (2.8) for $dX_{\delta,x_0}(t)$ yields the decomposition

$$\hat{\vartheta}_\delta^A(x_0) = \vartheta(x_0) + \|B^* K_{\delta,x_0}\| (\mathcal{I}_\delta^A)^{-1} M_\delta^A + (\mathcal{I}_\delta^A)^{-1} R_\delta^A, \tag{4.4}$$

with

$$\begin{aligned} M_\delta^A &= \int_0^T X_{\delta,x_0}^\Delta(t)\, d\overline{W}(t) \quad \text{(martingale part)},\\ \mathcal{I}_\delta^A &= \int_0^T X_{\delta,x_0}^\Delta(t)^2\, dt, \quad \text{(observed Fisher information)},\\ R_\delta^A &= \int_0^T X_{\delta,x_0}^\Delta(t) \langle X(t), (A_\vartheta^* - \vartheta(x_0)\Delta) K_{\delta,x_0}\rangle\, dt \quad \text{(remaining bias)}. \end{aligned}$$

Let us note that $\mathcal{I}_\delta^A$ is not the observed Fisher information in a strict sense (due to the appearance of m_ϑ in the likelihood), but it plays the same role, compare the analysis of the MLE for Ornstein–Uhlenbeck processes in Kutoyants [23]. In particular, it forms the quadratic variation of the martingale M_δ^A. In the specific case $A_\vartheta = \vartheta\Delta$ for some parametric $\vartheta > 0$ the term R_δ^A vanishes, otherwise it induces a bias due to the variations of ϑ around $\vartheta(x_0)$ and due to first and zero order differential operators that may appear in A_ϑ.

As the error structure suggests, the augmented MLE $\hat{\vartheta}_\delta^A(x_0)$ is a consistent estimator for $\delta \to 0$ if the observed Fisher information satisfies $\mathcal{I}_\delta^A \to \infty$. In the simple stationary case of (4.2) we obtain $\mathbb{E}[\mathcal{I}_\delta^A] = \frac{T}{2\vartheta}\langle(-\Delta)K_{\delta,x_0}, K_{\delta,x_0}\rangle$, which by the scaling properties is of order δ^{-2}. Physically, this can be interpreted as an increase in energy in X_{δ,x_0}^Δ under δ-localisation due to the Laplacian in the drift, while the energy from the space–time white noise remains unchanged. This is in fact the same phenomenon as the increasing signal-to-noise ratio for high Fourier modes in the spectral approach by Huebner and Rozovskii [19].

Proxy MLE. Consider $\hat{\vartheta}_\delta^P(x_0) = \hat{\vartheta}_\delta^P$ from (4.3). The only stochastic part is

$$\mathcal{I}_\delta^P := \delta^{-2} \int_0^T X_{\delta,x_0}(t)^2\, dt \tag{4.5}$$

in the denominator. In the general model (2.8) we shall see that $\mathcal{I}_\delta^P$ converges to $\vartheta(x_0)^{-1} T \Psi(K, K)$, compare also Remark 3.3 with $K = \Delta\tilde{K}$. Asking for consistency $\hat{\vartheta}_\delta^P(x_0) \to \vartheta(x_0)$ leads to requiring the identity $\|\nabla\tilde{K}\|_{L^2(\mathbb{R}^d)}^2 \|B_{0,x_0}^* K\|_{L^2(\mathbb{R}^d)}^2 = 2\|K\|_{L^2(\mathbb{R}^d)}^2 \Psi(K, K)$. This does not hold for any operator B_{0,x_0}. We therefore restrict to our main specification $B = M_\sigma$, for which the identity holds by (3.3). In contrast to the augmented MLE, the proxy MLE works with the observation of X_{δ,x_0} alone, but asks for new structural assumptions on B and K. If they are not fulfilled, other likelihood approximations should be pursued. Compare also the suboptimal behaviour of $\hat{\vartheta}_\delta^P(x_0)$ for the kernel $K^{(2)}$ in the simulations of Section 6 below.

5. Main results.

5.1. *Results for the augmented MLE.* Recall the function Ψ from (3.2) and the error decomposition (4.4). We show first that the observed Fisher information and the bias, after rescaling, converge to deterministic quantities. The propositions are proved in Appendix A.1.

PROPOSITION 5.1. *Grant Assumptions 3.2 and 3.7(K; 2). Then for any $d \geq 1$ as $\delta \to 0$*

$$\delta^2 \mathbb{E}[\mathcal{I}_\delta^A] \to T\vartheta(x_0)^{-1}\Psi(\Delta K, \Delta K), \qquad \mathcal{I}_\delta^A / \mathbb{E}[\mathcal{I}_\delta^A] \xrightarrow{\mathbb{P}} 1.$$

PROPOSITION 5.2. *Grant Assumptions 3.2, 3.7(K; 2), and for $d = 1$ assume $\vartheta \in C^{1+\alpha'}(\overline{\Lambda})$ for $\alpha' > 1/2$ and $\int_{\mathbb{R}} K(x)\,dx = 0$. Then for $\delta \to 0$*

$$\delta^{-1}(\mathcal{I}_\delta^A)^{-1} R_\delta^A \xrightarrow{\mathbb{P}} \mu^A \quad \textit{with } \mu^A := (\Psi(\Delta K, \Delta K))^{-1}\Psi(\Delta K, \beta),$$

where $\beta(x) = \Delta(\langle \nabla\vartheta(x_0), x\rangle_{\mathbb{R}^d} K)(x) - \langle \nabla\vartheta(x_0) - a(x_0), \nabla K(x)\rangle_{\mathbb{R}^d}$, $x \in \mathbb{R}^d$.

From this it follows that the augmented MLE $\hat{\vartheta}_\delta^A(x_0)$ satisfies a central limit theorem with rate δ.

THEOREM 5.3. *Grant Assumptions 3.2, 3.7(K; 2), and for $d = 1$ assume $\vartheta \in C^{1+\alpha'}(\overline{\Lambda})$ for $\alpha' > 1/2$ and $\int_{\mathbb{R}} K(x)\,dx = 0$. Then for $\delta \to 0$*

$$\begin{aligned} &\delta^{-1}(\hat{\vartheta}_\delta^A(x_0) - \vartheta(x_0)) \xrightarrow{d} N(\mu^A, \vartheta(x_0)\Sigma^A), \\ &\textit{with } \Sigma^A = T^{-1}(\Psi(\Delta K, \Delta K))^{-1} \|B_{0,x_0}^* K\|_{L^2(\mathbb{R}^d)}^2, \end{aligned} \tag{5.1}$$

and with μ^A from Proposition 5.2.

PROOF. In terms of $Y_t^{(\delta)} := X_{\delta,x_0}^\Delta(t)/\mathbb{E}[\mathcal{I}_\delta^A]^{1/2}$ we obtain $M_\delta^A/\mathbb{E}[\mathcal{I}_\delta^A]^{1/2} = \int_0^T Y_t^{(\delta)}\,d\bar{W}(t)$, the quadratic variation of which satisfies $\int_0^T (Y_t^{(\delta)})^2\,dt = \mathcal{I}_\delta^A/\mathbb{E}[\mathcal{I}_\delta^A] \xrightarrow{\mathbb{P}} 1$ by Proposition 5.1. A standard continuous martingale CLT, for example, Kutoyants [23], Theorem 1.19, shows $M_\delta^A/\mathbb{E}[\mathcal{I}_\delta^A]^{1/2} \xrightarrow{d} N(0, 1)$. Moreover,

$$\|B^* K_{\delta,x_0}\| = \|B_{\delta,x_0}^* K\|_{L^2(\Lambda_{\delta,x_0})} \to \|B_{0,x_0}^* K\|_{L^2(\mathbb{R}^d)} \tag{5.2}$$

due to Assumption 3.2 and $\delta^{-1}(\mathcal{I}_\delta^A)^{-1} R_\delta^A \xrightarrow{\mathbb{P}} \mu^A$ by Proposition 5.2. We conclude by applying Slutsky's lemma. □

REMARKS 5.4.

(i) Both, bias and standard deviation of $\hat{\vartheta}_\delta^A(x_0)$, are of order δ. The asymptotic bias μ^A is independent of T, while the variance Σ^A decays in T.

(ii) B, $\nabla\vartheta$ and a appear in the limit only via the localized terms B_{0,x_0}^*, $\nabla\vartheta(x_0)$, $a(x_0)$, while b does not appear at all. This demonstrates again the universality property of local measurements, in the spirit of Theorem 3.6.

(iii) The estimator and thus also its asymptotic bias and variance are invariant under constant scaling factors in the kernel. In fact, using the scaling such that $\|K_{\delta,x_0}\| = \|K\|_{L^2(\mathbb{R}^d)}$ is arbitrary, but simplifies the analysis.

(iv) The additional assumptions for the convergence of the remaining bias R_δ^A in $d = 1$ allow for compensating the slower decay of the heat kernel compared to $d > 1$; cf. Lemma A.6(ii), (iii). This is not necessary for constant ϑ and in that case Theorem 5.3 holds without these assumptions.

(v) When we dispose of observations at different locations x, then we can estimate $\vartheta(x)$ pointwise at each location x. In the case of multiplicative covariance $B = M_\sigma$ it can be shown that estimators at different locations become asymptotically independent. The argument relies on a multivariate martingale difference CLT, using that at points $x_0, x_1 \in \Lambda$ the corresponding Brownian motions $\overline{W}_0, \overline{W}_1$ in (2.8) are independent whenever $\operatorname{supp}(K_{\delta,x_0}) \cap \operatorname{supp}(K_{\delta,x_1}) = \varnothing$.

From Proposition 5.2 we see that μ^A vanishes if $A_\vartheta = \vartheta\Delta + b$ for parametric $\vartheta > 0$. Another important situation where $\mu^A = 0$ is given next.

EXAMPLE 5.5 (Example 3.4(a) ctd.). Let $B = M_\sigma$ and recall the identities $\Psi(\Delta K, \Delta K) = \frac{\sigma(x_0)^2}{2}\|\nabla K\|^2_{L^2(\mathbb{R}^d)}$, $\Psi(\Delta K, \beta) = -\frac{\sigma(x_0)^2}{2}\langle K, \beta\rangle_{L^2(\mathbb{R}^d)}$ with β from Proposition 5.2. By Lemma A.3 with $z = K$ this means

$$\langle K, \beta\rangle_{L^2(\mathbb{R}^d)} = -\langle\langle\nabla\vartheta(x_0), x\rangle_{\mathbb{R}^d}, |\nabla K(x)|^2\rangle_{L^2(\mathbb{R}^d)},$$

and Theorem 5.3 yields

$$\mu^A = \frac{\int_{\mathbb{R}^d}\langle\nabla\vartheta(x_0), x\rangle_{\mathbb{R}^d}|\nabla K(x)|^2\,dx}{\|\nabla K\|^2_{L^2(\mathbb{R}^d)}}, \qquad \Sigma^A = \frac{2\|K\|^2_{L^2(\mathbb{R}^d)}}{T\|\nabla K\|^2_{L^2(\mathbb{R}^d)}}.$$

In particular, if ∇K is symmetric in the sense $|\nabla K(-x)| = |\nabla K(x)|$, $x \in \mathbb{R}^d$, then the asymptotic bias vanishes:

$$\delta^{-1}(\hat{\vartheta}^A_\delta(x_0) - \vartheta(x_0)) \xrightarrow{d} N\left(0, \frac{2\vartheta(x_0)\|K\|^2_{L^2(\mathbb{R}^d)}}{T\|\nabla K\|^2_{L^2(\mathbb{R}^d)}}\right).$$

The rougher K is, the smaller is the asymptotic variance, which bears some similarity with deconvolution problems.

If the asymptotic bias μ^A vanishes, we can construct a simple confidence interval in terms of the augmented MLE. Note that in the setting of Example 5.5, $\Sigma^A = 2T^{-1}\|K\|^2_{L^2(\mathbb{R}^d)} \times \|\nabla K\|^{-2}_{L^2(\mathbb{R}^d)}$ is easily accessible.

COROLLARY 5.6. *Assume the setting of Theorem 5.3, $\mu^A = 0$ and let $\overline{\alpha} \in (0, 1)$. Then the confidence interval for $\vartheta(x_0)$*

$$I^A_{1-\overline{\alpha}} = [\hat{\vartheta}^A_\delta(x_0) - \delta(\hat{\vartheta}^A_\delta(x_0)\Sigma^A)^{1/2}q_{1-\overline{\alpha}/2}, \hat{\vartheta}^A_\delta(x_0) + \delta(\hat{\vartheta}^A_\delta(x_0)\Sigma^A)^{1/2}q_{1-\overline{\alpha}/2}],$$

with the standard normal $(1-\overline{\alpha}/2)$-quantile $q_{1-\overline{\alpha}/2}$, has asymptotic coverage $1-\overline{\alpha}$ for $\delta \to 0$.

PROOF. By Theorem 5.3 and Slutsky's lemma applied for $\hat{\vartheta}^A_\delta(x_0) \xrightarrow{\mathbb{P}} \vartheta(x_0)$, we have

$$\delta^{-1}(\hat{\vartheta}^A_\delta(x_0)\Sigma^A)^{-1/2}(\hat{\vartheta}^A_\delta(x_0) - \vartheta(x_0)) \xrightarrow{d} N(0, 1), \quad \delta \to 0,$$

noting $\mu^A = 0$. This yields $\mathbb{P}(\vartheta(x_0) \in I^A_{1-\overline{\alpha}}) \to 1 - \overline{\alpha}$. $\square$

5.2. *Results for the proxy MLE.* In the setting described in Section 4.2 we obtain a CLT for the quadratic functional $\mathcal{I}_\delta^P$. The proof uses very precise asymptotic moment calculations and the fourth moment theorem in Wiener chaos. Fundamental for this analysis is that $X_{\delta,x_0}(t)$ and $X_{\delta,x_0}(s)$ quickly decorrelate as $\delta^{-2}|t-s| \to \infty$, which is also predicted by the scaling limit in Corollary 3.6. Note that this method of proof might also cover time-discrete observations of X_{δ,x_0} if the sampling frequency increases sufficiently fast as $\delta \to 0$, but this is not pursued here.

The next assumption gathers the conditions required for the analysis of the proxy MLE.

ASSUMPTION 5.7. Let $K = \Delta \tilde{K}$ for $\tilde{K} \in H^4(\mathbb{R}^d)$ with compact support and let $B = M_\sigma$ with $\sigma \in C^1(\mathbb{R}^d)$. Grant Assumption 3.7($\tilde{K}$; 3), and for $d = 1$ assume $\vartheta \in C^{1+\alpha'}(\overline{\Lambda})$ for $\alpha' > 1/2$ and $\int_{\mathbb{R}} \tilde{K}(x)\,dx = 0$.

The following proposition is proved in Appendix A.1.

PROPOSITION 5.8. *Grant Assumption 5.7. Then for $\delta \to 0$:*

$$\delta^{-1}\big(\mathcal{I}_\delta^P - \vartheta(x_0)^{-1} C_{T,K}\big) \xrightarrow{d} N\big(-\vartheta^{-2}(x_0) C_{T,K}\mu_1^P, \vartheta^{-3}(x_0) C_{T,K}^2 \Sigma^P\big),$$

$$\text{with } C_{T,K} = \frac{T}{2}\sigma^2(x_0)\|\nabla\tilde{K}\|_{L^2(\mathbb{R}^d)}^2,$$

$$\mu_1^P = -\frac{\vartheta^2(x_0)}{\sigma^2(x_0)}\|\nabla\tilde{K}\|_{L^2(\mathbb{R}^d)}^{-2}\left\langle\left\langle\nabla\left(\frac{\sigma^2}{\vartheta}\right)(x_0), x\right\rangle_{\mathbb{R}^d}, |\nabla\tilde{K}|^2\right\rangle_{L^2(\mathbb{R}^d)},$$

$$\Sigma^P = \frac{4}{T}\|\nabla\tilde{K}\|_{L^2(\mathbb{R}^d)}^{-4}\int_0^\infty \|\nabla e^{(s/2)\Delta}\tilde{K}\|_{L^2(\mathbb{R}^d)}^4\,ds.$$

This yields asymptotic normality for the proxy MLE $\hat{\vartheta}_\delta^P(x_0)$.

THEOREM 5.9. *Grant Assumption 5.7. Then for $\delta \to 0$:*

$$\delta^{-1}\big(\hat{\vartheta}_\delta^P(x_0) - \vartheta(x_0)\big) \xrightarrow{d} N\big(\mu_1^P + \mu_2^P, \vartheta(x_0)\Sigma^P\big),$$

$$\text{with } \mu_2^P = \frac{\vartheta(x_0)}{\sigma^2(x_0)}\|K\|_{L^2(\mathbb{R}^d)}^{-2}\langle\langle\nabla\sigma^2(x_0), x\rangle_{\mathbb{R}^d}, |K|^2\rangle_{L^2(\mathbb{R}^d)},$$

and with μ_1^P, Σ^P from Proposition 5.8.

PROOF. Recall the quadratic variation $\langle X_{\delta,x_0}\rangle_T = T\|\sigma K_{\delta,x_0}\|^2$, the constant $C_{T,K}$ from Proposition 5.8 and set

$$D_{T,K} = \frac{T}{2}\|\nabla\tilde{K}\|_{L^2(\mathbb{R}^d)}^2\|K\|_{L^2(\mathbb{R}^d)}^{-2}.$$

Write $\hat{\vartheta}_\delta^P(x_0) = (\mathcal{I}_\delta^P)^{-1} D_{T,K}\|\sigma K_{\delta,x_0}\|^2$ and decompose

$$\begin{aligned}\delta^{-1}\big(\hat{\vartheta}_\delta^P(x_0) - \vartheta(x_0)\big) &= (\mathcal{I}_\delta^P)^{-1} D_{T,K}\delta^{-1}\big(\|\sigma K_{\delta,x_0}\|^2 - D_{T,K}^{-1}C_{T,K}\big)\\ &\quad + C_{T,K}\delta^{-1}\big((\mathcal{I}_\delta^P)^{-1} - \vartheta(x_0)C_{T,K}^{-1}\big).\end{aligned}$$

From the compact support of K we infer for $\delta \to 0$

$$\begin{aligned}\frac{\|\sigma K_{\delta,x_0}\|^2 - D_{T,K}^{-1}C_{T,K}}{\delta} &= \left\langle\frac{\sigma^2(x_0+\delta\cdot) - \sigma^2(x_0)}{\delta}K, K\right\rangle_{L^2(\mathbb{R}^d)}\\ &\to \langle\langle\nabla\sigma^2(x_0), x\rangle_{\mathbb{R}^d}, |K|^2\rangle_{L^2(\mathbb{R}^d)}.\end{aligned}$$

Proposition 5.8 and the delta method (Ferguson [12], Theorem 7) give

$$\delta^{-1}\big((\mathcal{I}_\delta^P)^{-1} - \vartheta(x_0)C_{T,K}^{-1}\big) \xrightarrow{d} N\big(C_{T,K}^{-1}\mu_1^P, \vartheta(x_0)C_{T,K}^{-2}\Sigma^P\big),$$

and, in particular, $(\mathcal{I}_\delta^P)^{-1} \xrightarrow{\mathbb{P}} \vartheta(x_0)C_{T,K}^{-1}$. The theorem follows from Slutsky's lemma. □

The dependencies on δ, T, ϑ, K in the CLT are similar as for $\hat{\vartheta}_\delta^A(x_0)$. It is interesting to note that the asymptotic bias depends locally at x_0 on σ^2, ϑ and their gradients, while a, b do not appear at all. The asymptotic bias vanishes when $\frac{\sigma^2}{\vartheta}$ and σ^2 are constant, but also similar to Example 5.5 if $|\nabla\tilde{K}(-x)| = |\nabla\tilde{K}(x)|$, $|K(-x)| = |K(x)|$, $x \in \mathbb{R}^d$. As for the augmented MLE in Corollary 5.6, we obtain an asymptotic $(1-\overline{\alpha})$-confidence interval.

COROLLARY 5.10. *Grant Assumption 5.7, suppose* $\mu_1^P + \mu_2^P = 0$ *and let* $\overline{\alpha} \in (0,1)$. *Then the confidence interval for* $\vartheta(x_0)$

$$I_{1-\overline{\alpha}}^P = \big[\hat{\vartheta}_\delta^P(x_0) - \delta\big(\hat{\vartheta}_\delta^P(x_0)\Sigma^P\big)^{1/2} q_{1-\overline{\alpha}/2}, \hat{\vartheta}_\delta^P(x_0) + \delta\big(\hat{\vartheta}_\delta^P(x_0)\Sigma^P\big)^{1/2} q_{1-\overline{\alpha}/2}\big],$$

with the standard normal $(1-\overline{\alpha}/2)$-*quantile* $q_{1-\overline{\alpha}/2}$, *has asymptotic coverage* $1-\overline{\alpha}$ *for* $\delta \to 0$.

Let us finally compare the variance factor Σ^P to Σ^A from Theorem 5.3.

LEMMA 5.11. *Under Assumption 5.7 the asymptotic variances of* $\hat{\vartheta}_\delta^P$ *and* $\hat{\vartheta}_\delta^A$ *always satisfy* $\vartheta(x_0)\Sigma^P \geq \vartheta(x_0)\Sigma^A$.

PROOF. Using the tensor products $\Delta \otimes \Delta$, $f \otimes f$ and $\Delta \oplus \Delta := I \otimes \Delta + \Delta \otimes I$, we can write for $f \in L^2(\mathbb{R}^d)$, identifying $L^2(\mathbb{R}^d) \otimes L^2(\mathbb{R}^d) = L^2(\mathbb{R}^{2d})$,

$$\begin{aligned}\int_0^\infty \|e^{(s/2)\Delta} f\|_{L^2(\mathbb{R}^d)}^4 \, ds &= \int_0^\infty \langle (e^{s\Delta} \otimes e^{s\Delta})(f \otimes f), f \otimes f\rangle_{L^2(\mathbb{R}^{2d})} \, ds \\ &= \int_0^\infty \langle e^{s(\Delta \oplus \Delta)}(f \otimes f), f \otimes f\rangle_{L^2(\mathbb{R}^{2d})} \, ds \\ &= \|(-\Delta \oplus \Delta)^{-1/2}(f \otimes f)\|_{L^2(\mathbb{R}^{2d})}^2,\end{aligned}$$

provided the last norm is finite, for example, if $f = (-\Delta)^{1/2}\tilde{K}$. With this f we conclude via two duality arguments, using $\Delta\tilde{K} = K$:

$$\begin{aligned}\Sigma^P &= \frac{4}{T} \frac{\|(-\Delta \oplus \Delta)^{-1/2}(f \otimes f)\|_{L^2(\mathbb{R}^{2d})}^2}{\|f \otimes f\|_{L^2(\mathbb{R}^{2d})}^2} \geq \frac{4}{T} \frac{\|f \otimes f\|_{L^2(\mathbb{R}^{2d})}^2}{\|(-\Delta \oplus \Delta)^{1/2}(f \otimes f)\|_{L^2(\mathbb{R}^{2d})}^2} \\ &= \frac{4}{T} \frac{\|f \otimes f\|_{L^2(\mathbb{R}^{2d})}^2}{\langle(-\Delta \oplus \Delta)(f \otimes f), f \otimes f\rangle_{L^2(\mathbb{R}^{2d})}} = \frac{2}{T} \frac{\|(-\Delta)^{1/2}\tilde{K}\|_{L^2(\mathbb{R}^d)}^2}{\|K\|_{L^2(\mathbb{R}^d)}^2} \\ &\geq \frac{2}{T} \frac{\|K\|_{L^2(\mathbb{R}^d)}^2}{\|(-\Delta)^{1/2}K\|_{L^2(\mathbb{R}^d)}^2} = \Sigma^A,\end{aligned}$$

which yields the result. □

Consequently, the proxy MLE has a larger variance than the augmented MLE, but the loss in precision is not severe if K has a well concentrated Fourier spectrum (consider Δ as a multiplier in the Fourier domain).

Let us point out that in the one-dimensional parametric case with $A_\vartheta = \vartheta\partial_{xx} + a\partial_x + b$ and $B = M_\sigma$ for constant σ, Bibinger and Trabs [4] construct least-squares estimators for $\sigma^2/\sqrt{\vartheta}$ and a/ϑ from discrete high frequency observations in time at two spatial points x_1, x_2. Compared to this, the proxy MLE uses spatial averages of the solution in infinitesimally small neighbourhoods, observed continuously in time, to estimate ϑ itself, without having to know a or σ. A similar phenomenon has been observed by Cialenco and Huang [7] for discrete observations when $a = 0$, but they achieve only consistent estimation of ϑ and σ. A more profound comparison of both approaches would be highly desirable.

5.3. *Rate optimality.* Let us address the question of optimality of the above estimators by providing a minimax lower bound. For minimax lower bounds it suffices to consider a subclass of the original model and we thus assume here that X_{δ,x_0} is observed with $A_\vartheta = \vartheta\Delta$, $B = I$ and a stationary initial condition X_{δ,x_0}. Then the following result establishes that the rate of convergence δ is optimal and gives some lower bound for the dependence on T, ϑ and K.

PROPOSITION 5.12. *Assume $A_\vartheta = \vartheta\Delta$, $\vartheta > 0$, $B = I$, $K \in H^1(\mathbb{R}^d)$ with compact support and that X_{δ,x_0} is stationary. For $\vartheta_0 > 0$ and $\delta \to 0$ we have the asymptotic local lower bound for the root mean squared error*

$$\inf_{\hat\vartheta} \sup_{\vartheta\in[\vartheta_0,\vartheta_0(1+\delta)]} \mathbb{E}_\vartheta\big[(\hat\vartheta-\vartheta)^2\big]^{1/2} \geq \bar c\bigg(\frac{(\vartheta_0\wedge 1)\|(I-\Delta)^{-1}K\|^2_{L^2(\mathbb{R}^d)}}{\sqrt{T}(\|K\|^2_{L^2(\mathbb{R}^d)} + \|\nabla K\|^2_{L^2(\mathbb{R}^d)})}\bigg)\delta,$$

where Δ is the Laplace operator on $L^2(\mathbb{R}^d)$, $\bar c > 0$ is some constant and the infimum is taken over all estimators $\hat\vartheta$ based on observing X_{δ,x_0}.

PROOF. The autocovariance function of the stationary process $(\delta^{-1}X_{\delta,x_0}(\delta^2 t))_{t\in\mathbb{R}}$ is given by

$$\begin{aligned} c_{\vartheta,\delta}(t) &:= \delta^{-2}\mathbb{E}\big[X_{\delta,x_0}(\delta^2 t)X_{\delta,x_0}(0)\big] \\ &= \delta^{-2}\int_{-\infty}^0 \langle S_\vartheta(\delta^2|t|-s)K_{\delta,x_0}, S_\vartheta(-s)K_{\delta,x_0}\rangle\, ds \\ &= \langle(-2A_{\vartheta,\delta,x_0})^{-1}S_{\vartheta,\delta,x_0}(|t|)K, K\rangle_{L^2(\Lambda_{\delta,x_0})}, \end{aligned}$$

using the scaling in Lemma 3.1 and $\frac{d}{ds}S_{\vartheta,\delta,x_0}(s) = A_{\vartheta,\delta,x_0}S_{\vartheta,\delta,x_0}(s)$ in the last line. The covariance operator for $\delta^{-1}X_{\delta,x_0}(\delta^2\cdot)$ on $L^2(\mathbb{R})$ is obtained by convolution:

$$C_{\vartheta,\delta}f(t) = (c_{\vartheta,\delta} * f)(t), \quad t\in\mathbb{R}. \tag{5.3}$$

The squared Hellinger distance $H^2(\vartheta,\vartheta_0)$ between two equivalent centered Gaussian measures can be bounded in terms of the Hilbert–Schmidt norm of the covariance operators; see, for example, the proof of the Feldman–Hajek Theorem in Da Prato and Zabczyk [9], Theorem 2.25. For the laws of $(\delta^{-1}X_{\delta,x_0}(\delta^2 t))_{t\in[0,T\delta^{-2}]}$ under ϑ_0 and ϑ we can thus bound the corresponding Hellinger distance via

$$H^2(\vartheta,\vartheta_0) \leq \|C^{-1}_{\vartheta_0,\delta}(C_{\vartheta,\delta} - C_{\vartheta_0,\delta})\|^2_{HS(L^2([0,T\delta^{-2}]))}.$$

Since the Hellinger distance is invariant under bi-measurable bijective transformations, $H(\vartheta,\vartheta_0)$ denotes equally the Hellinger distance between the observation laws of $(X_{\delta,x_0}(t))_{t\in[0,T]}$.

Let now $\vartheta_\delta = \vartheta_0 + c\delta$ for some small $c > 0$, which we choose below, and assume that we can show $H^2(\vartheta_\delta, \vartheta_0) \le 1$ for sufficiently small δ. Then we obtain from the general lower bound scheme in Tsybakov [36], using his Theorem 2.2(ii) and (2.9), that

$$\inf_{\hat{\vartheta}} \max_{\vartheta \in \{\vartheta_0, \vartheta_\delta\}} \mathbb{E}_\vartheta\big[(\hat{\vartheta} - \vartheta)^2\big] \ge \frac{2-\sqrt{3}}{4}(\vartheta_\delta - \vartheta_0)^2 = \frac{2-\sqrt{3}}{4} c^2\delta^2. \tag{5.4}$$

From this we will obtain the claimed lower bound.

In order to show $H^2(\vartheta_\delta, \vartheta_0) \le 1$, denote by $\iota_1 : H^1([0, T\delta^{-2}]) \to L^2([0, T\delta^{-2}])$ the Sobolev embedding operator. It is known from Maurin's Theorem (see, e.g., the proof of Adams and Fournier [1], Theorem 6.61) that ι_1 is Hilbert–Schmidt with

$$\|\iota_1\|^2_{HS(H^1([0,T\delta^{-2}]), L^2([0,T\delta^{-2}]))} \le K_{HS} T\delta^{-2}$$

for some constant $K_{HS} > 0$. By Hilbert–Schmidt norm calculus (in particular, $\|AB\|_{HS(H_2,H_3)} \le \|A\|_{HS(H_1,H_3)} \|B\|_{H_2 \to H_1}$ with obvious notation for the Hilbert–Schmidt and operator norms between Hilbert spaces H_1, H_2, H_3), the implicit restriction of the covariance operators and by the covariance bound of Lemma A.1 below we conclude for $\vartheta_\delta > \vartheta_0$ that

$$\begin{aligned}
H^2(\vartheta_\delta, \vartheta_0) &\le \big\|C^{-1}_{\vartheta_0,\delta}(C_{\vartheta_\delta,\delta} - C_{\vartheta_0,\delta})\big\|^2_{HS(L^2([0,T\delta^{-2}]))} \\
&\le \|\iota_1\|^2_{HS(H^1([0,T\delta^{-2}]),L^2([0,T\delta^{-2}]))} \big\|C^{-1}_{\vartheta_0,\delta}(C_{\vartheta_\delta,\delta} - C_{\vartheta_0,\delta})\big\|^2_{L^2([0,T\delta^{-2}]) \to H^1([0,T\delta^{-2}])} \\
&\le K_{HS} T\delta^{-2} \big\|C^{-1}_{\vartheta_0,\delta}(C_{\vartheta_\delta,\delta} - C_{\vartheta_0,\delta})\big\|^2_{L^2(\mathbb{R}) \to H^1(\mathbb{R})} \\
&\le K_{HS} T \left(\vartheta_0^{-2} + \vartheta_0^{-1} \frac{\|K\|^2_{L^2(\mathbb{R}^d)} + \|\nabla K\|^2_{L^2(\mathbb{R}^d)}}{\|(I - A_{1,\delta,x_0})^{-1}K\|^2_{L^2(\Lambda_{\delta,x_0})}}\right)^2 \frac{(\vartheta_\delta^2 - \vartheta_0^2)^2}{\delta^2}.
\end{aligned}$$

Hence, $H^2(\vartheta_\delta, \vartheta) \le 1$ holds whenever

$$\vartheta_\delta^2 - \vartheta_0^2 \le \frac{\vartheta_0^2}{\sqrt{K_{HS}T}} \left(1 + \vartheta_0 \frac{\|K\|^2_{L^2(\mathbb{R}^d)} + \|\nabla K\|^2_{L^2(\mathbb{R}^d)}}{\|(I - A_{1,\delta,x_0})^{-1}K\|^2_{L^2(\Lambda_{\delta,x_0})}}\right)^{-1} \delta.$$

Noting the convergence $\|(I - A_{1,\delta,x_0})^{-1}K\|_{L^2(\Lambda_{\delta,x_0})} \to \|(I - \Delta)^{-1}K\|_{L^2(\mathbb{R}^d)}$ from Lemma A.1 below, we can thus find a sufficiently small constant $c' > 0$ such that, with

$$c = c'\vartheta_0 \frac{(1 \wedge \vartheta_0^{-1})\|(I - \Delta)^{-1}K\|^2_{L^2(\mathbb{R}^d)}}{\sqrt{T}(\|K\|^2_{L^2(\mathbb{R}^d)} + \|\nabla K\|^2_{L^2(\mathbb{R}^d)})},$$

(5.4) holds for $\vartheta_\delta = \vartheta_0 + c\delta$. This yields the result. □

6. A numerical example. In this section we briefly illustrate the main results from above with simulation results. Let $\Lambda = (0, 1)$, $T = 1$, and consider the stochastic heat equation

$$dX(t) = \Delta_\vartheta X(t)\,dt + dW(t)$$

with Dirichlet boundary conditions and with spatially varying diffusivity ϑ, which is smooth (true diffusivity in Figure 1 (center)). Assume that X_0 is zero, except for two equally high "peaks" at $x = 0.2$ and $x = 0.8$. The heat map for a typical realisation is presented in Figure 1 (left) and we see already qualitatively that the heat diffusion is higher for $x \le 1/2$.

An approximate solution $\tilde{X}(t_k, y_j) \approx X(t_k)(y_j)$ is obtained with respect to a regular time-space grid $\{(t_k, y_j) : t_k = k/N, y_j = j/M, k = 0, \ldots, N, j = 0, \ldots, M\}$ by a semi-implicit Euler scheme and a finite difference approximation of Δ_ϑ (Lord, Powell and Shardlow [26],

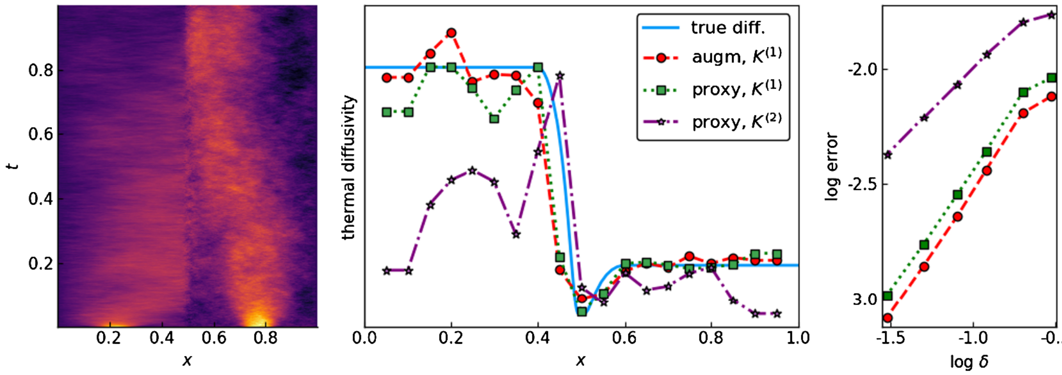

FIG. 1. *(left) heat map for a typical realisation of* $X(t,x)$; *(center) true* ϑ *compared to* $\hat{\vartheta}_\delta^A$ *and* $\hat{\vartheta}_\delta^P$ *at* $\delta = 0.12$ *with two different kernels*; *(right)* $\log_{10}$-$\log_{10}$ *plot of root mean squared estimation errors at* $x_0 = 0.6$ *for the estimators in the center.*

Section 10.5). Since the solution is tested against functions K_{δ,x_0} and $\Delta K_{\delta,x_0}$ with small support, M needs to be relatively large, while it is well known that accurate simulation requires $N \asymp M^2$, see Lord, Powell and Shardlow [26], p. 458. We therefore choose $M = 2000$, $N = 10^6$.

Consider the kernels $K^{(1)} = \varphi'''$, $K^{(2)} = \varphi'$ with a smooth bump function

$$\varphi(x) = \exp\left(-\frac{12}{1-x^2}\right), \quad x \in (-1, 1).$$

For $\delta \in \{0.03, 0.05, 0.08, 0.12, 0.2, 0.3\}$ and $x_0 \in (0, 1)$ on a regular grid we obtain approximate local measurements $\tilde{X}_{\delta,x_0}$, $\tilde{X}^\Delta_{\delta,x_0}$ for $K^{(1)}$ and $K^{(2)}$, respectively, from which the augmented MLE $\hat{\vartheta}_\delta^A(x_0)$ and the proxy MLE $\hat{\vartheta}_\delta^P(x_0)$ are computed. For x_0 near the boundary and $i = 1, 2$ set

$$K^{(i)}_{\delta,x_0} = \begin{cases} K^{(i)}_{\delta,\delta}, & x_0 < \delta, \\ K^{(i)}_{\delta,1-\delta}, & x_0 > 1-\delta. \end{cases}$$

Figure 1 (center) shows pointwise estimation results for $\vartheta(x_0)$ at $\delta = 0.12$ and for different x_0, while Figure 1 (right) presents a $\log_{10}$-$\log_{10}$ plot of root mean squared estimation errors at $x_0 = 0.6$ for $\delta \to 0$, obtained by 5000 Monte-Carlo runs.

Already at the relatively large resolution $\delta = 0.12$ both $\hat{\vartheta}_\delta^A(x_0)$ and $\hat{\vartheta}_\delta^P(x_0)$ perform surprisingly well. For $K^{(1)}$ both estimators are close together and achieve after a burn-in phase the convergence rate δ, as predicted by Theorems 5.3 and 5.9. Note that $K^{(1)} = \Delta\tilde{K}$ for $\tilde{K} = \varphi'$ and $\int_{\mathbb{R}} \tilde{K}(x)\,dx = 0$ such that the assumptions of Theorem 5.9 are satisfied. With respect to $K^{(2)}$ those assumptions are not met and indeed $\hat{\vartheta}_\delta^P(x_0)$ deviates considerably from $\vartheta(x_0)$, but still seems to be consistent with rate of convergence dropping to about $\delta^{3/4}$. Estimation by $\hat{\vartheta}_\delta^A(x_0)$ is unaffected by choosing $K^{(2)}$ instead of $K^{(1)}$ (not shown).

APPENDIX: PROOFS

For a better understanding we structure the appendix such that the proofs for the main theorems of Section 5 are given in Section A.1. Only afterwards, we provide the technical tools used for the main proofs. Section A.2 contains analytical results for rescaled semigroups and heat kernels, while Section A.3 assembles precise asymptotics for variance and covariance expressions.

From now on, without loss of generality replace Λ with $\Lambda - x_0$ and assume $x_0 = 0$. In particular, we estimate $\vartheta(0)$ and ease notation by removing the subindex x_0 and write $\Lambda_\delta =$

Λ_{δ,x_0}, $z_\delta = z_{\delta,x_0}$ and $X_\delta = X_{\delta,x_0}$. Unless stated otherwise, all limits are for $\delta \to 0$. C always denotes a generic positive constant, which may depend on T, if not made explicit otherwise, and changes from line to line. $A \lesssim B$ means $A \le CB$. For $z \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ define the norm

$$\|z\|_{L^1 \cap L^2(\mathbb{R}^d)} := \|z\|_{L^1(\mathbb{R}^d)} + \|z\|_{L^2(\mathbb{R}^d)},$$

and for z with partial derivatives up to second order in $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ set

$$\|z\|_{W^2_{1,2}(\mathbb{R}^d)} := \big\|z + |\nabla z| + \Delta z\big\|_{L^1 \cap L^2(\mathbb{R}^d)}.$$

We write throughout $\langle X(t), z\rangle = \langle \tilde{X}(t), z\rangle + \langle S_\vartheta(t)X_0, z\rangle$ for $z \in L^2(\Lambda)$ with $\langle \tilde{X}(t), z\rangle$ being defined as $\langle X(t), z\rangle$, but with $X_0 = 0$. Note that $\mathbb{E}[\langle \tilde{X}(t), z\rangle] = 0$ and $\mathbb{E}[\langle X(t), z\rangle] = \langle S_\vartheta(t)X_0, z\rangle$. Set also $\tilde{X}_\delta(t) = \langle \tilde{X}(t), K_\delta\rangle$, $\tilde{X}^\Delta_\delta(t) = \langle \tilde{X}(t), \Delta K_\delta\rangle$. We will use frequently, without explicit mention, that $\Delta K_\delta = \delta^{-2}(\Delta K)_\delta$ by Lemma 3.1.

A.1. Proofs for Section 5.

PROOF OF PROPOSITION 5.1. We show the result first for $\tilde{\mathcal{I}}^A_\delta = \int_0^T \tilde{X}^\Delta_\delta(t)^2\,dt$. Propositions A.8(ii) and A.9(ii) below with $w^{(\delta)} = \Delta K$ yield

$$\mathbb{E}\big[\delta^2 \tilde{\mathcal{I}}^A_\delta\big] = \delta^2 \int_0^T \operatorname{Var}\big(\tilde{X}^\Delta_\delta(t)\big)\,dt \to T\vartheta(0)^{-1}\Psi(\Delta K, \Delta K), \quad \operatorname{Var}\big(\delta^2 \tilde{\mathcal{I}}^A_\delta\big) \to 0.$$

In particular, $\operatorname{Var}(\tilde{\mathcal{I}}^A_\delta)/\mathbb{E}[\tilde{\mathcal{I}}^A_\delta]^2 \xrightarrow{\mathbb{P}} 0$ and thus $\tilde{\mathcal{I}}^A_\delta/\mathbb{E}[\tilde{\mathcal{I}}^A_\delta] \xrightarrow{\mathbb{P}} 1$. To finish the proof, decompose

$$\delta^2 \mathcal{I}^A_\delta = \delta^2 \tilde{\mathcal{I}}^A_\delta + \int_0^T \delta^2 \mathbb{E}\big[X^\Delta_\delta(t)\big]^2\,dt + 2\int_0^T \delta^2 \tilde{X}^\Delta_\delta(t)\mathbb{E}\big[X^\Delta_\delta(t)\big]\,dt. \tag{A.1}$$

Assumption 3.7(K; 2) gives

$$\int_0^T \delta^2 \mathbb{E}\big[X^\Delta_\delta(t)\big]^2\,dt = \delta^{-2}\int_0^T \big\langle S_\vartheta(t)X_0, (\Delta K)_\delta\big\rangle^2\,dt \to 0. \tag{A.2}$$

By the Cauchy–Schwarz inequality, the cross-term in (A.1) is therefore also negligible and the result follows. $\square$

PROOF OF PROPOSITION 5.2. Define $\tilde{R}^A_\delta$ as R^A_δ, but with respect to $\tilde{X}(\cdot)$. In terms of $\beta^{(\delta)} := \delta^{-1}(A^*_{\vartheta,\delta} - \vartheta(0)\Delta)K$ we have $\delta R^A_\delta = \int_0^T X^\Delta_\delta(t)\langle X(t), \beta^{(\delta)}_\delta\rangle\,dt$ and $\delta\tilde{R}^A_\delta = \int_0^T \tilde{X}^\Delta_\delta(t)\langle \tilde{X}(t), \beta^{(\delta)}_\delta\rangle\,dt$. $\beta^{(\delta)}$ and β correspond to $v^{(\delta)}$ and v from Lemma A.5 below with $z = K$, and therefore $\beta^{(\delta)} \to \beta$ in $L^2(\mathbb{R}^d)$. Decompose $\delta R^A_\delta = \delta\tilde{R}^A_\delta + V_{1,\delta} + V_{2,\delta} + V_{3,\delta}$, where

$$V_{1,\delta} = \int_0^T \big\langle S_\vartheta(t)X_0, \delta^{-2}(\Delta K)_\delta\big\rangle\big\langle \tilde{X}(t), \beta^{(\delta)}_\delta\big\rangle\,dt,$$

$$V_{2,\delta} = \int_0^T \big\langle S_\vartheta(t)X_0, \delta^{-2}(\Delta K)_\delta\big\rangle\big\langle S_\vartheta(t)X_0, \beta^{(\delta)}_\delta\big\rangle\,dt,$$

$$V_{3,\delta} = \int_0^T \tilde{X}^\Delta_\delta(t)\big\langle S_\vartheta(t)X_0, \beta^{(\delta)}_\delta\big\rangle\,dt.$$

By the Cauchy-Schwarz inequality and Wick's formula (Janson [20], Theorem 1.28) we infer

$$\mathbb{E}\big[V^2_{3,\delta}\big] = \operatorname{Var}(V_{3,\delta}) \lesssim \bigg(\int_0^T\int_0^T \operatorname{Cov}\big(\tilde{X}^\Delta_\delta(t), \tilde{X}^\Delta_\delta(s)\big)^2 dt\,ds\bigg)^{1/2} \int_0^T \big\langle S_\vartheta(t)X_0, \beta^{(\delta)}_\delta\big\rangle^2\,dt$$

$$= \frac{1}{\sqrt{2}}\operatorname{Var}\bigg(\int_0^T \tilde{X}^\Delta_\delta(t)^2 dt\bigg)^{1/2}\int_0^T \big\langle S_\vartheta(t)X_0, \beta^{(\delta)}_\delta\big\rangle^2\,dt. \tag{A.3}$$

If $d \geq 2$, then $V_{3,\delta} \xrightarrow{\mathbb{P}} 0$ follows from Proposition A.9(iv) below with $w^{(\delta)} = \Delta K$, $m = K$ and from

$$\int_0^T \langle S_\vartheta(t) X_0, \beta_\delta^{(\delta)} \rangle^2 \, dt = O\big(\ell_{d,2}(\delta)\delta^2 \|\beta^{(\delta)}\|_{L^1 \cap L^2(\mathbb{R})}\big), \tag{A.4}$$

which holds by Lemma A.7(i) with $u = \beta^{(\delta)}$. When $d = 1$, $\vartheta \in C^{1+\alpha'}(\overline{\Lambda})$ for $\alpha' > 1/2$ and $\int_{\mathbb{R}} K(x)\, dx = 0$, it follows from Lemma A.5(ii) that there is a compactly supported $\tilde{\beta} \in H^2(\mathbb{R})$ with $\beta = \Delta\tilde{\beta}$, $\|\beta^{(\delta)} - \Delta\tilde{\beta}\|_{L^1 \cap L^2(\mathbb{R})} \leq C\delta^{\alpha'}$. This, Lemma A.7(i) with $u = \beta^{(\delta)} - \Delta\tilde{\beta}$ and Lemma A.7(ii) with $z = \tilde{\beta}$ show that the left hand side in (A.4) is of order $o(\delta^2)$, and we obtain again $V_{3,\delta} \xrightarrow{\mathbb{P}} 0$ from Proposition A.9(iv). $V_{2,\delta} \xrightarrow{\mathbb{P}} 0$ follows by Assumption 3.7(K; 2) and the just proved upper bounds for the left hand side in (A.4). For $V_{1,\delta} \xrightarrow{\mathbb{P}} 0$ we argue in the same way as for $V_{2,\delta}$, using (A.3) and Assumption 3.7(K; 2), after replacing $\tilde{X}_\delta^\Delta(t)$ by $\langle \tilde{X}(t), \beta_\delta^{(\delta)} \rangle$ and noting for $d \geq 2$ by Proposition A.9(iii) with $u^{(\delta)} = \beta^{(\delta)}$, $u = \beta$

$$\operatorname{Var}\Big(\int_0^T \langle \tilde{X}(t), \beta_\delta^{(\delta)} \rangle^2 \, dt\Big) = O\big(\ell_{d,2}(\delta)^2 \delta^4\big), \tag{A.5}$$

while for $d = 1$ we conclude by Proposition A.9(ii) with $w^{(\delta)} = \beta^{(\delta)}$, $m = \tilde{\beta}$. In all, $V_{1,\delta}, V_{2,\delta}, V_{3,\delta} \xrightarrow{\mathbb{P}} 0$ holds for all $d \geq 1$. By Proposition 5.1 it therefore suffices to show

$$\mathbb{E}[\delta \tilde{R}_\delta^A] \to T\vartheta(0)^{-1}\Psi(\Delta K, \beta), \qquad \operatorname{Var}(\delta \tilde{R}_\delta^A) \to 0.$$

The convergence of $\mathbb{E}[\delta \tilde{R}_\delta^A]$ follows for $d \geq 2$ from Proposition A.8(iii) below with $w^{(\delta)} = \Delta K$, $z = K$, $u^{(\delta)} = \beta^{(\delta)}$, $u = \beta$. For $d = 1$ and $\tilde{\beta}$ as above, by polarisation and Proposition A.8(ii), $\mathbb{E}[\delta \tilde{R}_\delta^A]$ converges to

$$\frac{T}{4\vartheta(0)}\big(\Psi\big(\Delta(K + \tilde{\beta}), \Delta(K + \tilde{\beta})\big) - \Psi\big(\Delta(K - \tilde{\beta}), \Delta(K - \tilde{\beta})\big)\big) = \frac{T}{\vartheta(0)}\Psi(\Delta K, \beta).$$

Next, $\operatorname{Var}(\delta \tilde{R}_\delta^A) = \operatorname{Var}(\int_0^T X_\delta^\Delta(t) \langle X(t), \beta_\delta^{(\delta)} \rangle \, dt) \to 0$ follows for $d \geq 2$ by Proposition A.9(i) below with $z = K$, $u^{(\delta)} = \beta^{(\delta)}$, $u = \beta$, while for $d = 1$, $\operatorname{Var}(\delta \tilde{R}_\delta^A) \to 0$ holds by polarisation, the basic inequality $\operatorname{Var}(Z + \tilde{Z}) \leq 2\operatorname{Var}(Z) + 2\operatorname{Var}(\tilde{Z})$ for two random variables $Z, \tilde{Z}$ and Proposition A.9(ii) applied separately to $w^{(\delta)} = \Delta K - \beta^{(\delta)}$, $m = K - \tilde{\beta}$ and $w^{(\delta)} = \Delta K + \beta^{(\delta)}$, $m = K + \tilde{\beta}$. □

PROOF OF PROPOSITION 5.8. Define $\tilde{\mathcal{I}}_\delta^P$ as $\mathcal{I}_\delta^P$, but with respect to $\tilde{X}(\cdot)$. By Assumption 3.7($\tilde{K}$; 3) and $K_\delta = (\Delta\tilde{K})_\delta$ we have $\delta^{-3}\int_0^T \langle S_\vartheta(t) X_0, K_\delta \rangle^2 \, dt \to 0$ whence $\delta^{-1}(\mathcal{I}_\delta^P - \tilde{\mathcal{I}}_\delta^P) \to 0$ follows by $\tilde{\mathcal{I}}_\delta^P = O_{\mathbb{P}}(\delta)$ and the Cauchy–Schwarz inequality.

It remains to prove the result for $\tilde{\mathcal{I}}_\delta^P$. Note that

$$\delta^{-1}\big(\tilde{\mathcal{I}}_\delta^P - \vartheta(0)^{-1} C_{T,K}\big) = Z_\delta + \delta^{-1}\Big(\delta^{-2}\int_0^T \operatorname{Var}\big(\tilde{X}_\delta(t)\big)\, dt - \vartheta(0)^{-1} C_{T,K}\Big) \tag{A.6}$$

with $Z_\delta := \delta^{-3}\int_0^T (\tilde{X}_\delta(t)^2 - \mathbb{E}[\tilde{X}_\delta(t)^2])\, dt$. $\tilde{X}_\delta$ is a centered Gaussian process and Z_δ is an element of the second Wiener chaos. By the fourth moment theorem (Nualart and Peccati [30], Theorem 1) it suffices to prove $\operatorname{Var}(Z_\delta) \to \Sigma$ and $\mathbb{E}[Z_\delta^4] \to 3\Sigma^2$ to conclude $Z_\delta \xrightarrow{d} N(0, \Sigma)$. Propositions A.9(iv) and A.16 below (with $w^{(\delta)} = \Delta\tilde{K}$) provide exactly these convergences with

$$\Sigma = \frac{4T}{\vartheta^3(0)}\int_0^\infty \Psi(e^{s\Delta}\Delta\tilde{K}, \Delta\tilde{K})^2 \, ds = \frac{T\sigma^4(0)}{\vartheta^3(0)}\int_0^\infty \|\nabla e^{(s/2)\Delta}\tilde{K}\|_{L^2(\mathbb{R}^d)}^4 \, ds,$$

where the last identity is (3.3). The claim follows from applying Proposition A.10 below with $z = \tilde{K}$ to the second term in (A.6) and Slutsky's lemma. □

LEMMA A.1. *Assume the setting of Proposition 5.12 and recall the operator $C_{\vartheta,\delta}$ from (5.3). We have for $\vartheta > \vartheta_0 > 0$*

$$\|C_{\vartheta_0,\delta}^{-1}(C_{\vartheta,\delta} - C_{\vartheta_0,\delta})\|_{L^2(\mathbb{R})\to H^1(\mathbb{R})} \leq \frac{\vartheta^2 - \vartheta_0^2}{\vartheta_0^2}\left(1 + \vartheta_0 \frac{\|K\|_{L^2(\mathbb{R}^d)}^2 + \|\nabla K\|_{L^2(\mathbb{R}^d)}^2}{\|(I - A_{1,\delta})^{-1}K\|_{L^2(\Lambda_\delta)}^2}\right).$$

Moreover, we have $\|(I - A_{1,\delta})^{-1}K\|_{L^2(\Lambda_\delta)} \to \|(I - \Delta)^{-1}K\|_{L^2(\mathbb{R}^d)}$ for $\delta \to 0$, where Δ is the Laplace operator on $L^2(\mathbb{R}^d)$.

PROOF. For simplicity write in the following proof $\vartheta\Delta$ and $e^{\vartheta t\Delta}$ instead of $A_{\vartheta,\delta}$ and $S_{\vartheta,\delta}(t)$. In the Fourier domain, the convolution operator $C_{\vartheta,\delta}$ is given by

$$\begin{aligned}
\mathcal{F}c_{\vartheta,\delta}(w) &= \int_0^\infty \langle(-2\vartheta\Delta)^{-1}e^{\vartheta t\Delta}K, K\rangle_{L^2(\Lambda_\delta)}(e^{iwt} + e^{-iwt})\,dt \\
&= \left\langle(-2\vartheta\Delta)^{-1}\int_0^\infty (e^{t(\vartheta\Delta+iwI)} + e^{t(\vartheta\Delta-iwI)})K\,dt, K\right\rangle_{L^2(\Lambda_\delta)} \\
&= \langle(-2\vartheta\Delta)^{-1}(-(\vartheta\Delta + iwI)^{-1} - (\vartheta\Delta - iwI)^{-1})K, K\rangle_{L^2(\Lambda_\delta)} \\
&= \langle(\vartheta^2\Delta^2 + w^2 I)^{-1}K, K\rangle_{L^2(\Lambda_\delta)}.
\end{aligned}$$

The operator $C_{\vartheta_0,\delta}^{-1}(C_{\vartheta,\delta} - C_{\vartheta_0,\delta})$ is expressed in the Fourier domain by multiplication with

$$\frac{\mathcal{F}c_{\vartheta,\delta}(w) - \mathcal{F}c_{\vartheta_0,\delta}(w)}{\mathcal{F}c_{\vartheta_0,\delta}(w)} = (\vartheta^2 - \vartheta_0^2)\frac{\langle(\vartheta^2\Delta^2 + w^2 I)^{-1}\Delta^2(\vartheta_0^2\Delta^2 + w^2 I)^{-1}K, K\rangle_{L^2(\Lambda_\delta)}}{\langle(\vartheta_0^2\Delta^2 + w^2 I)^{-1}K, K\rangle_{L^2(\Lambda_\delta)}}.$$

Using the description of $H^1(\mathbb{R})$ in the Fourier domain and functional calculus for the Laplacian Δ on $L^2(\Lambda_\delta)$ yields therefore for $\vartheta > \vartheta_0$ that

$$\begin{aligned}
&\|C_{\vartheta_0,\delta}^{-1}(C_{\vartheta,\delta} - C_{\vartheta_0,\delta})\|_{L^2(\mathbb{R})\to H^1(\mathbb{R})} \\
&\quad= \sup_{w\in\mathbb{R}}\left|(1 + w^2)^{1/2}\frac{\mathcal{F}c_{\vartheta,\delta}(w) - \mathcal{F}c_{\vartheta_0,\delta}(w)}{\mathcal{F}c_{\vartheta_0,\delta}(w)}\right| \\
&\quad\leq (\vartheta^2 - \vartheta_0^2)\sup_{w'\in\mathbb{R}}\left|(1 + (\vartheta_0 w')^2)^{1/2}\frac{\langle\Delta^2(\vartheta_0^2\Delta^2 + \vartheta_0^2(w')^2 I)^{-2}K, K\rangle_{L^2(\Lambda_\delta)}}{\langle(\vartheta_0^2\Delta^2 + \vartheta_0^2(w')^2 I)^{-1}K, K\rangle_{L^2(\Lambda_\delta)}}\right| \\
&\quad= \frac{\vartheta^2 - \vartheta_0^2}{\vartheta_0^2}\sup_{w\in\mathbb{R}}\left|(1 + (\vartheta_0 w)^2)^{1/2}\frac{\|w^{-1}\Delta(w^{-2}\Delta^2 + I)^{-1}K\|_{L^2(\Lambda_\delta)}^2}{\|(w^{-2}\Delta^2 + I)^{-1/2}K\|_{L^2(\Lambda_\delta)}^2}\right| \\
&\quad\leq \frac{\vartheta^2 - \vartheta_0^2}{\vartheta_0^2}\left(1 + \vartheta_0\left(1 \vee \frac{\sup_{w>1}\|w^{-1/2}\Delta(w^{-2}\Delta^2 + I)^{-1}K\|_{L^2(\Lambda_\delta)}^2}{\inf_{w>1}\|(w^{-2}\Delta^2 + I)^{-1/2}K\|_{L^2(\Lambda_\delta)}^2}\right)\right) \\
&\quad\leq \frac{\vartheta^2 - \vartheta_0^2}{\vartheta_0^2}\left(1 + \vartheta_0\left(1 \vee \frac{\|(-\Delta)^{1/2}K\|_{L^2(\Lambda_\delta)}^2}{\|(I + \Delta^2)^{-1/2}K\|_{L^2(\Lambda_\delta)}^2}\right)\right),
\end{aligned}$$

where we used in the last line $w^{-1/2}\lambda(1 + w^{-2}\lambda^2)^{-1} \leq \lambda^{1/2}$ for all $\lambda, w > 0$. For this and similar arguments note that by spectral calculus with a self-adjoint operator A, for example, $-\Delta$, we have $\|f(A)K\| \leq \|g(A)K\|$ whenever $|f| \leq |g|$ for bounded f, g on the spectrum of A. Since $\langle -\Delta K, K\rangle_{L^2(\Lambda_\delta)} = \|\nabla K\|^2_{L^2(\mathbb{R}^d)}$, the numerator is independent of δ. For the denominator write again $A_{1,\delta} = \Delta$ and note similarly $(I + A_{1,\delta}^2)^{-1/2} \geq (I - A_{1,\delta})^{-1}$, where we have explicitly (cf. Pazy [32], Chapter 2.6)

$$(I - A_{1,\delta})^{-1} K = \int_0^\infty e^{-t} S_{1,\delta}(t) K\, dt.$$

Proposition 3.5 yields then first, approximating K by continuous functions, that $\|S_{1,\delta}(t)K\|_{L^2(\Lambda_\delta)} \lesssim \|K\|_{L^2(\mathbb{R}^d)}$ uniformly in δ, and second, the convergence

$$\|(I - A_{1,\delta})^{-1} K\|_{L^2(\Lambda_\delta)} \to \|(I - \Delta)^{-1} K\|_{L^2(\mathbb{R}^d)}, \quad \delta \to 0. \qquad \square$$

A.2. Analytical results. Recall that the solution of the heat equation $\frac{d}{dt} u(t) = \lambda \Delta u(t)$, $\lambda > 0$, on $\mathbb{R}^d$ with initial value $w \in L^2(\mathbb{R}^d)$ is given by the convolution

$$u(t) = e^{\lambda t \Delta} w = q_{\lambda t} * w, \tag{A.7}$$

with the heat kernel $q_t(x) = (4\pi t)^{-d/2} \exp(-|x|^2/(4t))$, $x \in \mathbb{R}^d$.

LEMMA A.2. *We have for $u \in L^2(\mathbb{R}^d)$, $t > 0$:*

(i) $\|e^{t\Delta} u\|_{L^2(\mathbb{R}^d)} \lesssim (1 \wedge t^{-d/4}) \|u\|_{L^1 \cap L^2(\mathbb{R}^d)}$, $\|\Delta e^{t\Delta} u\|_{L^2(\mathbb{R}^d)} \lesssim t^{-1} \|u\|_{L^2(\mathbb{R}^d)}$.
(ii) $x e^{\vartheta(0) t\Delta} u(x) = -2\vartheta(0) t \nabla e^{\vartheta(0) t \Delta} u(x) + e^{\vartheta(0) t\Delta}(xu)(x)$, $x \in \mathbb{R}^d$.
(iii) $\||x|^2 e^{\vartheta(0) t\Delta} u\|_{L^2(\mathbb{R}^d)} \lesssim (1 \vee t)(1 \wedge t^{-d/4}) \|(1 + |x| + |x|^2) u\|_{L^1 \cap L^2(\mathbb{R}^d)}$.

PROOF. (i) For the second part use functional calculus. The first part follows from

$$\|e^{t\Delta} u\|_{L^2(\mathbb{R}^d)} = \|q_t * u\|_{L^2(\mathbb{R}^d)} \lesssim \min\big(\|u\|_{L^2(\mathbb{R}^d)}, t^{-d/4} \|u\|_{L^1(\mathbb{R}^d)}\big).$$

(ii) Let $i \in \{1, \dots, d\}$. The result follows from

$$\begin{aligned} x_i\big(e^{\vartheta(0)t\Delta} u\big)(x) &= x_i (q_{\vartheta(0)t} * u)(x) \\ &= \vartheta(0) t \int_{\mathbb{R}^d} \frac{x_i - y_i}{\vartheta(0) t} q_{\vartheta(0)t}(x - y) u(y)\, dy + \int_{\mathbb{R}^d} y_i q_{\vartheta(0)t}(x - y) u(y)\, dy \\ &= -2\vartheta(0) t (\partial_i q_{\vartheta(0)t} * u)(x) + \big(q_{\vartheta(0)t} * (x_i u)\big)(x). \end{aligned}$$

(iii) Applying the proof in (ii) twice for $i \in \{1, \dots, d\}$ gives

$$\begin{aligned} & x_i^2\big(e^{\vartheta(0)t\Delta} u\big)(x) \\ &\quad = -2 x_i \vartheta(0) t (\partial_i q_{\vartheta(0)t} * u)(x) + x_i\big(q_{\vartheta(0)t} * (x_i u)\big)(x) \\ &\quad = 4\vartheta^2(0) t^2 \big(\partial_{ii}^2 q_{\vartheta(0)t} * u\big)(x) - 2\vartheta(0) t \big(\partial_i q_{\vartheta(0)t} * (x_i u)\big)(x) + \big(q_{\vartheta(0)t} * (x_i^2 u)\big)(x). \end{aligned}$$

Summing over i with $v_i = e^{\vartheta(0)t\Delta}(x_i u)$ obtain from this for $\||x|^2 e^{\vartheta(0)t\Delta} u\|_{L^2(\mathbb{R}^d)}$ up to a constant the upper bound

$$t^2 \|\Delta e^{\vartheta(0)t\Delta} u\|_{L^2(\mathbb{R}^d)} + t \sum_{i=1}^d \|\partial_i v_i\|_{L^2(\mathbb{R}^d)} + \|e^{\vartheta(0)t\Delta}(|x|^2 u)\|_{L^2(\mathbb{R}^d)}.$$

Using $e^{\vartheta(0)t\Delta} = e^{\vartheta(0)(t/2)\Delta} e^{\vartheta(0)(t/2)\Delta}$ and the two statements in (i) yield for the first and last terms the claimed bound. For the second term integration by parts implies $\|\partial_i v_i\|^2_{L^2(\mathbb{R}^d)} \leq \langle -\Delta v_i, v_i\rangle_{L^2(\mathbb{R}^d)} \leq \|\Delta v_i\|_{L^2(\mathbb{R}^d)} \|v_i\|_{L^2(\mathbb{R}^d)}$. The result follows again from applying (i). $\square$

LEMMA A.3. *If $z \in H^2(\mathbb{R}^d)$ has compact support, then $\langle z, \partial_i z\rangle_{L^2(\mathbb{R}^d)} = 0$, $\langle x_i \Delta z, z\rangle_{L^2(\mathbb{R}^d)} = -\langle x_i, |\nabla z|^2\rangle_{L^2(\mathbb{R}^d)}$ for $i = 1, \dots, d$. If $z \in H^4(\mathbb{R}^d)$, then also $\langle \Delta z, e^{t\Delta}\Delta\partial_i z\rangle_{L^2(\mathbb{R}^d)} = 0$, $t \geq 0$.*

PROOF. Integration by parts gives $\langle z, \partial_i z\rangle_{L^2(\mathbb{R}^d)} = -\langle \partial_i z, z\rangle_{L^2(\mathbb{R}^d)}$ (argue with compactly supported z first, then extend by continuity), implying $\langle z, \partial_i z\rangle_{L^2(\mathbb{R}^d)} = 0$ and $\langle x_i \partial_{jj} z, z\rangle_{L^2(\mathbb{R}^d)} = -\langle x_i, (\partial_j z)^2\rangle_{L^2(\mathbb{R}^d)}$ for $j = 1, \dots, d$. The last part follows from the first one for $\tilde{z}_t = e^{(t/2)\Delta} z \in H^2(\mathbb{R}^d)$ using

$$\langle e^{t\Delta}\Delta^2 z, \partial_i z\rangle_{L^2(\mathbb{R}^d)} = -\frac{d^2}{dt^2}\langle e^{t\Delta} z, \partial_i z\rangle_{L^2(\mathbb{R}^d)} = -\frac{d^2}{dt^2}\langle \tilde{z}_t, \partial_i \tilde{z}_t\rangle_{L^2(\mathbb{R}^d)}. \qquad \square$$

The upper bounds in the next Proposition are well known for analytic semigroups. The main difficulty is to ensure that they hold for growing domains, uniformly in $\delta > 0$.

PROPOSITION A.4. *There exist universal constants M_0, M_1 such that for $\delta, t > 0$:*

(i) $\|S^*_{\vartheta,\delta}(t)\|_{L^2(\Lambda_\delta)} \leq M_0 e^{C\delta^2 t}$,

(ii) $\|t A^*_{\vartheta,\delta} S^*_{\vartheta,\delta}(t)\|_{L^2(\Lambda_\delta)} \leq M_1 e^{C\delta^2 t}$.

PROOF. The claimed bounds in the statement follow from Proposition 2.1.1 of Lunardi [27], if we can show

$$\|(\lambda I - A^*_{\vartheta,\delta})^{-1}\|_{L^2(\Lambda_\delta)} \leq \frac{M}{|\lambda - w|}, \tag{A.8}$$

with $w = c_1\delta^2$ for all $\lambda \in \Sigma_{\sigma,w} := \{\rho \in \mathbb{C} : |\arg(\rho - w)| < \sigma\}\backslash\{w\}$ and with constants $c_1, M > 0, \sigma \in (\pi/2, \pi)$ independent of δ. Since the self-adjoint operator $\Delta_{\vartheta(\delta\cdot)}$ has strictly negative spectrum for all $\delta > 0$ (cf. Evans [11], Section 6.5), by functional calculus (A.8) holds indeed for $\Delta_{\vartheta(\delta\cdot)}$ with $w = 0$, $M = 1$ and any $\sigma \in (\pi/2, \pi)$.

In order to extend this to $A^*_{\vartheta,\delta}$, we consider it as a perturbation of $\Delta_{\vartheta(\delta\cdot)}$. We show first that $A^*_{0,\delta}$ is $\Delta_{\vartheta(\delta\cdot)}$-bounded, that is,

$$\|A^*_{0,\delta} v\|_{L^2(\Lambda_\delta)} \leq c_2\varepsilon\|\Delta_{\vartheta(\delta\cdot)} v\|_{L^2(\Lambda_\delta)} + \left(\frac{1}{4\varepsilon} + c_3\right)\delta^2\|v\|_{L^2(\Lambda_\delta)} \tag{A.9}$$

for $\varepsilon > 0$, $v \in H^1_0(\Lambda_\delta) \cap H^2(\Lambda_\delta)$ and absolute constants $c_2, c_3 > 0$. For this note that $\|A^*_{0,\delta} v\|_{L^2(\Lambda_\delta)}$ is upper bounded by

$$\|\delta\langle a(\delta\cdot), \nabla v\rangle\|_{L^2(\Lambda_\delta)} + \delta^2\big(\|v \operatorname{div}(a(\delta\cdot))\|_{L^2(\Lambda_\delta)} + \|b\|_\infty\|v\|_{L^2(\Lambda_\delta)}\big).$$

Moreover, $\|\delta\langle a(\delta\cdot), \nabla v\rangle\|_{L^2(\Lambda_\delta)}$ is upper bounded by

$$\begin{aligned}
&\delta d^{1/2} \sup_{i=1,\dots,d}\|a_i\|_\infty\left(\sum_{i=1}^d \|\partial_i v\|^2_{L^2(\Lambda_\delta)}\right)^{1/2}\\
&\quad\leq \delta\frac{d^{1/2}\sup_{i=1,\dots,d}\|a_i\|_\infty}{\min_x \vartheta(x)^{1/2}}\langle(-\Delta_{\vartheta(\delta\cdot)} v), v\rangle^{1/2}_{L^2(\Lambda_\delta)}\\
&\quad\leq \frac{d^{1/2}\sup_{i=1,\dots,d}\|a_i\|_\infty}{\min_x \vartheta(x)^{1/2}}\|\Delta_{\vartheta(\delta\cdot)} v\|^{1/2}_{L^2(\Lambda_\delta)}\delta\|v\|^{1/2}_{L^2(\Lambda_\delta)}\\
&\quad\leq c_2\varepsilon\|\Delta_{\vartheta(\delta\cdot)} v\|_{L^2(\Lambda_\delta)} + \frac{\delta^2}{4\varepsilon}\|v\|_{L^2(\Lambda_\delta)},
\end{aligned} \tag{A.10}$$

with $c_2 := \frac{d \sup_{i=1,\dots,d} \|a_i\|_\infty^2}{\min_x \vartheta(x)}$, where we used in the last line the basic inequality $xy \leq \varepsilon x^2 + \frac{1}{4\varepsilon} y^2$ for $x, y > 0$. This shows (A.9) with $c_3 := \sum_{i=1}^d \|\partial_i a_i\|_\infty + \|b\|_\infty$.

Choosing ε sufficiently small, the proof of Lemma III.2.6 in Engel and Nagel [10] implies (A.8) for all $\lambda \in \Sigma_{\sigma,0} \cap \{\rho \in \mathbb{C} : |\rho| > c_4\delta^2\}$ with $c_4 = \frac{(4\varepsilon)^{-1}+c_3}{1-2c_2\varepsilon}$, $\sigma = 3\pi/4$ and $M' > 0$ instead of M. Setting $w = (1 + c_5)c_4\delta^2$, for a suitable constant $c_5 > 0$ to be determined later, and assuming that for these λ

$$\lambda + w \in \Sigma_{\sigma,0} \cap \{\rho \in \mathbb{C} : |\rho| > c_4\delta^2\}, \quad |\lambda + w| \geq C|\lambda|, \tag{A.11}$$

with a universal constant C, we can therefore conclude for any $\lambda \in \Sigma_{\sigma,0} \cap \{\rho \in \mathbb{C} : |\rho| > c_4\delta^2\}$ that

$$\big\| \big((\lambda + w)I - A^*_{\vartheta,\delta}\big)^{-1} \big\|_{L^2(\Lambda_\delta)} \leq \frac{M'}{|\lambda + w|} \leq \frac{M'C}{|\lambda|}. \tag{A.12}$$

In order to obtain (A.8) from this let $\lambda \in \Sigma_{\sigma,w}$ such that $\lambda - w \in \Sigma_{\sigma,0}$. Assume that we can also show

$$|\lambda - w| > c_4\delta^2. \tag{A.13}$$

Then the result follows from (A.12) with $c_1 = (1 + c_5)c_4$, $M = M'C$, because

$$\big\| \big(\lambda I - A^*_{\vartheta,\delta}\big)^{-1} \big\|_{L^2(\Lambda_\delta)} = \big\| \big(((\lambda - w) + w)I - A^*_{\vartheta,\delta}\big)^{-1} \big\|_{L^2(\Lambda_\delta)} \leq \frac{M'C}{|\lambda - w|}.$$

We are left with showing (A.11) and (A.13). For (A.11) note that $\lambda \in \Sigma_{\sigma,0}$ already yields $\lambda + w \in \Sigma_{\sigma,0}$, because $w > 0$, while the inequality $|\lambda + w| > c_4\delta^2$ holds clearly, if $|\operatorname{Im}(\lambda)| > c_4\delta^2$. On the other hand, $|\arg(\lambda)| < \sigma$ implies $|\operatorname{Re}(\lambda)| < c_5|\operatorname{Im}(\lambda)|$ for a constant $c_5 > 0$ and thus, if $|\operatorname{Im}(\lambda)| \leq c_4\delta^2$, then

$$|\lambda + w| \geq w - |\operatorname{Re}(\lambda)| \geq w - c_5|\operatorname{Im}(\lambda)| > c_4\delta^2. \tag{A.14}$$

In order to find the constant C in (A.11), note that $|\lambda + w| \geq |\lambda|$ holds always if $\operatorname{Re}(\lambda) \geq 0$, and that $|\lambda + w| \geq \frac{1}{2}|\lambda|$ whenever $2w \leq |\lambda|$. Let now $\operatorname{Re}(\lambda) < 0$ and $|\lambda| < 2w$ such that by (A.11) $|\lambda + w| > c_4\delta^2 = \frac{2w}{2(1+c_5)} > C|\lambda|$, with $C := \frac{1}{2(1+c_5)}$. Finally, with respect to (A.13), $|\lambda - w| > c_4\delta^2$ holds always, if $|\operatorname{Im}(\lambda)| > c_4\delta^2$. On the other hand, $|\arg(\lambda - w)| < \sigma$ implies $|\arg(\lambda)| < \sigma$ and hence for $|\operatorname{Im}(\lambda)| \leq c_4\delta^2$, as in (A.14), $|\lambda - w| \geq w - |\operatorname{Re}(\lambda)| > c_4\delta^2$. □

With these preparations we can proceed to proving Proposition 3.5.

PROOF OF PROPOSITION 3.5. (i) The proof is based on giving a stochastic representation for $S^*_{\vartheta,\delta}(t)z$ via the Feynman–Kac formulas. Without loss of generality let $\vartheta \in C^{1+\alpha}(\mathbb{R}^d)$, $a \in C^{1+\alpha}(\mathbb{R}^d; \mathbb{R}^d)$, $b \in C^\alpha(\mathbb{R}^d)$, $\alpha > 0$, with $\min_{x \in \mathbb{R}^d} \vartheta(x) > 0$. Then for $f \in C_c^2(\mathbb{R}^d)$

$$A^*_{\vartheta,\delta} f(x) = \vartheta(\delta x)\Delta f(x) + \langle \tilde{a}_\delta(x), \nabla f(x)\rangle_{\mathbb{R}^d} + \tilde{b}_\delta(x) f(x), \quad x \in \mathbb{R}^d, \tag{A.15}$$

where $\tilde{a}_\delta = \delta(\nabla\vartheta(\delta\cdot) - a(\delta\cdot)) \in C^\alpha(\mathbb{R}^d)$, $\tilde{b}_\delta = \delta^2(b(\delta\cdot) - \operatorname{div}(a(\delta\cdot)) \in C^\alpha(\mathbb{R}^d)$. By Karatzas and Shreve ([22], Theorem 5.4.22), we can find a process $Y^{(\delta)} = (Y^{(\delta)}_t)_{t \geq 0}$ being a weak solution of the d-dimensional stochastic differential equation

$$dY^{(\delta)}_t = \tilde{a}_\delta(Y^{(\delta)}_t)\,dt + \sqrt{2}\vartheta(\delta Y^{(\delta)}_t)^{1/2}\,d\tilde{W}_t, \qquad Y^{(\delta)}_0 = x \in \mathbb{R}^d,$$

on a filtered probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t\geq 0}, \tilde{\mathbb{P}})$ carrying a scalar Brownian motion $(\tilde{W}_t)_{t\geq 0}$. We show below for $x \in \mathbb{R}^d$

$$(S^*_{\vartheta,\delta}(t)z)(x) = \tilde{\mathbb{E}}_x\Big[z\big(Y^{(\delta)}_t\big)\exp\Big(\int_0^t \tilde{b}_\delta\big(Y^{(\delta)}_s\big)\,ds\Big)\mathbf{1}_{\{t<\tau_\delta(Y^{(\delta)})\}}\Big], \tag{A.16}$$

where $\tilde{\mathbb{P}}_x$ and $\tilde{\mathbb{E}}_x$ indicate the initial value and $\tau_\delta(Y^{(\delta)}) := \inf\{t \geq 0 : Y^{(\delta)}_t \notin \Lambda_\delta\}$. Assume first this holds true. Denote the transition densities of $Y^{(\delta)}$ by $p_{\delta,t}(x, y)$, $x, y \in \mathbb{R}^d$. According to Sheu [34], equation (1.4), we have $p_{\delta,t}(x, y) \leq c_3 q_{c_2 t}(x-y)$ for universal constants $c_2, c_3 > 0$. Then by (A.16), using $\|\tilde{b}_\delta\|_\infty \leq c_1\delta^2$ for some constant $c_1 > 0$, it follows

$$\begin{aligned}\big|(S^*_{\vartheta,\delta}(t)z)(x)\big| &\leq e^{c_1 t\delta^2}\tilde{\mathbb{E}}_x\big[\big|z\big(Y^{(\delta)}_t\big)\big|\big] = e^{c_1 t\delta^2}\int_{\mathbb{R}^d}|z(y)|p_{\delta,t}(x,y)\,dy\\ &\leq c_3 e^{c_1 t\delta^2} q_{c_2 t} * |z|(x) = c_3 e^{c_1 t\delta^2}\big(e^{c_2 t\Delta}|z|\big)(x).\end{aligned}$$

We are left with showing (A.16). The proof is similar to Friedman [14], Theorem 6.5.2, and extends Peres and Mörters ([29], Theorem 7.44), which applies only to Brownian motion. It is enough to consider $x \in \overline{\Lambda}_\delta$, because otherwise $(S^*_{\vartheta,\delta}(t)z)(x) = 0$ and $\mathbf{1}_{\{t<\tau_\delta(Y_\delta)\}} = 0$ $\tilde{\mathbb{P}}_x$-a.s. and so (A.16) holds trivially. The function $u(t) = u(t,\cdot)$ with $u(t,x) := (S^*_{\vartheta,\delta}(t)z)(x)$ for $t \geq 0$, $x \in \overline{\Lambda}_\delta$, is the unique solution in $L^2(\Lambda_\delta)$ of

$$\begin{cases}\dfrac{d}{dt}u(t) = A^*_{\vartheta,\delta}u(t), & t > 0,\\ u(0) = z,\ u(t)|_{\partial\Lambda_\delta} = 0, & t \geq 0,\end{cases}$$

where the derivative is taken in $L^2(\Lambda_\delta)$. Classical PDE theory yields $u \in C([0,\infty), \overline{\Lambda}_\delta) \cap C^{1,2}([\varepsilon,\infty), \overline{\Lambda}_\delta)$ for any $\varepsilon > 0$; see, for example, Friedman [14], Theorem 6.3.6 (here we use the regularity assumptions on ϑ, a, b). Set $h(t) = \exp(\int_0^t \tilde{b}_\delta(Y^{(\delta)}_s)\,ds)$ and let $\rho = \inf\{t \geq 0 : Y^{(\delta)}_t \notin U\}$ for a compact set $U \subseteq \Lambda_\delta$. Set $g(t', x) = u(t-t', x)$, $0 \leq t' \leq t$. Noting that $A^*_{\vartheta,\delta} - \tilde{b}_\delta$ generates the transition semigroup of $Y^{(\delta)}$ and $h'(t) = \tilde{b}_\delta(Y^{(\delta)}_t)h(t)$, Itô's formula shows for any $0 \leq t' < t$

$$\begin{aligned}&g\big(t'\wedge\rho, Y^{(\delta)}_{t'\wedge\rho}\big)h\big(t'\wedge\rho\big)\\ &\quad= g\big(0, Y^{(\delta)}_0\big) + \int_0^{t'\wedge\rho}\Big(A^*_{\vartheta,\delta}g(s,\cdot)\big(Y^{(\delta)}_s\big) - \frac{d}{ds}g\big(s, Y^{(\delta)}_s\big)\Big)h(s)\,ds\\ &\qquad+ \int_0^{t'\wedge\rho}\big\langle\nabla g\big(s, Y^{(\delta)}_s\big)h(s), d\tilde{W}_s\big\rangle_{\mathbb{R}^d}.\end{aligned}$$

Using the previous display, the second term vanishes. Taking expectations and letting $t' \to t$ yields therefore

$$\begin{aligned}\big(S^*_{\vartheta,\delta}(t)z\big)(x) &= g\big(0, Y^{(\delta)}_0\big) = \tilde{\mathbb{E}}_x\big[u\big(t - t\wedge\rho, Y^{(\delta)}_{t\wedge\rho}\big)h(t\wedge\rho)\big]\\ &= \tilde{\mathbb{E}}_x\big[z\big(Y^{(\delta)}_t\big)h(t)\mathbf{1}_{\{\rho>t\}}\big] + \tilde{\mathbb{E}}_x\big[u\big(t-\rho, Y^{(\delta)}_\rho\big)h(\rho)\mathbf{1}_{\{\rho\leq t\}}\big].\end{aligned}$$

If U exhausts Λ_δ, then $\rho \to \tau_\delta(Y^{(\delta)})$ and $Y^{(\delta)}_\rho \to 0$ such that $u(t-\rho, Y^{(\delta)}_\rho) \to 0$. This implies (A.16).

(ii) We can assume $z \in C(\overline{\Lambda}_\delta)$ for sufficiently small δ. Indeed, for $z \in L^2(\mathbb{R}^d)$ let $z^{(\varepsilon)} \in C_c(\mathbb{R}^d)$ converge to z in $L^2(\mathbb{R}^d)$ as $\varepsilon \to 0$. For small δ we have $z^{(\varepsilon)} \in C(\overline{\Lambda}_\delta)$. Applying Proposition A.4(i) to $S^*_{\vartheta,\delta}(t)(z|_{\Lambda_\delta} - z^{(\varepsilon)})$, Lemma A.2(i) to $e^{\vartheta(0)t\Delta}(z - z^{(\varepsilon)})$ we have

$$\begin{aligned}&\big\|(S^*_{\vartheta,\delta}(t)(z|_{\Lambda_\delta}) - e^{\vartheta(0)t\Delta}z\big\|_{L^2(\mathbb{R}^d)}\\ &\quad\lesssim \big(e^{c_1\delta^2 t}+1\big)\big\|z - z^{(\varepsilon)}\big\|_{L^2(\mathbb{R}^d)} + \big\|\big(S^*_{\vartheta,\delta}(t) - e^{\vartheta(0)t\Delta}\big)z^{(\varepsilon)}\big\|_{L^2(\mathbb{R}^d)}.\end{aligned}$$

Using the statement with respect to $z^{(\varepsilon)}$, and letting first $\delta \to 0$ and then $\varepsilon \to 0$, the last line tends to zero.

For $z \in C(\overline{\Lambda}_\delta)$ it is enough to show $(S^*_{\vartheta,\delta}(t)z)(x) \to (e^{\vartheta(0)\Delta t}z)(x)$ pointwise for $x \in \mathbb{R}^d$. $L^2(\mathbb{R}^d)$-covergence follows then from (i) and dominated convergence. Using the notation from (i) we have the representation $(e^{\vartheta(0)t\Delta}z)(x) = \tilde{\mathbb{E}}_x[z(Y_t^{(0)})]$ for $Y_t^{(0)} = x + \sqrt{2}\vartheta(0)^{1/2}\tilde{W}_t$. (A.16) therefore allows us to write $S^*_{\vartheta,\delta}(t)z - e^{\vartheta(0)t\Delta}z =: T_1 + T_2 + T_3$ with

$$T_1 = \tilde{\mathbb{E}}_x\big[z\big(Y_t^{(\delta)}\big) - z\big(Y_t^{(0)}\big)\big],$$

$$T_2 = \tilde{\mathbb{E}}_x\left[z\big(Y_t^{(\delta)}\big)\left(\exp\left(\int_0^t \tilde{b}_\delta\big(Y_s^{(\delta)}\big)\,ds\right) - 1\right)\mathbf{1}_{\{t<\tau_\delta(Y^{(\delta)})\}}\right],$$

$$T_3 = -\tilde{\mathbb{E}}_x\big[z\big(Y_t^{(\delta)}\big)\mathbf{1}_{\{t\ge\tau_\delta(Y^{(\delta)})\}}\big].$$

We shall show that $T_i \to 0$, $i = 1, 2, 3$. The transition semigroup of $(Y_t^{(0)})_{t\ge0}$ is generated by $\mathcal{A}^{(0)} = \vartheta(0)\Delta$. Since $\mathcal{A}^{(\delta)} f \to \mathcal{A}^{(0)} f$ uniformly on $\mathbb{R}^d$ for $f \in C_c^\infty(\mathbb{R}^d)$ as $\delta \to 0$, it follows from Kallenberg [21], Theorem 19.25, that $Y^{(\delta)} \xrightarrow{d} Y^{(0)}$ with respect to the uniform topology on compacts in $\mathbb{R}_+$. This yields $T_1 \to 0$. As z is bounded and $\sup_{s>0} |\tilde{b}_\delta(Y_s^{(\delta)})| \lesssim \delta^2$, we also have $|T_2| \lesssim e^{Ct\delta^2}\delta^2$ and $|T_3| \lesssim \tilde{\mathbb{P}}_x(\tau_\delta(Y^{(\delta)}) \le t)$. To see why $\tilde{\mathbb{P}}_x(\tau_\delta(Y^{(\delta)}) \le t) \to 0$ holds let $Z_s^{(\delta)} = Y_s^{(\delta)} - x - \int_0^s \tilde{a}_\delta(Y_{s'}^{(\delta)})\,ds'$ and observe that $|\int_0^s \tilde{a}_\delta(Y_{s'}^{(\delta)})\,ds'| \lesssim \delta t$ such that

$$\tilde{\mathbb{P}}_x\big(\tau_\delta\big(Y^{(\delta)}\big) \le t\big) \le \sum_{i=1}^d \tilde{\mathbb{P}}_x\Big(\max_{0\le s\le t}\big|Z_s^{(\delta,i)}\big| \ge C\delta^{-1}\Big),$$

where $Z^{(\delta)} = (Z^{(\delta,i)})_{1\le i\le d}$. Since each $Z^{(\delta,i)}$ is a continuous martingale vanishing at 0 such that $\langle Z^{(\delta,i)}\rangle_s = 2\int_0^s \vartheta(\delta Y_{s'}^{(\delta)})\,ds' \le cs$, $c > 0$, uniformly in $i = 1, \dots, d$, we find for some scalar Brownian motion $(\tilde{B}_s)_{s\ge0}$ and $\tilde{c} > 0$

$$\tilde{\mathbb{P}}_x\big(\tau_\delta\big(Y^{(\delta)}\big) \le t\big) \le d\tilde{\mathbb{P}}_x\Big(\max_{0\le s\le t}|\tilde{B}_{cs}| \ge C\delta^{-1}\Big) \lesssim e^{-\tilde{c}\delta^{-2}t^{-1}},$$

because the density of the running maximum of a Brownian motion decays exponentially (Karatzas and Shreve [22], Chapter 2.8). This yields $T_3 \to 0$. □

LEMMA A.5. *Let $z \in H^2(\mathbb{R}^d)$ have compact support in $\Lambda_{\delta'}$ for some $\delta' > 0$. For $0 < \delta \le \delta'$ set $v^{(\delta)} := \delta^{-1}(A^*_{\vartheta,\delta} - \vartheta(0)\Delta)z$ and define*

$$v := \Delta\big(\langle\nabla\vartheta(0), x\rangle_{\mathbb{R}^d} z\big)(x) - \big\langle\nabla\vartheta(0) + a(0), \nabla z(x)\big\rangle_{\mathbb{R}^d}.$$

Then the following hold:

(i) $\|v^{(\delta)}\|_{L^1\cap L^2(\mathbb{R}^d)} \le C\|z\|_{W^2_{1,2}(\mathbb{R}^d)}$ *and* $v^{(\delta)} \to v$ *in* $L^2(\mathbb{R}^d)$ *for* $\delta \to 0$.

(ii) *If* $\int_{\mathbb{R}^d} z(x)\,dx = 0$, *then* $v = \Delta m$ *for* $m \in H^2(\mathbb{R}^d)$ *with compact support. Moreover, if* $\vartheta \in C^{1+\alpha'}(\overline{\Lambda})$ *for* $0 \le \alpha' \le 1$, *then* $\|v^{(\delta)} - v\|_{L^1\cap L^2(\mathbb{R}^d)} \le C\delta^{\alpha'}\|z\|_{W^2_{1,2}(\mathbb{R}^d)}$.

(iii) *If* $\int_{\mathbb{R}^d} z(x)\,dx = 0$ *and* $\int_{\mathbb{R}^d} xz(x)\,dx = 0$, *then* $z = \Delta m$ *for* $m \in H^4(\mathbb{R}^d)$ *with compact support*.

PROOF. (i) Without loss of generality let ϑ, a, b as well as the partial derivatives of ϑ, a be bounded on $\mathbb{R}^d$. Then for $x \in \mathbb{R}^d$

$$v^{(\delta)}(x) = \frac{\vartheta(\delta x) - \vartheta(0)}{\delta}\Delta z(x) + \big\langle\nabla\vartheta(\delta x) - a(\delta x), \nabla z(x)\big\rangle_{\mathbb{R}^d} + \delta\big(b(\delta x) - (\operatorname{div} a)(\delta x)\big)z(x). \tag{A.17}$$

From this obtain the upper bound on $v^{(\delta)}$ and the convergence in $L^2(\mathbb{R}^d)$ to

$$\langle \nabla\vartheta(0), x\rangle_{\mathbb{R}^d} \Delta z(x) - \langle -\nabla\vartheta(0) + a(0), \nabla z(x)\rangle_{\mathbb{R}^d} = v(x).$$

(ii) In order to find m, as z has compact support, it suffices to find a compactly supported function $g \in H^2(\mathbb{R}^d)$ with $\Delta g = \langle \nabla\vartheta(0) + a(0), \nabla z\rangle_{\mathbb{R}^d}$ in $L^2(\mathbb{R}^d)$ and to set $m := \langle \nabla\vartheta(0), x\rangle_{\mathbb{R}^d} z(x) - g(x)$. Using the Fourier transform $\mathcal{F}g(\omega) = \int_{\mathbb{R}^d} g(x)e^{i\langle x,\omega\rangle}\,dx$ this means by usual Fourier calculus

$$-|\omega|^2 \mathcal{F}g(\omega) = \langle \nabla\vartheta(0) + a(0), i\omega\rangle_{\mathbb{R}^d} \mathcal{F}z(\omega), \quad \omega \in \mathbb{R}^d.$$

By the compact support of z and $\int_{\mathbb{R}^d} z(x)\,dx = 0$ the Fourier transform $\mathcal{F}z$ is analytic with $\mathcal{F}z(0) = 0$. We can thus define

$$g(x) := \mathcal{F}^{-1}[u](x) \quad \text{with } u(\omega) := \langle \nabla\vartheta(0) + a(0), -|\omega|^{-1} i\omega\rangle_{\mathbb{R}^d} \frac{\mathcal{F}z(\omega)}{|\omega|}$$

as the inverse Fourier transform of the L^2-function u. Noting $z \in H^2(\mathbb{R}^d)$ and $|u(\omega)| \lesssim |\mathcal{F}z(\omega)|$ for $|\omega| \to \infty$, we see $g \in H^2(\mathbb{R}^d)$ and $\Delta g = \langle \nabla\vartheta(0) + a(0), \nabla z\rangle_{\mathbb{R}^d}$ in $L^2(\mathbb{R}^d)$.

For compactness of g we use the Paley–Wiener Theorem (Rudin [33], Theorem II.7.22) to deduce from the compact support of z that $\mathcal{F}z$ can be extended to an entire function on $\mathbb{C}^d$, satisfying the exponential growth condition $|\mathcal{F}z(\omega)| \le \gamma_N(1 + |\omega|)^{-N} \exp(r|\mathrm{Im}(\omega)|)$, $\omega \in \mathbb{C}^d$, for all $N \in \mathbb{N}$ and suitable positive constants γ_N, r. Hence, u is the quotient of an entire function and $|\omega|^2$, which is also entire. A meromorphic function with removable singularity extends continuously to an entire function. Consequently, we can work with an entire function u, which by definition satisfies the same exponential growth condition. A reverse application of the Paley–Wiener Theorem shows that g has compact support.

Finally, we can assume that $\nabla\vartheta, a$ are uniformly α'-Hölder continuous on $\mathbb{R}^d$. The upper bound on $\|v^{(\delta)} - v\|_{L^1 \cap L^2(\mathbb{R}^d)}$ follows then for $x \in \mathbb{R}^d$ using (A.17) from

$$\begin{aligned} |v^{(\delta)}(x) - v(x)| &\lesssim \left| \frac{\vartheta(\delta x) - \vartheta(0)}{\delta} - \langle \nabla\vartheta(0), x\rangle_{\mathbb{R}^d} \right| |\Delta z(x)| \\ &\quad + |\nabla\vartheta(\delta x) - \nabla\vartheta(0) + a(0) - a(\delta x)||\nabla z(x)| + \delta|z(x)|. \end{aligned}$$

(iii) The argument is similar to (ii). As above, the Fourier transform $\mathcal{F}z$ is analytic with $\mathcal{F}z(0) = 0$. The assumption $\int_{\mathbb{R}^d} x_i z(x)\,dx = 0$ gives also $\partial_i(\mathcal{F}z)(0) = 0$, $i = 1, \dots, d$. It follows for

$$m(x) := \mathcal{F}^{-1}[u](x) \quad \text{with } u(\omega) := -\frac{\mathcal{F}z(\omega)}{|\omega|^2}$$

that $m \in H^4(\mathbb{R}^d)$ and $\Delta m = z$. A Paley–Wiener argument as in (ii) shows that m has compact support. □

The following heat kernel bounds will be used frequently. The conditions in (iii) are essential for $d = 1$ to improve on (ii).

LEMMA A.6. *Let the functions $u, w \in L^2(\mathbb{R}^d)$, $z \in H^2(\mathbb{R}^d)$ have compact support in Λ_δ for some $\delta > 0$. Then for $0 < t \le T\delta^{-2}$:*

(i) $\|S^*_{\vartheta,\delta}(t)u\|_{L^2(\Lambda_\delta)} \le e^{CT}(1 \wedge t^{-d/4})\|u\|_{L^1 \cap L^2(\mathbb{R}^d)}$.

(ii) *If* $\|w - \Delta z\|_{L^1 \cap L^2(\mathbb{R}^d)} \le C\delta^{\alpha'} \|z\|_{W^2_{1,2}(\mathbb{R}^d)}$ *for* $0 \le \alpha' \le 1$, *then*

$$\|S^*_{\vartheta,\delta}(t)w\|_{L^2(\Lambda_\delta)} \le e^{CT}(1 \wedge t^{-\alpha'/2 - d/4})\|z\|_{W^2_{1,2}(\mathbb{R}^d)}.$$

(iii) *If* $\vartheta \in C^{1+\alpha'}(\overline{\Lambda})$ *for* $0 \le \alpha' \le 1$ *and* $\int_{\mathbb{R}^d} z(x)\,dx = 0$, *then*

$$\|S^*_{\vartheta,\delta}(t)\Delta z\|_{L^2(\Lambda_\delta)} \le e^{CT}\big(1 \wedge t^{-1/2-\alpha'/2-d/4}\big)\|z\|_{W^2_{1,2}(\mathbb{R}^d)}.$$

PROOF. (i) The semigroup bound in Proposition 3.5(i), applied to a sequence of continuous functions approximating u, and Lemma A.2(i) show for $t \le T\delta^{-2}$

$$\|S^*_{\vartheta,\delta}(t)u\|_{L^2(\Lambda_\delta)} \lesssim e^{CT}\|e^{c_2 t\Delta}|u|\|_{L^2(\mathbb{R}^d)} \lesssim e^{CT}\big(1 \wedge t^{-d/4}\big)\|u\|_{L^1\cap L^2(\mathbb{R}^d)}.$$

(ii) Write $w = u^{(\delta)} + \vartheta(0)^{-1}A^*_{\vartheta,\delta}z$ for $u^{(\delta)} = (w - \Delta z) - \vartheta(0)^{-1}\delta v^{(\delta)}$ and $v^{(\delta)} = \delta^{-1}(A^*_{\vartheta,\delta} - \vartheta(0)\Delta)z$ such that

$$\|S^*_{\vartheta,\delta}(t)w\|_{L^2(\Lambda_\delta)} \le \|S^*_{\vartheta,\delta}(t)u^{(\delta)}\|_{L^2(\Lambda_\delta)} + \vartheta(0)^{-1}\|A^*_{\vartheta,\delta}S^*_{\vartheta,\delta}(t)z\|_{L^2(\Lambda_\delta)}.$$

The second term is up to a constant bounded by $e^{CT}(1 \wedge t^{-1})\|S^*_{\vartheta,\delta}(t/2)z\|_{L^2(\Lambda_\delta)}$, using Proposition A.4(ii) for $t \le T\delta^{-2}$ and $S^*_{\vartheta,\delta}(t) = S^*_{\vartheta,\delta}(t/2)S^*_{\vartheta,\delta}(t/2)$. Applying (i) to $u = z$ gives the upper bound $e^{CT}(1 \wedge t^{-1-d/4})\|z\|_{L^2(\mathbb{R}^d)}$. The result follows from applying (i) to $u = u^{(\delta)}$ in the last display and noting that $\|u^{(\delta)}\|_{L^1\cap L^2(\mathbb{R}^d)} \le e^{CT}(1 \wedge t^{-\alpha'/2})\|z\|_{W^2_{1,2}(\mathbb{R}^d)}$ by Lemma A.5(i) and $\delta \le (T/t)^{1/2}$, as well as adjusting the constant C.

(iii) Following the proof of (ii) for $w = \Delta z$ it is enough to show the improved upper bound $\|S^*_{\vartheta,\delta}(t)u^{(\delta)}\|_{L^2(\Lambda_\delta)} \lesssim e^{CT}(1 \wedge t^{-1/2-\alpha'/2-d/4})\|z\|_{W^2_{1,2}(\mathbb{R}^d)}$. Lemma A.5(ii) shows the existence of a compactly supported $m \in H^2(\mathbb{R}^d)$ such that $\|v^{(\delta)} - \Delta m\|_{L^1\cap L^2(\mathbb{R}^d)} \le C\delta^{\alpha'}\|z\|_{W^2_{1,2}(\mathbb{R}^d)}$. With $\tilde{u}^{(\delta)} = v^{(\delta)} - \Delta m$ write $u^{(\delta)} = \vartheta(0)^{-1}\delta\tilde{u}^{(\delta)} + \vartheta(0)^{-1}\delta\Delta m$. Applying (i) to $u = \tilde{u}^{(\delta)}$ and (ii) to $z = m$ yields

$$\|S^*_{\vartheta,\delta}(t)u^{(\delta)}\|_{L^2(\Lambda_\delta)} \lesssim e^{CT}\big(\delta^{1+\alpha'}\big(1 \wedge t^{-d/4}\big) + \delta\big(1 \wedge t^{-1/2-d/4}\big)\big)\|z\|_{W^2_{1,2}(\mathbb{R}^d)}.$$

For $\delta \le (T/t)^{1/2}$ the order in t is $1 \wedge t^{-1/2-\alpha'/2-d/4}$, as claimed. □

LEMMA A.7. *Let* $u \in L^2(\mathbb{R}^d)$, $z \in H^2(\mathbb{R}^d)$ *have compact support in* Λ_δ *for some* $\delta > 0$. *Using* $\ell_{d,2}(\delta)$ *as in Assumption* 3.7, *we have*:

(i) $\int_0^T \langle S_\vartheta(t)X_0, u_\delta\rangle^2\,dt \le e^{CT}\|X_0\|^2\ell_{d,2}(\delta)\delta^{2\wedge d}\|u\|^2_{L^1\cap L^2(\mathbb{R}^d)}$.

(ii) *If* $X_0 \in L^p(\Lambda)$, $p \ge 2$, $1/p + 1/p' = 1$, *then, with* $\gamma(d,p) := 2\frac{1+d/p}{1+d/2} + d(1 - \frac{2}{p})$,

$$\int_0^T \langle S_\vartheta(t)X_0, (\Delta z)_\delta\rangle^2\,dt \le e^{CT}\|X_0\|^2_{L^p(\Lambda)}\delta^{\gamma(d,p)}\big(\|\Delta z\|^2_{L^{p'}(\mathbb{R}^d)} + \|z\|^2_{W^2_{1,2}(\mathbb{R}^d)}\big).$$

(iii) *If* $X_0 \in \mathcal{D}(A_\vartheta)$ *and for* $d = 1$ *also* $\vartheta \in C^{1+\alpha'}(\overline{\Lambda})$, $0 \le \alpha' \le 1$, *and* $\int_{\mathbb{R}} z(x)\,dx = 0$, *then*

$$\int_0^T \langle S_\vartheta(t)X_0, (\Delta z)_\delta\rangle^2\,dt \le e^{CT}\big(\|X_0\|^2 + \|A_\vartheta X_0\|^2\big)\ell_{d,2}(\delta)\delta^{2+2\alpha'}\|z\|^2_{W^2_{1,2}(\mathbb{R}^d)}.$$

PROOF. (i) By Lemma A.6(i) and the scaling in Lemma 3.1 we find

$$\begin{aligned}\int_0^T \langle S_\vartheta(t)X_0, u_\delta\rangle^2\,dt &\le \|X_0\|^2\delta^2\int_0^{T\delta^{-2}} \|S^*_{\vartheta,\delta}(t)u\|^2_{L^2(\Lambda_\delta)}\,dt\\ &\lesssim e^{CT}\|X_0\|^2\delta^2\int_0^{T\delta^{-2}}\big(1 \wedge t^{-d/2}\big)\,dt\,\|u\|^2_{L^1\cap L^2(\mathbb{R}^d)}.\end{aligned}$$

The claim follows, because the integral has order $O(1)$ for $d \geq 3$, order $O(\log(T\delta^{-2}))$ for $d = 2$ and order $O(T^{1/2}\delta^{-1})$ for $d = 1$.

(ii) It is enough to consider continuous X_0. Using the Hölder inequality and Proposition 3.5(i), we obtain

$$\langle S_\vartheta(t)X_0, (\Delta z)_\delta\rangle^2 \leq \|S_\vartheta(t)X_0\|^2_{L^p(\Lambda)}\|(\Delta z)_\delta\|^2_{L^{p'}(\Lambda)}$$
$$\lesssim \|q_{c_2t} * |X_0|\|^2_{L^p(\mathbb{R}^d)}\delta^{d(2/p'-1)}\|\Delta z\|^2_{L^{p'}(\mathbb{R}^d)}.$$

Here, $\|q_{c_2t} * |X_0|\|_{L^p(\mathbb{R}^d)} \leq \|X_0\|_{L^p(\Lambda)}$. For $\varepsilon > 0$, Lemmas 3.1 and A.6(ii) show

$$\int_\varepsilon^T \langle S_\vartheta(t)X_0, (\Delta z)_\delta\rangle^2\,dt \leq \|X_0\|^2 \int_\varepsilon^T \|S^*_{\vartheta,\delta}(t\delta^{-2})\Delta z\|^2_{L^2(\Lambda_\delta)}\,dt$$
$$\lesssim e^{CT}\|X_0\|^2 \int_\varepsilon^T (t\delta^{-2})^{-1-d/2}\,dt\|z\|^2_{W^2_{1,2}(\mathbb{R}^d)}.$$

Splitting up the integral and adjusting the constant C yields thus

$$\int_0^T \langle S_\vartheta(t)X_0, (\Delta z)_\delta\rangle^2\,dt \lesssim e^{CT}\|X_0\|^2\big(\delta^{d(2/p'-1)}\varepsilon + \delta^{2+d}\varepsilon^{-d/2}\big)$$
$$\times \big(\|\Delta z\|^2_{L^{p'}(\mathbb{R}^d)} + \|z\|^2_{W^2_{1,2}(\mathbb{R}^d)}\big).$$

The claim follows with $\|X_0\| \lesssim \|X_0\|_{L^p(\Lambda)}$ and $\varepsilon = \delta^{2\frac{1+d/p}{1+d/2}}$.

(iii) With $v^{(\delta)} := \delta^{-1}(A^*_{\vartheta,\delta} - \vartheta(0)\Delta)z$ converging to v as in Lemma A.5 and using the scaling in Lemma 3.1, write $\vartheta(0)(\Delta z)_\delta = -\delta v^{(\delta)}_\delta + \delta^2 A^*_\vartheta z_\delta$. Then

$$\langle S_\vartheta(t)X_0, (\Delta z)_\delta\rangle^2 \lesssim \delta^2\langle S_\vartheta(t)X_0, v^{(\delta)}_\delta\rangle^2 + \delta^4\langle S_\vartheta(t)A_\vartheta X_0, z_\delta\rangle^2,$$

and the claim follows for $d \geq 2$ from applying (i) with $u = v^{(\delta)}$ and $u = z$ (with $A_\vartheta X_0 \in L^2(\Lambda)$ instead of X_0), while for $d = 1$ apply (i) also with $u = v^{(\delta)} - v$ and (ii) with $z = m$, using in addition Lemma A.5(ii) with $v = \Delta m$. $\square$

A.3. Asymptotic results for the covariances. The general idea for the proofs in this section is to apply the scaling in Lemma 3.1 to the covariance function as in Section 3.3 and to deduce a limit for the integral using the heat kernel bounds and the convergence of the semigroups from the last section.

PROPOSITION A.8. *Grant Assumption 3.2. Consider functions $z \in H^2(\mathbb{R}^d)$, $u \in L^2(\mathbb{R}^d)$, $(w^{(\delta)})_{\delta>0}$, $(u^{(\delta)})_{\delta>0} \subseteq L^2(\mathbb{R}^d)$ with compact support in $\Lambda_{\delta'}$ for some $\delta' > 0$. Assume for $0 < \delta \leq \delta'$ that $\|w^{(\delta)} - \Delta z\|_{L^1\cap L^2(\mathbb{R}^d)} \leq C\delta^{\alpha'}$ for $\alpha' > 1/2$, $\|u^{(\delta)} - u\|_{L^2(\mathbb{R}^d)} \to 0$ as $\delta \to 0$. Then with Ψ from (3.2):*

(i) $\delta^{-2}\operatorname{Var}(\langle\tilde{X}(t), w^{(\delta)}_\delta\rangle) \to \vartheta(0)^{-1}\Psi(\Delta z, \Delta z)$, $t > 0$.
(ii) $\delta^{-2}\int_0^T \operatorname{Var}(\langle\tilde{X}(t), w^{(\delta)}_\delta\rangle)\,dt \to T\vartheta(0)^{-1}\Psi(\Delta z, \Delta z)$.
(iii) *If $d \geq 2$, then* $\delta^{-2}\int_0^T \operatorname{Cov}(\langle\tilde{X}(t), w^{(\delta)}_\delta\rangle, \langle\tilde{X}(t), u^{(\delta)}_\delta\rangle)\,dt \to T\vartheta(0)^{-1}\Psi(\Delta z, u)$.

PROOF. (i) By (3.4) for $t = t'$ we have

$$\delta^{-2}\operatorname{Var}(\langle\tilde{X}(t), w^{(\delta)}_\delta\rangle) = \int_0^{t\delta^{-2}} \|B^*_\delta S^*_{\vartheta,\delta}(s)w^{(\delta)}\|^2_{L^2(\Lambda_\delta)}\,ds = \int_0^\infty f_\delta(s)\,ds,$$

with $f_\delta(s) = \|B^*_\delta S^*_{\vartheta,\delta}(s)w^{(\delta)}\|^2_{L^2(\Lambda_\delta)}\mathbf{1}_{\{s\leq t\delta^{-2}\}}$. Set $f(s) = \|B^*_0 e^{\vartheta(0)s\Delta}\Delta z\|^2_{L^2(\mathbb{R}^d)}$ and note $\int_0^\infty f(s)\vartheta(0)\,ds = \Psi(\Delta z, \Delta z)$, substituting $ds' = \vartheta(0)\,ds$. By assumption $w^{(\delta)} \to \Delta z$ in

$L^2(\Lambda_\delta)$ and by Proposition 3.5(ii) above $S^*_{\vartheta,\delta}(s)\Delta z \to e^{\vartheta(0)s\Delta}\Delta z$ in $L^2(\mathbb{R}^d)$. From Proposition A.4(i) above we have $\sup_{0\le s\le t/\delta^2}\|S^*_{\vartheta,\delta}(s)\|_{L^2(\Lambda_\delta)} < \infty$, as well as $\sup_{0<\delta\le 1}\|B^*_\delta\|_{L^2(\mathbb{R}^d)} < \infty$ by Assumption 3.2 and the uniform boundedness principle. We deduce

$$\begin{aligned}&\|B^*_\delta S^*_{\vartheta,\delta}(s)w^{(\delta)} - B^*_0 e^{\vartheta(0)s\Delta}\Delta z\|_{L^2(\mathbb{R}^d)}\\ &\quad\le \|B^*_\delta\|\big(\|S^*_{\vartheta,\delta}(s)\|_{L^2(\Lambda_\delta)}\|w^{(\delta)}-\Delta z\|_{L^2(\mathbb{R}^d)} + \|S^*_{\vartheta,\delta}(s)\Delta z - e^{\vartheta(0)s\Delta}\Delta z\|_{L^2(\mathbb{R}^d)}\big)\\ &\qquad + \|(B^*_\delta - B^*_0)e^{\vartheta(0)s\Delta}\Delta z\|_{L^2(\mathbb{R}^d)} \to 0,\end{aligned}$$

which implies $f_\delta(s)\to f(s)$ pointwise. Lemma A.6(ii) yields $|f_\delta(s)|\lesssim 1\wedge s^{-\alpha'-d/2}$. Since $\alpha'>1/2$, $\sup_{0<\delta\le 1} f_\delta(\cdot)\in L^1([0,\infty))$, for any fixed t, and the result follows from the dominated convergence theorem.

(ii) By (i) and Fatou's lemma we obtain

$$\liminf_{\delta\to 0}\delta^{-2}\int_0^T \operatorname{Var}(\langle\tilde X(t), w^{(\delta)}_\delta\rangle)\,dt \ge T\vartheta(0)^{-1}\Psi(\Delta z,\Delta z).$$

On the other hand, $\operatorname{Var}(\langle\tilde X(t), w^{(\delta)}_\delta\rangle)$ is increasing in t; cf. (2.5). The result follows from

$$\begin{aligned}\limsup_{\delta\to 0}\delta^{-2}\int_0^T \operatorname{Var}(\langle\tilde X(t), w^{(\delta)}_\delta\rangle)\,dt &\le \lim_{\delta\to 0}\delta^{-2}T\operatorname{Var}(\langle\tilde X(T), w^{(\delta)}_\delta\rangle)\\ &= T\vartheta(0)^{-1}\Psi(\Delta z,\Delta z).\end{aligned}$$

(iii) Revisiting the derivations in (i) and (ii), we obtain

$$\begin{aligned}&\delta^{-2}\operatorname{Cov}(\langle\tilde X(t), w^{(\delta)}_\delta\rangle,\langle\tilde X(t), u^{(\delta)}_\delta\rangle) = \int_0^\infty f_\delta(s)\,ds,\\ &\quad\text{with } f_\delta(s) := \langle B^*_\delta S^*_{\vartheta,\delta}(s)w^{(\delta)}, B^*_\delta S^*_{\vartheta,\delta}(s)u^{(\delta)}\rangle_{L^2(\Lambda_\delta)}\mathbf{1}_{\{s\le t\delta^{-2}\}}.\end{aligned}$$

Putting $f(s) := \langle B^*_0 e^{\vartheta(0)s\Delta}\Delta z, B^*_0 e^{\vartheta(0)s\Delta}u\rangle$, we obtain as in (i), (iii) that $f_\delta(s)\to f(s)$ holds pointwise for $\delta\to 0$ by the L^2-continuity of the scalar product. Furthermore, the Cauchy–Schwarz inequality and Lemma A.6(i), (ii) yield the bound

$$\begin{aligned}|f_\delta(s)| &\lesssim \|S^*_{\vartheta,\delta}(s)w^{(\delta)}\|_{L^2(\Lambda_\delta)}\|S^*_{\vartheta,\delta}(s)u^{(\delta)}\|_{L^2(\Lambda_\delta)}\mathbf{1}_{\{s\le t/\delta^2\}}\\ &\lesssim e^{CT}(1\wedge s^{-\alpha'/2-d/2}) \lesssim 1\wedge s^{-5/4}\end{aligned}\tag{A.18}$$

for $d\ge 2$. Since this bound is integrable in $s\ge 0$, we conclude that

$$\delta^{-2}\operatorname{Cov}(\langle\tilde X(t), w^{(\delta)}_\delta\rangle,\langle\tilde X(t), u^{(\delta)}_\delta\rangle) \to \int_0^\infty f(s)\,ds = \vartheta(0)^{-1}\Psi(\Delta z, u),$$

meaning in particular that $\Psi(\Delta z, u)$ is well defined. What is more, the bound (A.18) also shows that the covariance is uniformly bounded in $t\in[0,T]$ so that another application of the dominated convergence theorem shows that the integral over $t\in[0,T]$ converges to $T\vartheta(0)^{-1}\Psi(\Delta z, u)$. □

PROPOSITION A.9. *Grant Assumption 3.2. Consider functions $z, m\in H^2(\mathbb{R}^d)$, $u\in L^2(\mathbb{R}^d)$, $(w^{(\delta)})_{\delta>0}$, $(u^{(\delta)})_{\delta>0}\subseteq L^2(\mathbb{R}^d)$ with compact supports in $\Lambda_{\delta'}$ for some $\delta'>0$. Assume for $0<\delta\le\delta'$ that $\|w^{(\delta)}-\Delta m\|_{L^1\cap L^2(\mathbb{R}^d)}\le C\delta^{\alpha'}$ for $\alpha'>1/2$, $\|u^{(\delta)}-u\|_{L^2(\mathbb{R}^d)}\to 0$ as $\delta\to 0$. Then*:

(i) *For $d\ge 2$,* $\operatorname{Var}(\int_0^T\langle\tilde X(t),(\Delta z)_\delta\rangle\langle\tilde X(t), u^{(\delta)}_\delta\rangle\,dt) = O(\delta^6\ell_{d,2}(\delta)^3)$.

(ii) $\operatorname{Var}(\int_0^T\langle\tilde X(t), w^{(\delta)}_\delta\rangle^2\,dt) = o(\delta^4)$.

(iii) *For* $d \geq 2$, $\operatorname{Var}(\int_0^T \langle \tilde{X}(t), u_\delta^{(\delta)} \rangle^2 \, dt) = O(\delta^4 \ell_{d,2}(\delta)^2)$.

(iv) *Let* $d \geq 2$, *or let* $\vartheta \in C^{1+\alpha'}(\overline{\Lambda})$ *and* $\int_{\mathbb{R}^d} m(x)\, dx = 0$. *Then with* Ψ *from* (3.2)

$$\delta^{-6} \operatorname{Var}\biggl(\int_0^T \langle \tilde{X}(t), w_\delta^{(\delta)} \rangle^2 \, dt\biggr) \to 4T\vartheta(0)^{-3} \int_0^\infty \Psi\bigl(e^{s\Delta}\Delta m, \Delta m\bigr)^2 \, ds.$$

PROOF. We first make some preliminary remarks. For $v, \tilde{v} \in L^2(\Lambda_\delta)$ set $\xi(t) = \langle \tilde{X}(t), v_\delta \rangle$, $\tilde{\xi}(t) = \langle \tilde{X}(t), \tilde{v}_\delta \rangle$. The random variables $\{\xi(t) | t \geq 0\} \cup \{\tilde{\xi}(t) | t \geq 0\}$ are jointly Gaussian and centered and so it follows from Wick's formula (Janson [20], Theorem 1.28) that

$$\begin{aligned} &\delta^{-6} \operatorname{Var}\biggl(\int_0^T \xi(t)\tilde{\xi}(t)\, dt\biggr) \\ &\quad = \delta^{-6} \int_0^T \int_0^T \operatorname{Cov}\bigl(\xi(t)\tilde{\xi}(t), \xi(s)\tilde{\xi}(s)\bigr)\, dt\, ds \\ &\quad = 2\delta^{-6} \int_0^T \int_0^t \bigl(\operatorname{Cov}\bigl(\xi(t), \xi(s)\bigr) \operatorname{Cov}\bigl(\tilde{\xi}(t), \tilde{\xi}(s)\bigr) \\ &\qquad + \operatorname{Cov}\bigl(\xi(t), \tilde{\xi}(s)\bigr) \operatorname{Cov}\bigl(\tilde{\xi}(t), \xi(s)\bigr)\bigr)\, dt\, ds \\ &\quad =: 2V_1 + 2V_2, \end{aligned} \tag{A.19}$$

with $V_1 = V(v, v, \tilde{v}, \tilde{v})$, $V_2 = V(v, \tilde{v}, \tilde{v}, v)$, where for $k, \tilde{k} \in L^2(\Lambda_\delta)$

$$\begin{aligned} V(v, \tilde{v}, k, \tilde{k}) := \delta^{-6} \int_0^T \int_0^t &\operatorname{Cov}\bigl(\langle \tilde{X}(t), v_\delta \rangle, \langle \tilde{X}(s), \tilde{v}_\delta \rangle\bigr) \\ &\times \operatorname{Cov}\bigl(\langle \tilde{X}(t), k_\delta \rangle, \langle \tilde{X}(s), \tilde{k}_\delta \rangle\bigr)\, ds\, dt, \end{aligned} \tag{A.20}$$

and $V(v) := V(v, v, v, v)$. It is thus enough to study V_1, V_2. Set

$$f_\delta\bigl((s, v), (s', \tilde{v})\bigr) = \langle B_\delta^* S_{\vartheta,\delta}^*(s) v, B_\delta^* S_{\vartheta,\delta}^*(s')\tilde{v} \rangle_{L^2(\Lambda_\delta)} \mathbf{1}_{\{0 < s, s' \leq T\delta^{-2}\}}, \quad s, s' \geq 0.$$

Then by (3.4), $V(v, \tilde{v}, k, \tilde{k})$ equals

$$\begin{aligned} &\delta^{-2} \int_0^T \int_0^t \biggl(\int_0^{s\delta^{-2}} f_\delta\bigl((t\delta^{-2} - r', v), (s\delta^{-2} - r', \tilde{v})\bigr)\, dr'\biggr) \\ &\qquad \times \biggl(\int_0^{s\delta^{-2}} f_\delta\bigl((t\delta^{-2} - r'', k), (s\delta^{-2} - r'', \tilde{k})\bigr)\, dr''\biggr)\, ds\, dt \\ &\quad = \int_0^T \int_0^{t\delta^{-2}} \biggl(\int_0^{t\delta^{-2} - s} f_\delta\bigl((s + s', v), (s', \tilde{v})\bigr)\, ds'\biggr) \\ &\qquad \times \biggl(\int_0^{t\delta^{-2} - s} f_\delta\bigl((s + s'', k), (s'', \tilde{k})\bigr)\, ds''\biggr)\, ds\, dt, \end{aligned} \tag{A.21}$$

substituting $ds' = s\delta^{-2} - dr'$, $ds'' = s\delta^{-2} - dr''$ and $ds''' = \delta^{-2}(t - ds)$, but writing again s for s'''. With this we prove now the proposition.

(i) Let $v = \Delta z$, $\tilde{v} = u^{(\delta)}$. By the rescaling with δ^{-6} it is enough to show $V_i = O(\ell_{d,2}(\delta)^3)$, $i = 1, 2$. Observe by Lemma A.6(i), (ii) that

$$\begin{aligned} \bigl|f_\delta\bigl((s + s'', u^{(\delta)}), (s'', u^{(\delta)})\bigr)\bigr| &\lesssim \|S_{\vartheta,\delta}^*(s + s'')u^{(\delta)}\|_{L^2(\Lambda_\delta)} \|S_{\vartheta,\delta}^*(s'')u^{(\delta)}\|_{L^2(\Lambda_\delta)} \\ &\lesssim \bigl(1 \wedge (s + s'')^{-d/4}\bigr)\bigl(1 \wedge (s'')^{-d/4}\bigr) \lesssim 1 \wedge (s'')^{-d/2}, \end{aligned} \tag{A.22}$$

$$\bigl|f_\delta\bigl((s + s', \Delta z), (s', \Delta z)\bigr)\bigr| \lesssim \|S_{\vartheta,\delta}^*(s + s')\Delta z\|_{L^2(\Lambda_\delta)} \|S_{\vartheta,\delta}^*(s')\Delta z\|_{L^2(\Lambda_\delta)}$$

$$\lesssim \big(1 \wedge (s+s')^{-1/2-d/4}\big)\big(1 \wedge (s')^{-1/2-d/4}\big) \tag{A.23}$$
$$\le \big(1 \wedge s^{-1/2-d/4}\big)\big(1 \wedge (s')^{-1/2-d/4}\big).$$

These bounds yield in (A.21) when $d \geq 2$ for $V(\Delta z, \Delta z, u^{(\delta)}, u^{(\delta)})$ up to a constant the upper bound

$$\Big(\int_0^{T\delta^{-2}} \big(1 \wedge (s'')^{-d/2}\big)\, ds''\Big)\Big(\int_0^{T\delta^{-2}} \big(1 \wedge s^{-1/2-d/4}\big)\, ds\Big) \tag{A.24}$$
$$\times \Big(\int_0^{T\delta^{-2}} \big(1 \wedge (s')^{-1/2-d/4}\big)\, ds'\Big) \lesssim \ell_{d,2}(\delta)^3.$$

Similarly, $|f_\delta((s+s'', \Delta z), (s'', u^{(\delta)}))| \lesssim (1 \wedge s^{-d/4})(1 \wedge (s'')^{-1/2-d/4})$, $|f_\delta((s+s', u^{(\delta)}), (s', \Delta z))| \lesssim (1 \wedge s^{-d/4})(1 \wedge (s')^{-1/2-d/4})$, implying the upper bound $\ell_{d,2}(\delta)^3$ also for $V(\Delta z, u^{(\delta)}, u^{(\delta)}, \Delta z)$. In all, we find $|V_1|, |V_2| \lesssim \ell_{d,2}(\delta)^3$.

(ii) Let $v = \tilde{v} = w^{(\delta)}$. By the rescaling it is enough to show $\delta^2 V_1 \to 0$. We have by Lemma A.6(ii)

$$\big|f_\delta\big((s+s'', v), (s'', v)\big)\big| \lesssim 1 \wedge (s'')^{-\alpha'-d/2},$$
$$\big|f_\delta\big((s+s', v), (s', v)\big)\big| \lesssim \big(1 \wedge s^{-\alpha'/2-d/4}\big)\big(1 \wedge (s')^{-\alpha'/2-d/4}\big).$$

Therefore, since $\alpha' > 1/2$ and as in (A.24) but this time for all $d \geq 1$, $V_1 = o(\delta^{-2})$.

(iii) Let $v = \tilde{v} = u^{(\delta)}$. The claim is a direct consequence of (A.22) and (A.24) for $d \geq 2$, with the ds-integral of order $O(\delta^{-2})$ this time.

(iv). Let $v = \tilde{v} = w^{(\delta)}$. Since $V_1 = V_2 = V(w^{(\delta)})$, it is enough to show

$$V_1 = V\big(w^{(\delta)}\big) \to T\vartheta(0)^{-3} \int_0^\infty \Psi\big(e^{s\Delta}\Delta m, \Delta m\big)^2\, ds. \tag{A.25}$$

We argue by dominated convergence. Set

$$f\big((s, \Delta m), (s', \Delta m)\big) = \langle B_0^* e^{\vartheta(0)s\Delta}\Delta m, B_0^* e^{\vartheta(0)s'\Delta}\Delta m\rangle_{L^2(\mathbb{R}^d)}.$$

Exactly as in the proof of Proposition A.8(i) we get pointwise by polarisation

$$f_\delta\big((s+s', w^{(\delta)}), (s', w^{(\delta)})\big) \to f\big((s+s', \Delta m), (s', \Delta m)\big).$$

In order to conclude we need to upper bound the f_δ-terms. For $d \geq 2$ this has been done in (ii). If $\vartheta \in C^{1+\alpha'}(\overline{\Lambda})$ for $\alpha' > 1/2$ and $\int_{\mathbb{R}^d} m(x)\, dx = 0$, the improved bound in Lemma A.6(iii) for (A.23) gives

$$\big|f_\delta\big((s+s', w^{(\delta)}), (s', w^{(\delta)})\big)\big| \lesssim \big(1 \wedge s^{-1/2-\alpha'/2-d/4}\big)\big(1 \wedge (s')^{-1/2-\alpha'/2-d/4}\big).$$

In both cases (A.25) follows from dominated convergence, noting

$$\int_0^\infty f\big((s+s', \Delta m), (s', \Delta m)\big)\, ds' = \vartheta(0)^{-1}\Psi\big(e^{\vartheta(0)s\Delta}\Delta m, \Delta m\big). \qquad \square$$

The next result improves on Proposition A.8(ii) when B is a multiplication operator, by making lower order terms explicit. This is necessary for the proof of Theorem 5.9. The main difficulty is to work around not having a rate of convergence in Proposition 3.5(ii).

PROPOSITION A.10. *Let $z \in H^4(\mathbb{R}^d)$ have compact support in $\Lambda_{\delta'}$ for some $\delta' > 0$ and suppose that $B = M_\sigma$ with $\sigma \in C^1(\overline{\Lambda})$. If $d = 1$, then assume $\vartheta \in C^{1+\alpha'}(\overline{\Lambda})$ for $\alpha' > 1/2$*

and $\int_{\mathbb{R}} z(x)\,dx = 0$. *Then for* $0 < \delta \le \delta'$ *and* $\delta \to 0$

$$\begin{aligned}\delta^{-2} \int_0^T & \operatorname{Var}\big(\langle \tilde{X}(t), (\Delta z)_\delta \rangle\big)\,dt \\ &= \frac{T\sigma^2(0)}{2\vartheta(0)} \|\nabla z\|^2_{L^2(\mathbb{R}^d)} \\ &\quad + \frac{\delta T}{2} \Big\langle \Big\langle \nabla\Big(\frac{\sigma^2}{\vartheta}\Big)(0), x \Big\rangle_{\mathbb{R}^d}, |\nabla z|^2 \Big\rangle_{L^2(\mathbb{R}^d)} + o(\delta).\end{aligned}$$

PROOF. Let $\langle \overline{X}(t), \cdot\rangle$ be defined as $\langle \tilde{X}(t), \cdot\rangle$ in (2.4), but with semigroup $(\overline{S}_\vartheta(t))_{t\ge 0}$ on $L^2(\Lambda)$ generated by $A_\vartheta = \Delta_\vartheta$. As before, $(\overline{S}_{\vartheta,\delta}(t))_{t\ge 0}$ is the corresponding semigroup on $L^2(\Lambda_\delta)$ generated by $\Delta_{\vartheta(\delta\cdot)}$. Note that the $\overline{S}_{\vartheta,\delta}(t)$ are self-adjoint. With $v^{(\delta)} := \delta^{-1}(\Delta_{\vartheta(\delta\cdot)} - \vartheta(0)\Delta)z$ set for $0 \le t \le T$

$$\begin{aligned}
T_1(t) &= -\frac{2\delta^{-1}}{\vartheta^2(0)} \operatorname{Cov}\big(\langle \overline{X}(t), (\Delta_{\vartheta(\delta\cdot)} z)_\delta\rangle, \langle \overline{X}(t), v_\delta^{(\delta)}\rangle\big),\\
T_2(t) &= \frac{\delta^{-2}}{\vartheta^2(0)} \operatorname{Var}\big(\langle \overline{X}(t), (\Delta_{\vartheta(\delta\cdot)} z)_\delta\rangle\big) - \frac{\sigma^2(0)}{\vartheta^2(0)} \int_0^{t\delta^{-2}} \|\overline{S}_{\vartheta,\delta}(s)\Delta_{\vartheta(\delta\cdot)} z\|^2_{L^2(\Lambda_\delta)}\,ds,\\
T_3(t) &= \frac{\sigma^2(0)}{\vartheta^2(0)} \int_0^{t\delta^{-2}} \|\overline{S}_{\vartheta,\delta}(s)\Delta_{\vartheta(\delta\cdot)} z\|^2_{L^2(\Lambda_\delta)}\,ds - \frac{\sigma^2(0)}{2\vartheta(0)} \|\nabla z\|^2_{L^2(\mathbb{R}^d)},\\
R_1(t) &= \delta^{-2} \operatorname{Cov}\big(\langle \tilde{X}(t), (\Delta z)_\delta\rangle - \langle \overline{X}(t), (\Delta z)_\delta\rangle, \langle \tilde{X}(t), (\Delta z)_\delta\rangle + \langle \overline{X}(t), (\Delta z)_\delta\rangle\big),\\
R_2(t) &= \frac{1}{\vartheta^2(0)} \operatorname{Var}\big(\langle \overline{X}(t), v_\delta^{(\delta)}\rangle\big),
\end{aligned}$$

and introduce the decompositions

$$\begin{aligned}
\delta^{-2} \int_0^T & \operatorname{Var}\big(\langle \tilde{X}(t), (\Delta z)_\delta\rangle\big)\,dt \\
&= \delta^{-2} \int_0^T \operatorname{Var}\big(\langle \overline{X}(t), (\Delta z)_\delta\rangle\big)\,dt + \int_0^T R_1(t)\,dt \\
&= \frac{\delta^{-2}}{\vartheta^2(0)} \int_0^T \operatorname{Var}\big(\langle \overline{X}(t), (\Delta_{\vartheta(\delta\cdot)} z)_\delta\rangle\big)\,dt + \int_0^T \big(T_1(t) + R_2(t)\big)\,dt + o(\delta) \\
&= \frac{T\sigma^2(0)}{2\vartheta(0)} \|\nabla z\|^2_{L^2(\mathbb{R}^d)} + \int_0^T \big(T_1(t) + T_2(t) + T_3(t)\big)\,dt + o(\delta),
\end{aligned}$$

where we use for the remainder terms that $\int_0^T R_i(t)\,dt = o(\delta)$, $i = 1, 2$, by Lemmas A.14 and A.15 below. The claim follows from Lemmas A.11, A.12 and A.13 below, which show

$$\delta^{-1} \int_0^T \big(T_1(t) + T_2(t) + T_3(t)\big)\,dt \to \frac{T}{2} \Big\langle \Big\langle \frac{\vartheta(0)}{\sigma^2(0)} \nabla\sigma^2(0) - \nabla\vartheta(0), x \Big\rangle_{\mathbb{R}^d}, |\nabla z|^2 \Big\rangle_{L^2(\mathbb{R}^d)}. \qquad \square$$

LEMMA A.11. *In Proposition* A.10 *we have*

$$\delta^{-1} \int_0^T T_1(t)\,dt \to -\frac{T\sigma^2(0)}{\vartheta^2(0)} \langle \langle \nabla\vartheta(0), x\rangle_{\mathbb{R}^d}, |\nabla z|^2 \rangle_{L^2(\mathbb{R}^d)}.$$

PROOF. Lemma A.5 above with $A_\vartheta = \Delta_\vartheta$ yields $v^{(\delta)} \to v := \Delta(\langle \nabla\vartheta(0), x\rangle_{\mathbb{R}^d} z) - \langle \nabla\vartheta(0), \nabla z\rangle_{\mathbb{R}^d}$ in L^2. Moreover, since $\vartheta \in C^1(\overline{\Lambda})$, we have $\|\Delta_{\vartheta(\delta\cdot)} z - \vartheta(0)\Delta z\|_{L^1\cap L^2(\mathbb{R}^d)} \le$

$C\delta$. If $d \geq 2$, then Proposition A.8(iii) with $w^{(\delta)} = \vartheta(0)^{-1}\Delta_{\vartheta(\delta\cdot)}z$, $u^{(\delta)} = v^{(\delta)}$ already implies $\delta^{-1}\int_0^T T_1(t)\,dt \to -2T\vartheta^{-2}(0)\Psi(\Delta z, v)$, and the claim follows from Lemma A.3 above, recalling the identity $\Psi(\Delta z, v) = -\frac{\sigma^2(0)}{2}\langle z, v\rangle_{L^2(\mathbb{R}^d)}$ from (3.3). For $d = 1$, on the other hand, the properties $\vartheta \in C^{1+\alpha'}(\overline{\Lambda})$ for $\alpha' > 1/2$ and $\int_{\mathbb{R}} z(x)\,dx = 0$ ensure by Lemma A.5(ii) that $v = \Delta m$ for a compactly supported $m \in H^2(\mathbb{R})$ with $\|v^{(\delta)} - \Delta m\|_{L^1\cap L^2(\mathbb{R})} \leq C\delta^{\alpha'}$. By polarisation and Proposition A.8(ii) (with $w^{(\delta)} = \vartheta(0)^{-1}\Delta_{\vartheta(\delta\cdot)}z$ and $w^{(\delta)} = v^{(\delta)}$) $\delta^{-1}\int_0^T T_1(t)\,dt$ converges again to the claimed limit. □

LEMMA A.12. *In Proposition A.10 we have*

$$\delta^{-1}\int_0^T T_2(t)\,dt \to \frac{T}{2\vartheta(0)}\langle\langle\nabla\sigma^2(0), x\rangle_{\mathbb{R}^d}, |\nabla z|^2\rangle_{L^2(\mathbb{R}^d)}.$$

PROOF. Using (3.4), we have $\vartheta^2(0)T_2(t) = \int_0^{t\delta^{-2}} f_\delta(s)\,ds$ for

$$\begin{aligned} f_\delta(s) &= \langle(\sigma^2(\delta\cdot) - \sigma^2(0))\overline{S}_{\vartheta,\delta}(s)\Delta_{\vartheta(\delta\cdot)}z, \overline{S}_{\vartheta,\delta}(s)\Delta_{\vartheta(\delta\cdot)}z\rangle_{L^2(\Lambda_\delta)} \\ &= \delta\int_0^1 \langle\langle\nabla\sigma^2(\delta r x), x\rangle_{\mathbb{R}^d}\overline{S}_{\vartheta,\delta}(s)\Delta_{\vartheta(\delta\cdot)}z, \overline{S}_{\vartheta,\delta}(s)\Delta_{\vartheta(\delta\cdot)}z\rangle_{L^2(\Lambda_\delta)}\,dr. \end{aligned}$$

By the Cauchy–Schwarz inequality and the semigroup bounds in Proposition A.4(i), (ii) above this means

$$\begin{aligned} |\delta^{-1}f_\delta(s)| &\lesssim \||x|\overline{S}_{\vartheta,\delta}(s)\Delta_{\vartheta(\delta\cdot)}z\|_{L^2(\Lambda_\delta)}(1\wedge s^{-1}) \\ &\lesssim \||x|^2\overline{S}_{\vartheta,\delta}(s)\Delta_{\vartheta(\delta\cdot)}z\|_{L^2(\Lambda_\delta)}^{1/2}(1\wedge s^{-3/2}) \\ &\lesssim \||x|^2 e^{\vartheta(0)s\Delta}|\Delta_{\vartheta(\delta\cdot)}z|\|_{L^2(\mathbb{R}^d)}^{1/2}(1\wedge s^{-3/2}) \\ &\lesssim (1\vee s)^{1/2}(1\wedge s^{-d/4})^{1/2}(1\wedge s^{-3/2}) \lesssim (1\wedge s^{-1-d/8}), \end{aligned} \tag{A.26}$$

where we used Proposition 3.5(i) for an approximating sequence of $\Delta_{\vartheta(\delta\cdot)}z$ with continuous functions and Lemma A.2(iii) in the last two lines. We conclude from Proposition 3.5(ii) that

$$\delta^{-1}f_\delta(s) \to f(s) := \vartheta^2(0)\langle\langle\nabla\sigma^2(0), x\rangle_{\mathbb{R}^d}e^{\vartheta(0)s\Delta}\Delta z, e^{\vartheta(0)s\Delta}\Delta z\rangle_{L^2(\mathbb{R}^d)}.$$

Combining this with (A.26), the dominated convergence theorem shows $\vartheta^2(0) \times \delta^{-1}\int_0^T T_2(t)\,dt \to T\int_0^\infty f(s)\,ds$. For the result note that by Lemmas A.2(ii) and A.3 (here we need $z \in H^4(\mathbb{R}^d)$) $\vartheta^{-2}(0)\int_0^\infty f(s)\,ds$ equals

$$\begin{aligned} &\int_0^\infty \langle\langle\nabla\sigma^2(0), -2\vartheta(0)s\nabla e^{\vartheta(0)s\Delta}\Delta z + e^{\vartheta(0)s\Delta}(x\Delta z)\rangle_{\mathbb{R}^d}, e^{\vartheta(0)s\Delta}\Delta z\rangle_{L^2(\mathbb{R}^d)}\,ds \\ &= \int_0^\infty \langle\langle\nabla\sigma^2(0), x\rangle_{\mathbb{R}^d}\Delta z, e^{2\vartheta(0)s\Delta}\Delta z\rangle_{L^2(\mathbb{R}^d)}\,ds \\ &= -\frac{1}{2\vartheta(0)}\langle\langle\nabla\sigma^2(0), x\rangle_{\mathbb{R}^d}z, \Delta z\rangle_{L^2(\mathbb{R}^d)} = \frac{1}{2\vartheta(0)}\langle\langle\nabla\sigma^2(0), x\rangle_{\mathbb{R}^d}, |\nabla z|^2\rangle_{L^2(\mathbb{R}^d)}. \end{aligned}$$ □

LEMMA A.13. *In Proposition A.10 we have*

$$\delta^{-1}\int_0^T T_3(t)\,dt \to \frac{T\sigma^2(0)}{2\vartheta^2(0)}\langle\langle\nabla\vartheta(0), x\rangle_{\mathbb{R}^d}, |\nabla z|^2\rangle_{L^2(\mathbb{R}^d)}.$$

PROOF. Since $\overline{S}_{\vartheta,\delta}(s)$ is self-adjoint, we have

$$\frac{2\vartheta^2(0)}{\sigma^2(0)}T_3(t) = \langle \Delta_{\vartheta(\delta\cdot)}\overline{S}_{\vartheta,\delta}(2t\delta^{-2})z, z\rangle_{L^2(\Lambda_\delta)} - \langle (\Delta_{\vartheta(\delta\cdot)} - \vartheta(0)\Delta)z, z\rangle_{L^2(\mathbb{R}^d)}.$$

Integrating over $0 \le t \le T$ and using the semigroup bound in Proposition A.4(i) the first term is of order $O(\delta^2)$. Since $\vartheta \in C^1(\overline{\Lambda})$, the result follows from Lemma A.3 and

$$\begin{aligned}(\Delta_{\vartheta(\delta\cdot)} - \vartheta(0)\Delta)z &= \delta\left(\int_0^1 \langle \nabla\vartheta(s\delta\cdot), x\rangle_{\mathbb{R}^d}\,ds\right)\Delta z + \langle \nabla\vartheta(\delta\cdot), \nabla z\rangle_{\mathbb{R}^d} \\ &= \delta\langle \nabla\vartheta(0), x\rangle_{\mathbb{R}^d}\Delta z + \langle \nabla\vartheta(0), \nabla z\rangle_{\mathbb{R}^d} + o(\delta). \qquad \square\end{aligned}$$

LEMMA A.14. *In Proposition A.10 we have* $\delta^{-1}\int_0^T R_1(t)\,dt \to 0$.

PROOF. By (3.4) write $R_1(t) = \int_0^{t\delta^{-2}} f_\delta(s)\,ds$ with $f_\delta(s) = \langle g_\delta(s), h_\delta(s)\rangle_{L^2(\Lambda_\delta)}$ for $s \ge 0$, where

$$\begin{aligned}g_\delta(s) &= (S^*_{\vartheta,\delta}(s) - \overline{S}_{\vartheta,\delta}(s))\Delta z, \\ h_\delta(s) &= \sigma^2(\delta\cdot)(S^*_{\vartheta,\delta}(s) + \overline{S}_{\vartheta,\delta}(s))\Delta z.\end{aligned}$$

An application of the dominated convergence theorem then proves the result, if

$$|\delta^{-1}f_\delta(s)| \lesssim 1 \wedge s^{-3/2}, \quad 0 < s \le t\delta^{-2}, \tag{A.27}$$

$$\delta^{-1}f_\delta(s) \to 0. \tag{A.28}$$

In order to show (A.27) and (A.28) we use the *variation of parameters formula* (Engel and Nagel [10], p. 162): The function $[0,s] \ni s' \mapsto S^*_{\vartheta,\delta}(s')\overline{S}_{\vartheta,\delta}(s-s')\Delta z$ has derivative $S^*_{\vartheta,\delta}(s')(A^*_{\vartheta,\delta} - \Delta_{\vartheta(\delta\cdot)})\overline{S}_{\vartheta,\delta}(s-s')\Delta z$, implying

$$g_\delta(s) = \int_0^s S^*_{\vartheta,\delta}(s')(A^*_{\vartheta,\delta} - \Delta_{\vartheta(\delta\cdot)})\overline{S}_{\vartheta,\delta}(s-s')\Delta z\,ds'.$$

Since the operator $A^*_{\vartheta,\delta} - \Delta_{\vartheta(\delta\cdot)} = A^*_{0,\delta}$ is not bounded, a careful analysis is required. Decomposing it into first and zero order terms we have

$$\begin{aligned}g_\delta(s) = &-\delta\int_0^s S^*_{\vartheta,\delta}(s')\langle a(\delta\cdot), \nabla\overline{S}_{\vartheta,\delta}(s-s')\Delta z\rangle_{\mathbb{R}^d}\,ds' \\ &+ \delta^2\int_0^s S^*_{\vartheta,\delta}(s')(b(\delta\cdot) - (\operatorname{div} a)(\delta\cdot))\overline{S}_{\vartheta,\delta}(s-s')\Delta z\,ds'.\end{aligned} \tag{A.29}$$

The semigroup bounds in Proposition A.4(ii) and in Lemma A.6(i), (ii), (iii) above, subject to $d \ge 2$ or $d = 1$ with $\vartheta \in C^{1+\alpha'}(\overline{\Lambda})$ for $\alpha' > 1/2$ and $\int_{\mathbb{R}} z(x)\,dx = 0$, show for sufficiently small δ and $0 \le s' < s \le t\delta^{-2}$ that

$$\begin{aligned}&\|\Delta_{\vartheta,\delta}\overline{S}_{\vartheta,\delta}(s-s')\Delta z\|_{L^2(\Lambda_\delta)} \\ &\quad = \left\|\Delta_{\vartheta,\delta}\overline{S}_{\vartheta,\delta}\left(\frac{s-s'}{2}\right)\overline{S}_{\vartheta,\delta}\left(\frac{s-s'}{2}\right)\Delta z\right\|_{L^2(\Lambda_\delta)} \\ &\quad \lesssim (s-s')^{-1}\left\|\overline{S}_{\vartheta,\delta}\left(\frac{s-s'}{2}\right)\Delta z\right\|_{L^2(\Lambda_\delta)} \lesssim (s-s')^{-1}(1 \wedge (s-s')^{-1}).\end{aligned}$$

Hence, the computations in (A.10) above show

$$
\begin{aligned}
&\|\langle a(\delta\cdot), \nabla\overline{S}_{\vartheta,\delta}(s-s')\Delta z\rangle_{\mathbb{R}^d}\|_{L^2(\Lambda_\delta)} \\
&\quad\lesssim \|\Delta_{\vartheta(\delta\cdot)}\overline{S}_{\vartheta,\delta}(s-s')\Delta z\|_{L^2(\Lambda_\delta)}^{1/2}\|\overline{S}_{\vartheta,\delta}(s-s')\Delta z\|_{L^2(\Lambda_\delta)}^{1/2} \\
&\quad\lesssim (s-s')^{-1/2}\big(1\wedge (s-s')^{-1}\big).
\end{aligned}
$$

Because of Proposition A.4(i) this yields for $0 < s \le t\delta^{-2}$

$$
\begin{aligned}
\|\delta^{-1}g_\delta(s)\|_{L^2(\Lambda_\delta)} &\lesssim \int_0^s \big((s-s')^{-1/2}\big(1\wedge (s-s')^{-1}\big) + \delta\big(1\wedge (s-s')^{-1}\big)\big)\,ds' \\
&\lesssim \int_0^s (s')^{-1/2}\big(1\wedge (s')^{-1}\big)\,ds' + t^{1/4}\delta^{1/2}\int_0^s \big(1\wedge (s')^{-1-1/4}\big)\,ds'.
\end{aligned}
$$

In all, this is of order $1\wedge s^{-1/2}$. Since also $\|h_\delta(s)\|_{L^2(\Lambda_\delta)} \lesssim 1\wedge s^{-1}$ by Lemma A.6(ii), (iii), subject to $d\ge 2$ or the conditions in $d=1$, we obtain (A.27). With respect to (A.28) fix s and observe by the convergence of the semigroups in Proposition 3.5(ii) above and $\sigma^2(\delta\cdot)\to\sigma^2(0)$ that $h_\delta(s)\to 2\sigma^2(0)e^{\vartheta(0)s\Delta}\Delta z$ in $L^2(\mathbb{R}^d)$. Therefore, $f_\delta(s) = -2\delta\sigma^2(0)f_\delta^{(1)}(s) + o(\delta)$, uniformly in s, for

$$
\begin{aligned}
f_\delta^{(1)}(s) &= \int_0^s \langle S^*_{\vartheta,\delta}(s')\langle a(\delta\cdot), \nabla\overline{S}_{\vartheta,\delta}(s-s')\Delta z\rangle_{\mathbb{R}^d}, e^{\vartheta(0)s\Delta}\Delta z\rangle_{L^2(\Lambda_\delta)}\,ds' \\
&= \sum_{i=1}^d \int_0^s \langle \partial_i\overline{S}_{\vartheta,\delta}(s-s')\Delta z, a_i(\delta\cdot)S_{\vartheta,\delta}(s')\big(e^{\vartheta(0)s\Delta}\Delta z\big)|_{\Lambda_\delta}\rangle_{L^2(\Lambda_\delta)}\,ds'.
\end{aligned}
$$

In the same way, since $a_i(\delta\cdot)S_{\vartheta,\delta}(s')(v|_{\Lambda_\delta}) \to a_i(0)e^{\vartheta(0)s'\Delta}v$ for $v\in L^2(\mathbb{R}^d)$ by Proposition 3.5(ii), we have $f_\delta^{(1)}(s) = f_\delta^{(2)}(s) + o(\delta)$ for

$$
\begin{aligned}
f_\delta^{(2)}(s) &= \sum_{i=1}^d a_i(0)\int_0^s \langle \partial_i\overline{S}_{\vartheta,\delta}(s-s')\Delta z, e^{\vartheta(0)(s'+s)\Delta}\Delta z\rangle_{L^2(\Lambda_\delta)}\,ds' \\
&= -\sum_{i=1}^d a_i(0)\int_0^s \langle \overline{S}_{\vartheta,\delta}(s-s')\Delta z, e^{\vartheta(0)(s'+s)\Delta}\Delta\partial_i z\rangle_{L^2(\Lambda_\delta)}\,ds'.
\end{aligned}
$$

Noting that $\overline{S}_{\vartheta,\delta}(s-s')\Delta z \to e^{\vartheta(0)(s-s')\Delta}\Delta z$ in $L^2(\mathbb{R}^d)$, we finally obtain

$$
\begin{aligned}
\delta^{-1}f_\delta(s) &\to 2\sigma^2(0)\sum_{i=1}^d a_i(0)\int_0^s \langle e^{\vartheta(0)(s-s')\Delta}\Delta z, e^{\vartheta(0)(s'+s)\Delta}\Delta\partial_i z\rangle_{L^2(\Lambda_\delta)}\,ds' \\
&= \frac{2\sigma^2(0)s}{\vartheta(0)}\sum_{i=1}^d a_i(0)\langle\Delta z, e^{2\vartheta(0)s\Delta}\Delta\partial_i z\rangle_{L^2(\mathbb{R}^d)},
\end{aligned}
$$

which vanishes according to Lemma A.3. □

LEMMA A.15. *In Proposition A.10 we have* $\delta^{-1}\int_0^T R_2(t)\,dt \to 0$.

PROOF. By (3.4) we have for $0\le t\le T$

$$
R_2(t) \lesssim \delta^2\int_0^{t\delta^{-2}} \|\overline{S}_{\vartheta,\delta}(s)v^{(\delta)}\|_{L^2(\Lambda_\delta)}^2\,ds.
$$

Recall from Lemma A.11 that $v^{(\delta)}$ converges in $L^2(\mathbb{R}^d)$. For $d \geq 2$, Lemma A.6(i) then shows uniformly in $0 \leq t \leq T$ that

$$\delta^{-1} R_2(t) \lesssim \delta \int_0^{t\delta^{-2}} 1 \wedge s^{-1} \, ds \lesssim \delta^{1/2} \int_0^{\infty} 1 \wedge \left(s^{-1-1/4}\right) ds \to 0,$$

implying the claim in this case. For $d = 1$ it is enough to recall from Lemma A.11 that $\|v^{(\delta)} - \Delta m\|_{L^1 \cap L^2(\mathbb{R})} \leq C\delta^{\alpha'}$, and so the claim follows from Lemma A.6(iii). □

PROPOSITION A.16. *Grant Assumption 3.2. Let $z \in H^2(\mathbb{R}^d)$ have compact support in $\Lambda_{\delta'}$ for some $\delta' > 0$ and for $d = 1$ assume $\vartheta \in C^{1+\alpha'}(\overline{\Lambda})$ for $\alpha' > 1/2$, $\int_{\mathbb{R}} z(x)\,dx = 0$. For $0 < \delta \leq \delta'$ set $\xi_\delta(t) = \langle \tilde{X}(t), (\Delta z)_\delta \rangle$. Then the fourth moment of $\delta^{-3} \int_0^T (\xi_\delta(t)^2 - \mathbb{E}[\xi_\delta(t)^2])\,dt$ converges, with Ψ from (3.2), for $\delta \to 0$ to*

$$3\left(4T\vartheta(0)^{-3} \int_0^{\infty} \Psi\left(e^{s\Delta}\Delta z, \Delta z\right)^2 ds\right)^2.$$

PROOF. In view of Proposition A.9(iv) it is enough to show

$$\begin{aligned} R_\delta &:= \mathbb{E}\left[\left(\int_0^T (\xi_\delta(t)^2 - \mathbb{E}[\xi_\delta(t)^2])\,dt\right)^4\right] - 3\operatorname{Var}\left(\int_0^T \xi_\delta(t)^2\,dt\right)^2 \\ &= o\left(\operatorname{Var}\left(\int_0^T \xi_\delta(t)^2\,dt\right)^2\right) = o\left(\delta^{12}\right). \end{aligned}$$

Abbreviating $c_\delta(t,s) = \operatorname{Cov}(\langle \tilde{X}(t), (\Delta z)_\delta \rangle, \langle \tilde{X}(s), (\Delta z)_\delta \rangle)$, recall from (A.19) that

$$3\operatorname{Var}\left(\int_0^T \xi_\delta(t)^2\,dt\right)^2 = 3\left(\int_0^T \int_0^T 2c_\delta(t,s)^2\,dt\,ds\right)^2.$$

Wick's formula (Janson [20], Theorem 1.28) for 8th centered Gaussian moments $\mathbb{E}[\prod_{i=1}^8 Z_i] = \sum_{\pi \in \Pi_2(8)} \prod_{(i,j)\in\pi} \mathbb{E}[Z_i Z_j]$ applied to $Z_i = \xi_\delta(t_i)$ for $0 \leq t_i \leq T$, with $\Pi_2(8)$ being the set of all partitions π of $\{1, \dots, 8\}$ into 2-tuples, therefore yields, using the symmetry of the integrand in (t_1, t_2, t_3, t_4),

$$R_\delta = 48 \int_0^T \int_0^T \int_0^T \int_0^T c_\delta(t_1,t_2) c_\delta(t_2,t_3) c_\delta(t_3,t_4) c_\delta(t_4,t_1)\,dt_1\,dt_2\,dt_3\,dt_4.$$

The calculations in the proof of Proposition A.9(iv) show for $s \leq t$, both when $d \geq 2$ and when $d = 1$, $\vartheta \in C^{1+\alpha'}(\overline{\Lambda})$ for $\alpha' > 1/2$, $\int_{\mathbb{R}} z(x)\,dx = 0$, that

$$\begin{aligned} |c_\delta(t,s)| &= \left| \int_0^s f_\delta\left(\left(\delta^{-2}(t-s'), \Delta z\right), \left(\delta^{-2}(s-s'), \Delta z\right)\right) ds' \right| \\ &= \delta^2 \left| \int_0^{s\delta^{-2}} f_\delta\left(\left(\delta^{-2}(t-s) + s', \Delta z\right), \left(s', \Delta z\right)\right) ds' \right| \\ &\lesssim \delta^2 \int_0^{s\delta^{-2}} \left(1 \wedge \left(\delta^{-2}|t-s| + s'\right)^{-1}\right)\left(1 \wedge (s')^{-1}\right) ds' \\ &\lesssim \delta^2 \ell_{d,2}\left(1 \wedge \left(\delta^{-2}|t-s|\right)^{-1}\right) =: \bar{c}_\delta\left(\delta^{-2}(t-s)\right), \end{aligned} \tag{A.30}$$

so that as in (A.21), substituting $s_i = \delta^{-2}(t_{i+1} - t_i)$,

$$\begin{aligned} R_\delta &\lesssim \delta^6 \int_0^T \int_0^{T\delta^{-2}} \int_0^{T\delta^{-2}} \int_0^{T\delta^{-2}} \bar{c}_\delta(s_1)\bar{c}_\delta(s_2)\bar{c}_\delta(s_3) \\ &\qquad \times \bar{c}_\delta(s_1+s_2+s_3)\,dt_1\,ds_1\,ds_2\,ds_3 \\ &\lesssim T\delta^{14}\ell_{d,2}^4 \left(\int_0^{T\delta^{-2}} \left(1 \wedge s^{-1}\right) ds\right)^3 = o\left(\delta^{12}\right). \end{aligned}$$

□

Acknowledgments. We are grateful to Sylvie Roelly, Nicolas Perkowski, Josef Janák and two anonymous referees for very helpful comments and questions. This research has been partially funded by Deutsche Forschungsgemeinschaft (DFG) - SFB1294/1 - 318763901.

REFERENCES

[1] ADAMS, R. A. and FOURNIER, J. J. F. (2003). *Sobolev Spaces*, 2nd ed. *Pure and Applied Mathematics* (*Amsterdam*) **140**. Elsevier/Academic Press, Amsterdam. MR2424078

[2] AIHARA, S. and BAGCHI, A. (1989). Infinite-dimensional parameter identification for stochastic parabolic systems. *Statist*. *Probab*. *Lett*. **8** 279–287. MR1024040 https://doi.org/10.1016/0167-7152(89)90134-X

[3] AIHARA, S.-I. and SUNAHARA, Y. (1988). Identification of an infinite-dimensional parameter for stochastic diffusion equations. *SIAM J*. *Control Optim*. **26** 1062–1075. MR0957653 https://doi.org/10.1137/0326058

[4] BIBINGER, M. and TRABS, M. (2020). Volatility estimation for stochastic PDEs using high-frequency observations. *Stochastic Process*. *Appl*. **130** 3005–3052. MR4080737 https://doi.org/10.1016/j.spa.2019.09.002

[5] CIALENCO, I. (2018). Statistical inference for SPDEs: An overview. *Stat*. *Inference Stoch*. *Process*. **21** 309–329. MR3824970 https://doi.org/10.1007/s11203-018-9177-9

[6] CIALENCO, I. and GLATT-HOLTZ, N. (2011). Parameter estimation for the stochastically perturbed Navier–Stokes equations. *Stochastic Process*. *Appl*. **121** 701–724. MR2770904 https://doi.org/10.1016/j.spa.2010.12.007

[7] CIALENCO, I. and HUANG, Y. (2020). A note on parameter estimation for discretely sampled SPDEs. *Stoch*. *Dyn*. **20** 2050016, 28. MR4101083 https://doi.org/10.1142/S0219493720500161

[8] CONT, R. (2005). Modeling term structure dynamics: An infinite dimensional approach. *Int*. *J*. *Theor. Appl*. *Finance* **8** 357–380. MR2144706 https://doi.org/10.1142/S0219024905003049

[9] DA PRATO, G. and ZABCZYK, J. (2014). *Stochastic Equations in Infinite Dimensions*, 2nd ed. *Encyclopedia of Mathematics and Its Applications* **152**. Cambridge Univ. Press, Cambridge. MR3236753 https://doi.org/10.1017/CBO9781107295513

[10] ENGEL, K.-J. and NAGEL, R. (2000). *One-Parameter Semigroups for Linear Evolution Equations*. *Graduate Texts in Mathematics* **194**. Springer, New York. MR1721989

[11] EVANS, L. C. (2010). *Partial Differential Equations*, 2nd ed. *Graduate Studies in Mathematics* **19**. Amer. Math. Soc., Providence, RI. MR2597943 https://doi.org/10.1090/gsm/019

[12] FERGUSON, T. S. (1996). *A Course in Large Sample Theory*. *Texts in Statistical Science Series*. CRC Press, London. MR1699953 https://doi.org/10.1007/978-1-4899-4549-5

[13] FRANKIGNOUL, C. (1985). Sea surface temperature anomalies, planetary waves, and air-sea feedback in the middle latitudes. *Reviews of Geophysics* **23** 357.

[14] FRIEDMAN, A. (1975). *Stochastic Differential Equations and Applications*. *Vol*. 1. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London. MR0494490

[15] HAIRER, M. (2009). An introduction to stochastic PDEs. Preprint. Available at arXiv:0907.4178.

[16] HILDEBRANDT, F. and TRABS, M. (2019). Parameter estimation for SPDEs based on discrete observations in time and space. Preprint. Available at arXiv:1710.01649.

[17] HÜBNER, M., KHASMINSKII, R. and ROZOVSKII, B. L. (1993). Two examples of parameter estimation for stochastic partial differential equations. In *Stochastic Processes* 149–160. Springer, New York. MR1427311

[18] HUEBNER, M. and LOTOTSKY, S. (2000). Asymptotic analysis of a kernel estimator for parabolic SPDE's with time-dependent coefficients. *Ann*. *Appl*. *Probab*. **10** 1246–1258. MR1810873 https://doi.org/10.1214/aoap/1019487615

[19] HUEBNER, M. and ROZOVSKIĬ, B. L. (1995). On asymptotic properties of maximum likelihood estimators for parabolic stochastic PDE's. *Probab*. *Theory Related Fields* **103** 143–163. MR1355054 https://doi.org/10.1007/BF01204212

[20] JANSON, S. (1997). *Gaussian Hilbert Spaces*. *Cambridge Tracts in Mathematics* **129**. Cambridge Univ. Press, Cambridge. MR1474726 https://doi.org/10.1017/CBO9780511526169

[21] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd ed. *Probability and Its Applications* (*New York*). Springer, New York. MR1876169 https://doi.org/10.1007/978-1-4757-4015-8

[22] KARATZAS, I. and SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. Springer, New York. MR1121940 https://doi.org/10.1007/978-1-4612-0949-2

[23] KUTOYANTS, Y. A. (2004). *Statistical Inference for Ergodic Diffusion Processes*. *Springer Series in Statistics*. Springer London, Ltd., London. MR2144185 https://doi.org/10.1007/978-1-4471-3866-2

[24] LIPTSER, R. S. and SHIRYAEV, A. N. (2001). *Statistics of Random Processes*. *I*: *General Theory*, expanded ed. *Applications of Mathematics* (*New York*) **5**. Springer, Berlin. MR1800857

[25] LIU, W. and RÖCKNER, M. (2015). *Stochastic Partial Differential Equations*: *An Introduction*. *Universitext*. Springer, Cham. MR3410409 https://doi.org/10.1007/978-3-319-22354-4

[26] LORD, G. J., POWELL, C. E. and SHARDLOW, T. (2014). *An Introduction to Computational Stochastic PDEs*. *Cambridge Texts in Applied Mathematics*. Cambridge Univ. Press, New York. MR3308418 https://doi.org/10.1017/CBO9781139017329

[27] LUNARDI, A. (1995). *Analytic Semigroups and Optimal Regularity in Parabolic Problems*. *Modern Birkhäuser Classics*. Birkhäuser/Springer Basel AG, Basel. MR3012216

[28] MASLOWSKI, B. and TUDOR, C. A. (2013). Drift parameter estimation for infinite-dimensional fractional Ornstein–Uhlenbeck process. *Bull*. *Sci*. *Math*. **137** 880–901. MR3116218 https://doi.org/10.1016/j.bulsci.2013.04.008

[29] MÖRTERS, P. and PERES, Y. (2010). *Brownian Motion*. *Cambridge Series in Statistical and Probabilistic Mathematics* **30**. Cambridge Univ. Press, Cambridge. MR2604525 https://doi.org/10.1017/CBO9780511750489

[30] NUALART, D. and PECCATI, G. (2005). Central limit theorems for sequences of multiple stochastic integrals. *Ann*. *Probab*. **33** 177–193. MR2118863 https://doi.org/10.1214/009117904000000621

[31] PASEMANN, G. and STANNAT, W. (2020). Drift estimation for stochastic reaction-diffusion systems. *Electron*. *J*. *Stat*. **14** 547–579. MR4056266 https://doi.org/10.1214/19-EJS1665

[32] PAZY, A. (1983). *Semigroups of Linear Operators and Applications to Partial Differential Equations*. *Applied Mathematical Sciences* **44**. Springer, New York. MR0710486 https://doi.org/10.1007/978-1-4612-5561-1

[33] RUDIN, W. (1991). *Functional Analysis*, 2nd ed. *International Series in Pure and Applied Mathematics*. McGraw-Hill, Inc., New York. MR1157815

[34] SHEU, S. J. (1991). Some estimates of the transition density of a nondegenerate diffusion Markov process. *Ann*. *Probab*. **19** 538–561. MR1106275

[35] SIGRIST, F., KÜNSCH, H. R. and STAHEL, W. A. (2015). Stochastic partial differential equation based modelling of large space–time data sets. *J*. *R*. *Stat*. *Soc*. *Ser*. *B*. *Stat*. *Methodol*. **77** 3–33. MR3299397 https://doi.org/10.1111/rssb.12061

[36] TSYBAKOV, A. B. (2009). *Introduction to Nonparametric Estimation*. *Springer Series in Statistics*. Springer, New York. MR2724359 https://doi.org/10.1007/b13794

[37] WALSH, J. B. (1981). A stochastic model of neural response. *Adv*. *in Appl*. *Probab*. **13** 231–281. MR0612203 https://doi.org/10.2307/1426683

[38] WILKINSON, D. R. and EDWARDS, S. F. (1982). Spontaneous interparticle percolation. *Proc*. *R*. *Soc*. *Lond*. *Ser*. *A* **381** 33–51. MR0661715 https://doi.org/10.1098/rspa.1982.0057