On statistical Calderón problems

Kweku Abraham and Richard Nickl

Abstract. For D a bounded domain in $\mathbb{R}^d$, $d \geq 2$, with smooth boundary ∂D, the non-linear inverse problem of recovering the unknown conductivity γ determining solutions $u = u_{\gamma, f}$ of the partial differential equation

$$\nabla \cdot (\gamma \nabla u) = 0 \quad \text{in } D,$$
$$u = f \quad \text{on } \partial D,$$

from noisy observations Y of the Dirichlet-to-Neumann map

$$f \mapsto \Lambda_\gamma(f) = \gamma \frac{\partial u_{\gamma, f}}{\partial \nu}\Big|_{\partial D},$$

with $\partial/\partial \nu$ denoting the outward normal derivative, is considered. The data Y consists of Λ_γ corrupted by additive Gaussian noise at noise level $\varepsilon > 0$, and a statistical algorithm $\hat{\gamma}(Y)$ is constructed which is shown to recover γ in supremum-norm loss at a statistical convergence rate of the order $\log(1/\varepsilon)^{-\delta}$ as $\varepsilon \to 0$. It is further shown that this convergence rate is optimal, up to the precise value of the exponent $\delta > 0$, in an information theoretic sense. The estimator $\hat{\gamma}(Y)$ has a Bayesian interpretation in terms of the posterior mean of a suitable Gaussian process prior and can be computed by MCMC methods.

Mathematics Subject Classification (2010). 35J25, 35R30, 62G05.

Keywords. Non-linear inverse problems, elliptic partial differential equations, electrical impedance tomography, asymptotics of nonparametric Bayes procedures.

Contents

1. Introduction

Let $D \subset \mathbb{R}^d, d \geq 2$, be a bounded domain, which we understand here to be a connected open set with smooth boundary ∂D. For $\gamma \colon D \to (0, \infty)$ a *conductivity coefficient*, consider solutions u to the *Dirichlet problem*

$$\begin{aligned} \nabla \cdot (\gamma \, \nabla \, u) &= 0 \quad \text{in } D, \\ u &= f \quad \text{on } \partial D, \end{aligned} \tag{1}$$

where ∇ denotes the usual gradient operator and where $f \colon \partial D \to \mathbb{C}$ prescribes some boundary values. The parameter spaces considered in the sequel are of the form

$$\Gamma_{m,D'} = \{ \gamma \in C(D) : \inf_{x \in D} \gamma(x) \geq m, \ \gamma = 1 \text{ on } D \setminus D' \}, \tag{2}$$

$$\Gamma_{m,D'}^\alpha(M) = \{ \gamma \in \Gamma_{m,D'} : \|\gamma\|_{H^\alpha(D)} \leq M \}, \quad M > 0, \tag{3}$$

where $m \in (0, 1)$ is a fixed constant, D' is a domain compactly supported in D (that is, its closure $\bar{D}'$ is contained in D), and $\alpha \geq 0$ measures the regularity of γ in the Sobolev scale. The Sobolev spaces $H^\alpha(D), H^\alpha(\partial D)$ of complex-valued functions (and variations thereof) are defined in detail in Appendix A; the standard $L^2(D), L^2(\partial D)$ Lebesgue spaces arise as the case $\alpha = 0$, with inner products $\langle \cdot, \cdot \rangle_{L^2(D)}, \langle \cdot, \cdot \rangle_{L^2(\partial D)}$, respectively, and $C(D)$ denotes the space of bounded continuous *real-valued* functions on D, equipped with the sup-norm $\| \cdot \|_\infty$. Except where otherwise stated, all integrals are taken with respect to Lebesgue and surface measures on D and ∂D respectively.

The elliptic partial differential equation (PDE) in (1) has, for $\gamma \in \Gamma_{m,D'}$ and $f \in H^{s+1}(\partial D)/\mathbb{C}, s \in \mathbb{R}$, a unique weak solution $u_{\gamma,f}$ in the space

$$\mathscr{H}_s := \left(H^{\min\{1,s+3/2\}}(D) \cap H^1_{\text{loc}}(D) \right)/\mathbb{C};$$

that is (for $\overline{v}$ denoting the complex conjugate of v), the equations

$$\int_D \gamma \, \nabla u \cdot \nabla \overline{v} = 0 \quad \forall v \in H_0^1(D),$$
$$u = f \quad \text{on } \partial D, \tag{4}$$

hold simultaneously (for $u \in \mathcal{H}_s$) if and only if $u = u_{\gamma,f}$, where the boundary values of u are defined in a trace sense. Here and below $/\mathbb{C}$ means that we identify functions $f, f + c$ which are equal up to a scalar $c \in \mathbb{C}$. See Lemma 19 in Appendix C and its proof for details.

Given a solution $u_{\gamma,f}$ to the Dirichlet problem, one can measure the *Neumann (boundary) data*

$$\gamma \frac{\partial u_{\gamma,f}}{\partial \nu}\Big|_{\partial D} \equiv \frac{\partial u_{\gamma,f}}{\partial \nu}\Big|_{\partial D}, \quad \gamma \in \Gamma_{m,D'},$$

where $\frac{\partial}{\partial \nu}$ denotes the outward normal derivative on ∂D (again to be understood in a trace sense). It can be shown (see Lemma 20) that for any $s \in \mathbb{R}$ and any $f \in H^{s+1}(\partial D)/\mathbb{C}$, the Neumann data lies in the space

$$H_\diamond^s(\partial D) := \big\{ g \in H^s(\partial D) : \langle g, 1 \rangle_{L^2(\partial D)} = 0 \big\}. \tag{5}$$

Thus, we may define the so-called *Dirichlet-to-Neumann map*,

$$\Lambda_\gamma \colon H^{s+1}(\partial D)/\mathbb{C} \to H_\diamond^s(\partial D),$$
$$f \mapsto \gamma \frac{\partial u_{\gamma,f}}{\partial \nu}\Big|_{\partial D}, \tag{6}$$

which associates to each prescribed boundary value f the Neumann data of the solution of the PDE (1). The choice to quotient the domain of Λ_γ by $\mathbb{C}$ is natural as the Neumann data is invariant with respect to addition of scalars.

The *Calderón problem* [5] is a well studied inverse problem that addresses the task of recovering interior conductivities γ from knowledge of the boundary data Λ_γ. Note that while Λ_γ itself is a linear operator between Hilbert spaces, the 'forward map' $\gamma \to \Lambda_\gamma$ is non-linear. Landmark injectivity results by Sylvester and Uhlmann ($d \geq 3$) and by Nachman ($d = 2$) show, however, that recovery is in principle possible.

Theorem (Sylvester & Uhlmann [45]; Nachman [32]). *If $\Lambda_{\gamma_1} = \Lambda_{\gamma_2}$, then $\gamma_1 = \gamma_2$.*

Nachman [31] and Novikov [38] studied elaborate inversion algorithms that allow recovery of γ if exact knowledge of the entire operator Λ_γ is available. Moreover Alessandrini [2] (and later Novikov and Santacesaria [37] for $d = 2$) gave 'stability estimates' providing quantitative continuity bounds for the inverse map, and Mandache [28] gave an 'instability estimate' showing that these bounds are nearly sharp.

The Calderón problem has since been vigorously studied and an excellent survey can be found in Uhlmann [47] and also in the lecture notes by Salo [43]. Its importance partly stems from its applications to electrical impedance tomography (EIT) — described in more detail in the next section — where *discrete* boundary measurements of the operator Λ_γ are performed to infer the interior conductivity γ. Any such data comes with error, and arguably the most natural mathematical description of such approximate measurements is by a *statistical* noise model. As the superposition of many independent errors is well described by a normal distribution (via the central limit theorem), it is further natural to postulate that this noise follows a Gaussian law. In algorithmic practice this has already been widely acknowledged in the general setting of inverse problems, where statistical, and in particular Bayesian, inversion approaches have flourished in the last decade since the influential work of Stuart [44]. In the context of EIT we refer to the articles [10, 12, 21, 22, 24, 42] and the many references therein. Currently, little theory giving statistical guarantees for the performance of such Bayesian de-noising methodology is available, particularly for non-linear problems. Some recent progress has been made in non-linear settings (see [17,30,33–36,49]) but no results are available at present for the Calderón problem described above, and the purpose of the present paper is to at least partially fill this gap.

We will introduce three natural noise models for such statistical Caldéron problems, all asymptotically closely related, in the next section. We prove our main theorems initially in one of these models, and show in Appendix D.2 that the results are in fact valid in the other models too. The preferred model for the theoretical development is (12), wherein one observes Λ_γ corrupted by a Gaussian white noise in an appropriate space of Hilbert–Schmidt operators. The noise is described by the scalar quantity $\varepsilon > 0$ governing its magnitude and a parameter $r \in \mathbb{R}$ determining its 'spectral heteroscedasticity'. If we denote by $P_\varepsilon^\gamma = P_{\varepsilon,r}^\gamma$ the resulting probability law of the noisy observations Y of Λ_γ, then our main results can be summarised in the following two theorems.

Theorem 1. *Let $\alpha > 3 + d$ be an integer, let $m_0 \in (0, 1)$, $M > 1$ be given, and let D_0 be a domain in $\mathbb{R}^d$ such that $\bar{D}_0$ is contained in D.*

There exists a measurable function $\widehat{\gamma} = \widehat{\gamma}_\varepsilon(Y)$ of the observations $Y \sim P_\varepsilon^\gamma$ such that

$$\sup_{\gamma \in \Gamma_{m_0, D_0}^\alpha(M)} P_\varepsilon^\gamma \left(\|\widehat{\gamma} - \gamma\|_\infty > C \log(1/\varepsilon)^{-\delta} \right) \to 0 \quad as \ \varepsilon \to 0,$$

where $\delta > 0$ depends only on d and α, and C depends only on α, M, m_0, D, D_0, and r.

The estimator $\widehat{\gamma}$ in the previous theorem has a natural Bayesian interpretation in terms of the posterior mean of a suitable Gaussian process based prior for γ. Such priors are most effective for recovery of sufficiently smooth γ, and the precise bound on α is chosen here for convenience (see Remark 3 for further discussion).

The derivation and implementation of $\widehat{\gamma}$ are described in Section 3, where we give the more concrete Theorem 3, which implies Theorem 1. We note that $\widehat{\gamma}$ can be calculated without knowledge of the bound M for $\|\gamma\|_{H^\alpha(D)}$.

The slow (logarithmic) convergence rate is not surprising in view of the folklore that the Calderón problem is a severely ill-posed inverse problem (cf. also [28]). The following result makes this folklore information-theoretically precise – it shows that the convergence rate obtained by the estimator $\widehat{\gamma}$ is optimal in the statistical minimax sense, at least up to the precise value of the exponent δ, for the prototypical case where D_0, D are nested balls in $\mathbb{R}^d$. We denote by $\|\cdot\|$ the standard Euclidean norm on $\mathbb{R}^d$.

Theorem 2. *Let $D_0 = \{x \in \mathbb{R}^d : \|x\| < 1/2\} \subset D = \{x \in \mathbb{R}^d : \|x\| < 1\}$, let α be an integer greater than or equal to 2, and let $m_0 \in (0, 1)$ be arbitrary. For any $\delta' > \alpha(2d - 1)/d$ and all M large enough there exists $c = c(\delta', \alpha, d, m_0, r, M)$ such that*

$$\inf_{\widetilde{\gamma}} \sup_{\gamma \in \Gamma^\alpha_{m_0, D_0}(M)} P^\gamma_\varepsilon \left(\|\widetilde{\gamma} - \gamma\|_\infty > c \log(1/\varepsilon)^{-\delta'} \right) > 1/4$$

for all ε small enough, where the infimum extends over all measurable functions $\widetilde{\gamma} = \widetilde{\gamma}(Y)$ of the data $Y \sim P^\gamma_\varepsilon$.

The particular value $1/4$ in the lower bound is chosen for convenience. Determining the exact exponent δ in the minimax convergence rate is a delicate PDE question related directly to the stability estimates of [2, 37], and beyond the scope of the present paper. For $d \geq 3$, one could use an explicit bound of Novikov [39] to show that a choice of δ satisfying $\delta/\alpha \to 1$ as $\alpha \to \infty$ is permitted in Theorem 1. This scales proportionally (in the regularity index α) to the exponent $\delta' = \alpha(2d - 1)/d$ from the lower bound Theorem 2.

This paper is structured as follows. In Section 2 we introduce the measurement model we consider in our theorems, and discuss its relationship to physical measurement models arising in medical imaging practice. In Section 3 we give the construction of the Bayesian algorithm $\widehat{\gamma}$ that solves our noisy version of the Calderón problem. Section 4 contains some concluding remarks concerning the main theorems. All proofs and related background material are relegated to later sections. For convenience, the notation used is informally gathered in Appendix E.

2. Noise model and electrical impedance tomography

We now introduce three scenarios for noisy observations of the operator Λ_γ from (6), and discuss their relationship at the end of this section. Define

$$\widetilde{\Lambda}_\gamma = \Lambda_\gamma - \Lambda_1,$$

where the fixed (deterministic and known) operator Λ_1 is the Dirichlet-to-Neumann map for the standard Laplace equation, that is, eq. (1) with $\gamma = 1$ identically on D. We then equivalently consider measuring a noisy version of $\widetilde{\Lambda}_\gamma$.

Real-world data involving the Calderón problem arises for example in medical imaging, namely in *electrical impedance tomography*, see [10, 12, 21, 22, 24, 42] and references therein. Electrodes are attached to a patient (or some other physical medium), and are used both to apply voltages and to record the resulting currents. If we assume the applied voltages are uniform across the surface of any given electrode, and the electrodes measure the average current across their surface, we are led to the observation model

$$Y_{p,q} = \langle \widetilde{\Lambda}_\gamma[\psi_p], \psi_q \rangle_{L^2(\partial D)} + \varepsilon g_{p,q}, \quad p,q \leq P, \ g_{p,q} \overset{iid}{\sim} N(0,1), \ \varepsilon > 0, \quad (7)$$

where the ψ_p are, up to scaling factors, indicator functions $\mathbb{1}_{I_p}$ of some disjoint measurable subsets $(I_p)_{p \leq P}$ of ∂D representing the locations of the electrodes. Throughout $N(0,1)$ denotes the standard normal distribution. In principle the noise level $\varepsilon > 0$ could vary with p and q, but choosing scaling factors c_p so that the $\psi_p = c_p \mathbb{1}_{I_p}$ are $L^2(\partial D)$-orthonormal we expect to be able to realise the above homoscedastic noise model. Also note that while the inner product $\langle \cdot, \cdot \rangle_{L^2(\partial D)}$ is defined with respect to complex scalars, $\widetilde{\Lambda}_\gamma[\psi_p]$ takes real values since ψ_p does (see before Lemma 21), and hence it is natural to consider real-valued noise $g_{p,q}$ and data $Y_{p,q}$.

An alternative noise model considers spectral measurements. Denote by $(\phi_k = \phi_k^{(0)} : k \in \mathbb{N} \cup \{0\})$ an orthonormal basis of $L^2(\partial D)$ consisting of real-valued eigenfunctions of the Laplace–Beltrami operator on the compact manifold ∂D, described in more detail in Appendix A. (If D is a disc in $\mathbb{R}^2$ these comprise the usual trigonometric basis, while for D a ball in $\mathbb{R}^3$, they are the spherical harmonics.) By discarding the constant function ϕ_0 we obtain a basis of the spaces $L^2(\partial D)/\mathbb{C}$ and $L^2_\diamond(\partial D) = H^0_\diamond(\partial D)$. Moreover, appropriate rescaling of these basis functions also provides orthonormal bases $(\phi_k^{(r)} : k \in \mathbb{N})$ of all $H^r(\partial D)/\mathbb{C}$ and $H^r_\diamond(\partial D)$ spaces, $r \in \mathbb{R}$. For some $r \in \mathbb{R}$, we then consider the noisy matrix measurement model

$$Y_{j,k} = \langle \widetilde{\Lambda}_\gamma[\phi_j^{(r)}], \phi_k^{(0)} \rangle_{L^2(\partial D)} + \varepsilon g_{j,k}, \quad j \leq J, k \leq K, \ g_{j,k} \overset{iid}{\sim} N(0,1), \ \varepsilon > 0, \tag{8}$$

where again it is natural to consider real-valued noise $g_{j,k}$ only. The parameter r can in principle be chosen by the experimenter and reflects how the signal-to-noise ratio varies with frequency: as r increases, the signal at high frequencies (i.e. at larger values of j) decreases compared to the signal at low frequencies. Likely the most realistic choices are $r = 0$ (which will allow for comparison of the models (7) and (8)), and $r = 1$, in which case the signal-to-noise ratio is the same across all frequencies: since Λ_γ maps $H^1(\partial D)/\mathbb{C}$ to $L^2(\partial D)$ isomorphically (Lemma 20),

the signal magnitude $\|\Lambda_\gamma[\phi_j^{(1)}]\|_{L^2(\partial D)}$ is of order 1 for all j. A similar reasoning $(\|\phi_k^{(0)}\|_{L^2(\partial D)} = 1$ for all k) underpins the choice of the $L^2(\partial D)$-inner product in (8).

If we formally take the limit $J, K \to \infty$ in (8), we obtain a model of Gaussian white noise on a space of Hilbert–Schmidt operators as follows. For $j, k \in \mathbb{N}$, let $b_{jk}^{(r)} \colon H^r(\partial D) \to L^2(\partial D)$ denote the tensor product operator

$$b_{jk}^{(r)}(f) = \phi_j^{(r)} \otimes \phi_k^{(0)}(f) := \langle f, \phi_j^{(r)} \rangle_{H^r(\partial D)} \phi_k^{(0)}, \quad f \in H^r(\partial D), \qquad (9)$$

and define the space of linear operators

$$\mathbb{H}_r := \left\{ T \colon H^r(\partial D) \to L^2(\partial D), \ T = \sum_{j,k=1}^\infty t_{jk} b_{jk}^{(r)} : t_{jk} \in \mathbb{R}, \ \sum_{j,k=1}^\infty t_{jk}^2 < \infty \right\}. \tag{10}$$

The elements of $\mathbb{H}_r$ are the 'Hilbert–Schmidt' operators between (the real-valued subsets of) the Hilbert spaces $H^r(\partial D)$ and $L^2(\partial D)$, see [4, Chapter 12]. Moreover, $\mathbb{H}_r$ is itself a (real) Hilbert space for the inner product

$$\langle S, T \rangle_{\mathbb{H}_r} = \sum_{j,k=1}^\infty s_{jk} t_{jk} \equiv \sum_{j,k=1}^\infty \langle S\phi_j^{(r)}, \phi_k^{(0)} \rangle_{L^2(\partial D)} \langle T\phi_j^{(r)}, \phi_k^{(0)} \rangle_{L^2(\partial D)}.$$

We then consider observing a realisation of the Gaussian process

$$\big(Y(T) = \langle \widetilde{\Lambda}_\gamma, T \rangle_{\mathbb{H}_r} + \varepsilon \mathbb{W}(T) : T \in \mathbb{H}_r \big), \quad \varepsilon > 0, \tag{11}$$

where

$$\mathbb{W}(T) \equiv \langle \mathbb{W}, T \rangle_{\mathbb{H}_r} := \sum_{j,k=1}^\infty g_{jk} \langle T\phi_j^{(r)}, \phi_k^{(0)} \rangle_{L^2(\partial D)}, \quad \text{for } g_{jk} \overset{iid}{\sim} N(0, 1).$$

This makes sense rigorously only if $\widetilde{\Lambda}_\gamma \in \mathbb{H}_r$, and it is proved in Appendix C, Lemma 21, that this is indeed the case for any $\gamma \in \Gamma_{m,D'}$ and any $r \in \mathbb{R}$.

The process $\mathbb{W}$ so defined is a Gaussian white noise (*isonormal process*; see e.g. [16, p. 19]) indexed by the (real) Hilbert space $\mathbb{H}_r$. A closed form description of the data in (11) is therefore

$$Y = \widetilde{\Lambda}_\gamma + \varepsilon \mathbb{W}, \quad \varepsilon > 0. \tag{12}$$

We write $P_{\varepsilon,r}^\gamma$ for the law of Y in this last model. We often suppress the parameter r, and write P_ε^γ for the probability law and E_ε^γ for the corresponding expectation operator.

The continuous model (12) is more convenient for the application of PDE techniques and facilitates a clearer exposition in the proofs to follow. We prove our main

results Theorems 1 and 2 in that model initially. We will show that the models (12) and (7) are asymptotically closely related to each other in a rigorous 'Le Cam' sense, and that as a consequence, Theorem 1 and 2 also hold in the 'electrode model' (7): see Theorem 27 for a precise statement. The intuitive idea can be summarised as follows: Model (12) with $r = 0$ contains all the information available in model (7) simply by evaluating $Y(T)$ with $T = \psi_p \otimes \psi_q$ for each $p, q \leq P$ in (11); conversely, as $P \to \infty$, one can approximate Laplace–Beltrami eigenfunctions via linear combinations of indicator functions (under appropriate conditions on the sets I_p, $p \leq P$), and in doing so, given data from model (7) we approximately recover data from model (8) and ultimately then also from (12). We refer to Appendix D for rigorous details, where also the (simpler) equivalence of the models (8) and (12) is established.

3. The Bayesian approach to the noisy Calderón problem

We now construct the estimator $\widehat{\gamma}$ featuring in Theorem 1. Following the Bayesian approach to inverse problems advocated by A. Stuart [44], we will construct $\widehat{\gamma}$ in terms of the posterior mean arising from a certain Gaussian process prior. In the context of the EIT inverse problem a Bayesian approach was proposed already in [22], and conceptually related work appears in fact much earlier in Diaconis [9] who further traces some of the key ideas back to H. Poincaré; see [40, Chapter XV, §216] for what is possibly the first proposal of an infinite-dimensional Gaussian series prior in a numerical analysis context.

To this end we need to first establish the existence of a posterior distribution in our measurement setting. In the Gaussian white noise model (12), the log-likelihood function can be derived from the Cameron–Martin theorem in a suitable Hilbert space: precisely, the law P_ε^γ of Y is dominated by the law P_ε^1 of $\varepsilon \mathbb{W}$, with log-likelihood function

$$\ell(\gamma) \equiv \log p_\varepsilon^\gamma(Y) := \log \frac{dP_\varepsilon^\gamma}{dP_\varepsilon^1}(Y) = \frac{1}{\varepsilon^2}\langle Y, \widetilde{\Lambda}_\gamma \rangle_{\mathbb{H}_r} - \frac{1}{2\varepsilon^2}\|\widetilde{\Lambda}_\gamma\|_{\mathbb{H}_r}^2, \quad \gamma \in \Gamma_{m,D'}.$$
(13)

See [33, Section 7.4] for a detailed derivation, which requires Borel-measurability (ensured by Lemma 6 below) of the map $\gamma \mapsto \widetilde{\Lambda}_\gamma$ from the (Polish) space $\Gamma_{m,D'}$ equipped with the $\|\cdot\|_\infty$-topology into the Hilbert space $\mathbb{H}_r$.

Then for any prior (Borel) probability measure Π on $\Gamma_{m,D'}$, the posterior distribution given observations Y is given by

$$\Pi(B \mid Y) = \frac{\int_B p_\varepsilon^\gamma(Y)\, d\Pi(\gamma)}{\int_{\Gamma_{m,D'}} p_\varepsilon^\gamma(Y)\, d\Pi(\gamma)}, \quad B \subset \Gamma_{m,D'} \text{ Borel-measurable,} \qquad (14)$$

see again [33, Section 7.4] (and also [14, eq. (1.1)]). We denote by $E^\Pi[\cdot]$ the expectation operator according to the prior, and by $E^\Pi[\cdot \mid Y]$ the expectation according to the posterior.

What precedes is stated for a prior defined on the conductivities γ. The priors introduced below will be of the form $\gamma = \Phi \circ \theta$ for a suitable link function Φ, where θ takes values in a *linear* space, so that a Gaussian process prior can be assigned to θ. Composition with the map Φ described in Section 3.1 gives rise to a measurable bijection between the parameter spaces for θ and for γ, and the comments above therefore equivalently yield the existence of the posterior distribution for θ.

3.1. Prior construction. We define the prior on θ in terms of a base prior Π'. For the base prior we assume the following — we refer, e.g., to [16, Sections 2.1 and 2.6] for the basic definitions of Gaussian measures and processes and their reproducing kernel Hilbert spaces (RKHS).

Assumption 1. Let Π' be a centred Gaussian Borel probability measure on the Banach space $C_u(D)$ of uniformly continuous real-valued functions on D, and let α, β be integers satisfying $\alpha > \beta > 2 + d/2$. Assume $\Pi'(H^\beta(D)) = 1$ and that the RKHS $(\mathcal{H}, \|\cdot\|_\mathcal{H})$ of Π' is continuously embedded into the Sobolev space $H^\alpha(D)$.

Natural candidates for such priors are restrictions to D of Gaussian processes whose covariances are given by Whittle–Matérn kernels, see [14, pp. 313, 575] — in these cases one can satisfy the assumption for any $2 + d/2 < \beta < \alpha - d/2$ by taking $\mathcal{H}$ to *coincide* with the Sobolev space $H^\alpha(D)$. The restriction to integer-valued α, β is convenient to simplify some proofs.

In the proofs that follow we will require that the true γ_0 is in the 'interior' of the support of the induced prior on γ, so recalling that Theorem 1 is stated uniformly over $\Gamma^\alpha_{m_0, D_0}(M)$, we choose $0 < m_1 < m_0 < 1$ and a domain D_1 such that $\bar{D}_0 \subset D_1, \bar{D}_1 \subset D$. Then let $\zeta \colon D \to [0, 1]$ be a smooth cutoff function, identically one on D_0 and compactly supported in D_1. For a link function $\Phi \colon \mathbb{R} \to (m_1, \infty)$ to be specified, we define the prior $\Pi = \Pi_\varepsilon$ as the (Borel) law in $C_u(D)$ of the random function

$$\gamma = \Phi \circ \theta, \quad \theta(x) = \theta_\varepsilon(x) = \varepsilon^{d/(\alpha+d)} \zeta(x) \theta'(x), \quad x \in D, \ \theta' \sim \Pi', \quad (15)$$

where, in a slight abuse of notation, we use the notation Π for the prior laws both of θ and of the induced conductivity $\gamma = \Phi \circ \theta$. The link function Φ will be required to be *regular* in the sense of [36], that is to say, Φ is a smooth bijective function satisfying $\Phi(0) = 1$, $\Phi' > 0$ on $\mathbb{R}$, and $\|\Phi^{(j)}\|_\infty < \infty$ for all integers $j \geq 1$, with inverse function denoted by $\Phi^{-1} \colon (m_1, \infty) \to \mathbb{R}$. We refer to [36, Example 8], where a regular link function is exhibited, and to [36, Lemma 29] for basic properties of such functions. In particular we note that there are constants $C = C(\Phi), c = c(\Phi, m_0, m)$,

$C' = C'(\Phi, \alpha)$, and $c' = c'(\Phi, \alpha, m_0, m)$ such that for any bounded functions θ, θ_0, any integer $\alpha \geq d/2$ and any $\gamma_0, \gamma \in \Gamma_{m, D_1}$, $m \in (m_1, m_0)$,

$$\|\Phi \circ \theta - \Phi \circ \theta_0\|_\infty \leq C \|\theta - \theta_0\|_\infty, \tag{16}$$

$$\|\Phi^{-1} \circ \gamma - \Phi^{-1} \circ \gamma_0\|_\infty \leq c \|\gamma - \gamma_0\|_\infty, \tag{17}$$

$$\|\Phi \circ \theta\|_{H^\alpha(D)} \leq C'\big(1 + \|\theta\|_{H^\alpha(D)}^\alpha\big), \tag{18}$$

$$\|\Phi^{-1} \circ \gamma_0\|_{H^\alpha(D)} \leq c'\big(1 + \|\gamma_0\|_{H^\alpha(D)}^\alpha\big). \tag{19}$$

The first inequality is an immediate consequence of the mean value theorem and the third is given in [36, Lemma 29]. The second and fourth inequalites follow from the same arguments, applied to the function Φ^{-1} (this can be seen to be regular on the domain $[m, \infty)$ for $m > m_1$ by considering explicit formulas for its derivatives).

3.2. Posterior contraction result.
For the following result we define

$$\xi_{\varepsilon, \delta} = \log(1/\varepsilon)^{-\delta}, \quad \varepsilon, \delta > 0. \tag{20}$$

Theorem 3. *For some $m_0 \in (0, 1)$, D_0 compactly contained in D, and $M > 0$, suppose that the true conductivity γ_0 belongs to the set*

$$\Gamma_{m_0, D_0} \cap \big\{\Phi \circ \theta : \theta \in \mathcal{H}, \ \|\theta\|_{\mathcal{H}} \leq M\big\}, \tag{21}$$

and define $\theta_0 = \Phi^{-1} \circ \gamma_0$. For Π' satisfying Assumption 1 let Π be the prior arising from (15), and denote by $\Pi(\cdot \mid Y)$ the posterior distribution for θ arising from observations Y in the model (12). Then there exist constants $C = C(M, m_0, m_1, D, D_0, D_1, \Phi, \zeta, r, \alpha, \beta) > 0$ and $\delta = \delta(d, \beta) > 0$ such that

$$\Pi\big(\|\theta - \theta_0\|_\infty > C \xi_{\varepsilon, \delta} \mid Y\big) \to^{P_\varepsilon^{\gamma_0}} 0 \quad \text{as } \varepsilon \to 0. \tag{22}$$

Moreover, if $E^\Pi[\theta \mid Y]$ denotes the (Bochner) mean of $\Pi(\cdot \mid Y)$, then for any $K > C$,

$$\sup_{\gamma_0} P_\varepsilon^{\gamma_0}\big(\|E^\Pi[\theta \mid Y] - \theta_0\|_\infty > K \xi_{\varepsilon, \delta}\big) \to 0 \quad \text{as } \varepsilon \to 0, \tag{23}$$

where the supremum extends over all γ_0 in the set (21).

Theorem 3, whose proof is given in Section 5.4, immediately implies Theorem 1: Indeed, given an integer $\alpha > 3 + d$, let Π be a prior from (15) whose base prior Π' satisfies Assumption 1 with RKHS $\mathcal{H} = H^\alpha(D)$. (For instance, take a Whittle–Matérn prior, noting that a choice of integer $\beta > 2 + d/2$ is then admissible.) Let $\widehat{\theta} = E^\Pi[\theta \mid Y]$ be the associated posterior mean and define $\widehat{\gamma} = \Phi \circ \widehat{\theta}$. Then (16) and (23) will imply Theorem 1, so it suffices to show that the conditions of Theorem 1 imply those of Theorem 3, in particular that for any $\gamma_0 \in \Gamma_{m_0, D_0}^\alpha(M)$, there exists an $M' = M'(\alpha, M, m_0, D, D_0)$ such that $\|\theta_0\|_{H^\alpha(D)} \leq M'$. But this is immediate from (19).

4. Concluding remarks

Remark 1 (Computational aspects). The posterior mean $E^\Pi[\theta \mid Y]$ can be calculated via MCMC or expectation-propagation methods (naturally in the discretisations (7) or (8) of our continuous model (12)). This allows one to bypass potential difficulties encountered by optimisation based methods in non-linear settings such as the present one. For instance, the MAP estimates studied in [21, 36] may not provably recover global optima in the EIT setting, since the non-linearity of the map $\theta \mapsto \Lambda_{\Phi(\theta)} \equiv \Lambda_\gamma$ implies that the associated least squares criterion is not necessarily convex. Variational methods such as those proposed in [20] for EIT also typically require a convex relaxation to be efficiently computable, see [23].

The pCN algorithm [8] allows one to sample from posterior distributions in general inverse problems as long as the forward map $\theta \mapsto \Lambda_{\Phi(\theta)}$ can be evaluated, which in our setting has the basic cost of (numerically) solving the standard elliptic PDE (1). Even in the absence of log-concavity of the posterior measure one can give sampling guarantees for this algorithm, see [18], so that the approximate computation of $E^\Pi[\theta \mid Y]$ by the sample average $(1/M)\sum_m \theta_m$ of the pCN Markov chain is provably possible at any given noise level ε. Related work on MCMC-based approaches in the setting of electrical impedance tomography can be found in [10, 22, 42], wherein also many further references can be found. Instead of MCMC methods one can also resort to variational Bayes methods; see for example [12], where computation of the posterior mean is addressed specifically for the EIT problem relevant in the present paper.

Remark 2 (Estimation of γ at the boundary). In Theorem 1 we assume that the true conductivity γ_0 equals 1 on the complement of some known set $D_0 \subset D$. In the proofs we work mostly with the *differences* $\Lambda_\gamma - \Lambda_{\gamma_0}$ and the results can be expected to extend to γ_0 being known (and positive) on $D \setminus D_0$. Estimation of the value of γ_0 at ∂D is an 'easier' (less ill-posed) problem and is addressed, in a statistical setting, in [6].

Remark 3 (Smoothness of γ). In Theorem 1 we assume $\alpha > 3 + d$, and hence that the true conductivity is sufficiently smooth in a Sobolev sense. In terms of the proofs, this requirement primarily arises from the stability estimates employed ([2] requires a minimal Sobolev smoothness of γ of degree $\beta > 2 + d/2$), and further from analytical support properties of Gaussian process priors (for a Whittle–Matérn process with RKHS $H^\alpha(D)$ to be supported in $H^\beta(D)$, one needs $\alpha > \beta + d/2$), necessitating $\alpha > 2 + d$. To avoid technicalities with non-integer Sobolev spaces, we strengthen our hypothesis to $\alpha > 3 + d$.

An interesting direction for further research would seek to replace the smoothness assumption on γ_0 with different structural assumptions. For example, if we consider a class of functions that are piecewise constant on a known finite collection of subsets of D, better than logarithmic stability results are available in [3], and faster

convergence rates may then be obtained. Note that this would require different methods, e.g. prior measures that can model discontinuous functions, which is beyond the scope of the present paper.

5. Proofs

The proof of Theorem 3 follows the template devised in [30] for a very different inverse problem. We first show that the posterior distribution induced on the operators Λ_γ contracts around Λ_{γ_0} (see Theorem 13), by combining tools from Bayesian nonparametric statistics [14, 48] with analytical properties of the 'regression' operators Λ_γ (specifically of their low-rank approximations). Dealing with the unboundedness of Gaussian priors for θ and with the non-linearity of the composite forward map $\theta \mapsto \Lambda_{\Phi(\theta)}$ poses a main challenge in the proof. As in [30] this challenge is overcome using the rescaling in (15) of the base prior Π'. For such priors the posterior distribution concentrates on sufficiently regular conductivities γ that the stability estimates of [2,37] — which we show to hold also for the information-theoretically relevant norms here — can be applied, resulting in contraction of the posterior distributions of γ and θ about γ_0 and θ_0 respectively. Finally, to deduce consistency of the posterior mean, we adapt a quantitative uniform integrability argument from [30] to the present situation. The proof of Theorem 2 follows from the 'instability' estimate in [28] and common information-theoretic lower bound techniques for statistical estimators as in [16, 46].

Remark. The model (12) naturally considers Λ_γ as an element of the *real* Hilbert space $\mathbb{H}_r$ consisting of Hilbert–Schmidt operators between the sets of *real-valued* functions of $H^r(\partial D)/\mathbb{C}$ and $L^2_\diamond(\partial D)$, respectively. Various results in the literature — such as the stability and instability results of [2] and [28] to be used below — regard Λ_γ as an element of the space $\mathbb{H}_{r,\mathbb{C}}$ of Hilbert–Schmidt operators between *complex-valued* function spaces $H^r(\partial D)/\mathbb{C}, L^2_\diamond(\partial D)$. This difference is inconsequential since $\mathbb{H}_r$ embeds into $\mathbb{H}_{r,\mathbb{C}}$ via the unique extension

$$T_\mathbb{C}(f + ig) = T(f) + iT(g)$$

for $T \in \mathbb{H}_r$, and the Hilbert–Schmidt norm of this larger space, when restricted to $\mathbb{H}_r$, coincides with the $\mathbb{H}_r$ norm. In fact, since we regard $b_{jk}^{(r)} \in \mathbb{H}_r$ in (9) as a map from $H^r(\partial D)$ to $L^2_\diamond(\partial D)$, strictly speaking we have already embedded $\mathbb{H}_r$ into $\mathbb{H}_{r,\mathbb{C}}$ in this way.

5.1. Low rank approximation of $\widetilde{\Lambda}_\gamma$.
A key idea used in various proofs that follow is that we can project the operator $\widetilde{\Lambda}_\gamma$ onto a finite-dimensional subspace and incur only a small error. We recall from (9) the orthonormal basis $(b_{jk}^{(r)})_{j,k\in\mathbb{N}}$ of $\mathbb{H}_r$

consisting of tensor product operators

$$b_{jk}^{(r)}(f) = \phi_j^{(r)} \otimes \phi_k^{(0)}(f) := \langle f, \phi_j^{(r)} \rangle_{H^r(\partial D)} \phi_k^{(0)}, \quad f \in H^r(\partial D)/\mathbb{C}, \ j, k \in \mathbb{N},$$

where the Laplace–Beltrami eigenfunctions $(\phi_j^{(r)})_{j \in \mathbb{N}}$ were introduced before (8). An operator $U \in \mathbb{H}_r$ has coefficients

$$\langle U, b_{jk}^{(r)} \rangle_{\mathbb{H}_r} = \langle U\phi_j^{(r)}, \phi_k^{(0)} \rangle_{L^2(\partial D)}$$

with respect to this basis, and we define the projection map π_{JK} by

$$\pi_{JK} U = \sum_{j \leq J} \sum_{k \leq K} \langle U\phi_j^{(r)}, \phi_k^{(0)} \rangle_{L^2(\partial D)} b_{jk}^{(r)}. \tag{24}$$

Lemma 4. *For constants $m \in (0, 1)$, $M > 1$ and some domain D' compactly contained in D, let $\gamma \in \Gamma_{m,D'}$ be bounded by M on D. For any $\nu > 0$ there is a constant $C = C(\nu, D, D', r) > 0$ such that*

$$\|\widetilde{\Lambda}_\gamma - \pi_{JK}\widetilde{\Lambda}_\gamma\|_{\mathbb{H}_r} \leq C \frac{M}{m} \min(J, K)^{-\nu}.$$

Proof. Apply Lemma 18 from Appendix B with $s = 0$, $p = r - \nu(d - 1)$, and $q = \nu(d - 1)$, and note

$$\|\widetilde{\Lambda}_\gamma\|_{H^{p-(d-1)}(D) \to H^q(D)} \leq C \frac{M}{m}$$

for such a constant C by Lemma 20. $\qquad \square$

The proofs of the stability results for the Calderón problem in the next section involve the $H^{1/2}(\partial D)/\mathbb{C} \to H^{-1/2}(\partial D)$ operator norm, which we denote by $\|\cdot\|_*$. To connect this norm to the information-theoretically relevant $\mathbb{H}_r$-norm, the following consequence of Lemma 4 will be useful.

Lemma 5. *For $m \in (0, 1)$, $M_0, M_1 > 1$ and D' a domain compactly contained in D, let $\gamma_0, \gamma \in \Gamma_{m,D'}$ be bounded on D by M_0 and M_1 respectively. Then there are constants C_1 and C_2 depending only on r, D and D_0 such that if $\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_* \leq 1$ then*

$$\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r} \leq C_1 \left(\frac{M_1 + M_0}{m} \|\Lambda_\gamma - \Lambda_{\gamma_0}\|_* \right)^{1/2}, \tag{25}$$

and if $\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r} \leq 1$ then

$$\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_* \leq C_2 \left(\frac{M_1 + M_0}{m} \|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r} \right)^{1/2}. \tag{26}$$

Proof. For $J > 0$ and $\nu > 0$ to be chosen, by Lemma 4 we have

$$\|\widetilde{\Lambda}_\gamma - \pi_{JJ}\widetilde{\Lambda}_\gamma\|_{\mathbb{H}_r} \leq C \frac{M_1}{m} J^{-\nu},$$

for a constant $C = C(\nu, D, D', r)$, and a corresponding bound holds for

$$\|\widetilde{\Lambda}_{\gamma_0} - \pi_{JJ}\widetilde{\Lambda}_{\gamma_0}\|_{\mathbb{H}_r}.$$

An application of Lemma 17 with $s = 0$, $p = d - 1/2$, and $q = -1/2$, also yields (with $x_+ = \max(x, 0)$, and $\mathcal{L}_2(H^{d-1/2}, H^{-1/2})$ a space of Hilbert–Schmidt operators as defined in Appendix B)

$$\|\pi_{JJ}\widetilde{\Lambda}_\gamma - \pi_{JJ}\widetilde{\Lambda}_{\gamma_0}\|_{\mathbb{H}_r}$$
$$\leq C'\big(1 + J^{1/(d-1)}\big)^{1/2 + (d-1/2-r)_+}\|\pi_{JJ}(\widetilde{\Lambda}_\gamma - \widetilde{\Lambda}_{\gamma_0})\|_{\mathcal{L}_2(H^{d-1/2}, H^{-1/2})}$$
$$\leq C'\big(1 + J^{1/(d-1)}\big)^{(d+|r|)}\|\widetilde{\Lambda}_\gamma - \widetilde{\Lambda}_{\gamma_0}\|_{\mathcal{L}_2(H^{d-1/2}, H^{-1/2})}$$
$$\leq c'\big(1 + J^{1/(d-1)}\big)^{(d+|r|)}\|\widetilde{\Lambda}_\gamma - \widetilde{\Lambda}_{\gamma_0}\|_*,$$

for constants C', c' depending on D, r, where we use Lemma 18 to obtain the final inequality. Since $\Lambda_\gamma - \Lambda_{\gamma_0} = \widetilde{\Lambda}_\gamma - \widetilde{\Lambda}_{\gamma_0}$, we deduce, for a constant C'' that

$$\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r}$$
$$\leq \|\widetilde{\Lambda}_\gamma - \pi_{JJ}\widetilde{\Lambda}_\gamma\|_{\mathbb{H}_r} + \|\widetilde{\Lambda}_{\gamma_0} - \pi_{JJ}\widetilde{\Lambda}_{\gamma_0}\|_{\mathbb{H}_r} + \|\pi_{JJ}\widetilde{\Lambda}_\gamma - \pi_{JJ}\widetilde{\Lambda}_{\gamma_0}\|_{\mathbb{H}_r}$$
$$\leq C''\big(\big(\tfrac{M_1+M_0}{m}\big)J^{-\nu} + J^{(d+|r|)/(d-1)}\|\widetilde{\Lambda}_\gamma - \widetilde{\Lambda}_{\gamma_0}\|_*\big).$$

Since $\|\widetilde{\Lambda}_\gamma - \widetilde{\Lambda}_{\gamma_0}\|_* \leq 1$, we can choose an integer J to balance the two terms up to a constant (take $J = \lfloor\big(\tfrac{m}{M_0+M_1}\|\widetilde{\Lambda}_\gamma - \widetilde{\Lambda}_{\gamma_0}\|_*\big)^{-(d-1)/(\nu(d-1)+d+|r|)}\rfloor$). This yields, for a constant c'',

$$\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r} \leq c''\big(\tfrac{M_1+M_0}{m}\big)\Big(\big(\tfrac{M_1+M_0}{m}\big)^{-1}\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_*\Big)^{\nu(d-1)/(\nu(d-1)+d+|r|)}.$$

Choosing $\nu = (d + |r|)/(d - 1)$ yields (25).

For (26), given that

$$\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_* \leq \|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathcal{L}_2(H^{1/2}(\partial D)/\mathbb{C}, H^{-1/2}(\partial D))}$$

(using the general fact that the Hilbert–Schmidt norm upper bounds the operator norm of a linear operator between separable Hilbert spaces), and the observation that the proof of Lemma 4 can be adapted to apply with the $\mathcal{L}_2(H^{1/2}(\partial D)/\mathbb{C}, H^{-1/2}(\partial D))$ norm in place of the $\mathbb{H}_r$ norm, an almost identical argument to the above yields

$$\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_* \leq C\Big(\big(\tfrac{M_1+M_0}{m}\big)J^{-\nu} + J^{(r-1/2)_+/(d-1)}\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r}\Big).$$

Choosing J to balance the terms yields

$$\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_* \leq C'\big(\tfrac{M_1+M_0}{m}\big)\Big(\big(\tfrac{M_1+M_0}{m}\big)^{-1}\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r}\Big)^{\nu(d-1)/(\nu(d-1)+(r-1/2)_+)},$$

and the result follows from noting that the exponent is at least $1/2$ for

$$\nu > (r - 1/2)_+/(d - 1). \qquad \square$$

5.2. Forward and inverse continuity results. We now prove the following estimates for the maps $\gamma \mapsto \Lambda_\gamma, \Lambda_\gamma \rightarrow \gamma$.

Lemma 6. *For $m \in (0,1)$, $M_0, M_1 > 1$ and D' a domain compactly contained in D, let $\gamma, \gamma_0 \in \Gamma_{m,D'}$ be bounded on D by M_1 and M_0 respectively. Then there exist constants $C = C(r, D, D')$, $\tau = \tau(D)$ such that*

$$\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r} \leq C \, \tfrac{M_0 M_1}{m^2} \|\gamma - \gamma_0\|_\infty^{1/2},$$

whenever $\|\gamma - \gamma_0\|_\infty \leq \tau \tfrac{m^2}{M_0 M_1}$.

Lemma 7. *For some $\beta > 2 + d/2$, some $m \in (0,1)$, $M > 0$ and some domain D' compactly contained in D, suppose*

$$\gamma, \gamma_0 \in \Gamma^\beta_{m,D'}(M) = \{\gamma \in \Gamma_{m,D'} : \|\gamma\|_{H^\beta(D)} \leq M\}.$$

Then there exist constants C and τ depending only on M, D, D', m, β and r such that, for a constant $\delta = \delta(d, \beta) > 0$,

$$\|\gamma - \gamma_0\|_\infty \leq C \, |\log\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r}|^{-\delta},$$

whenever $\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r} \leq \tau$.

The explicit form of the dependence of the constant in Lemma 6 on M_1 and M_0 will be convenient in the proof of Lemma 11.

Proof of Lemma 6. We initially show, for some $C = C(D)$, that

$$\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_* \leq C M_0 \tfrac{m + M_1}{m^2} \|\gamma - \gamma_0\|_\infty. \tag{27}$$

The result then follows from Lemma 5, noting that

$$(M_0 + M_1) M_0 (m + M_1)/m^3 \leq 4 M_0^2 M_1^2 / m^4.$$

For γ, γ_0 as given, let $f \in H^{1/2}(\partial D)/\mathbb{C}$, and recall we write $u_{\gamma,f}$ for the unique solution in $\mathcal{H}_{-1/2} \equiv H^1(D)/\mathbb{C}$ to the Dirichlet problem on D with conductivity γ and boundary data f, whose existence is guaranteed by Lemma 19, and similarly for $u_{\gamma_0,f}$. The equivalence class of functions $u_{\gamma,f} - u_{\gamma_0,f}$ has a representative $w \in H_0^1(D)$, which is easily seen to (weakly) solve the PDE

$$\begin{aligned}
\nabla \cdot (\gamma \nabla w) &= \nabla \cdot \big((\gamma_0 - \gamma) \nabla u_{\gamma_0,f}\big) &&\text{in } D, \\
w &= 0 &&\text{on } \partial D.
\end{aligned} \tag{28}$$

We have the dual representation (see Remark ii. in Appendix A)

$$\left\|\frac{\partial w}{\partial \nu}\right\|_{H^{-1/2}(\partial D)} = \sup\left\{\left|\left\langle\frac{\partial w}{\partial \nu}, v\right\rangle_{L^2(\partial D)}\right| : v \in H^{1/2}(\partial D), \|v\|_{H^{1/2}(\partial D)} = 1\right\}.$$

For $v \in H^{1/2}(\partial D)$, by a standard trace theorem (e.g. [27, Chapter I, Theorem 9.4]) there exists $V \in H^1(D)$ such that

$$V|_{\partial D} = v \quad \text{and} \quad \|V\|_{H^1(D)} \leq C\|v\|_{H^{1/2}(\partial D)}$$

for a constant $C = C(D)$. Repeatedly applying the divergence theorem (recalling that $\gamma = \gamma_0 = 1$ on ∂D) and the Cauchy–Schwarz inequality, and using (28), we deduce

$$\left| \int_{\partial D} \bar{v} \frac{\partial w}{\partial \nu} \right| = \left| \int_D \bar{V} \nabla \cdot (\gamma \nabla w) + \int_D \gamma \nabla \bar{V} \cdot \nabla w \right|$$

$$\leq \left| \int_D \bar{V} \nabla \cdot \big((\gamma_0 - \gamma) \nabla u_{\gamma_0, f} \big) \right| + \|\gamma\|_\infty \|V\|_{H^1(D)} \|\nabla w\|_{L^2(D)}$$

$$\leq \left| \int_D (\gamma_0 - \gamma) \nabla \bar{V} \cdot \nabla u_{\gamma_0, f} \right| + C\|\gamma\|_\infty \|v\|_{H^{1/2}(\partial D)} \|\nabla w\|_{L^2(D)}$$

$$\leq C\|v\|_{H^{1/2}(\partial D)} \big(\|\gamma_0 - \gamma\|_\infty \|\nabla u_{\gamma_0, f}\|_{L^2(D)} + M_1 \|\nabla w\|_{L^2(D)} \big),$$

hence

$$\left\| \frac{\partial w}{\partial \nu} \right\|_{H^{-1/2}(\partial D)} \leq C \big(\|\gamma_0 - \gamma\|_\infty \|\nabla u_{\gamma_0, f}\|_{L^2(D)} + M_1 \|\nabla w\|_{L^2(D)} \big). \qquad (29)$$

A weak solution w to (28) by definition satisfies, for any $v \in H_0^1(D)$,

$$\int_D \gamma \nabla w \cdot \nabla \bar{v} = \int_D (\gamma_0 - \gamma) \nabla u_{\gamma_0, f} \cdot \nabla \bar{v}.$$

In particular this applies with $v = w$, hence, since $\gamma \geq m$ on D, we apply the Cauchy–Schwarz inequality to deduce

$$m\|\nabla w\|_{L^2(D)}^2 \leq \|\gamma_0 - \gamma\|_\infty \|\nabla u_{\gamma_0, f}\|_{L^2(D)} \|\nabla w\|_{L^2(D)},$$

which, returning to (29), shows that

$$\left\| \frac{\partial w}{\partial \nu} \right\|_{H^{-1/2}(\partial D)} \leq C \|\nabla u_{\gamma_0, f}\|_{L^2(D)} \|\gamma_0 - \gamma\|_\infty \big(1 + \tfrac{M_1}{m} \big).$$

Applying the trace result [27, Theorem 9.4, Chapter I] now to each representative of the equivalence class $f \in H^{1/2}(D)/\mathbb{C}$ and optimising, there exists $F \in H^1(D)/\mathbb{C}$ such that

$$F|_{\partial D} = f \quad \text{and} \quad \|F\|_{H^1(D)/\mathbb{C}} \leq C\|f\|_{H^{1/2}(\partial D)/\mathbb{C}}$$

for a constant $C = C(D)$. By definition of a weak solution to (1) we have

$$\int_D \gamma_0 \nabla u_{\gamma_0, f} \cdot \nabla \overline{(u_{\gamma_0, f} - F)} = 0,$$

and again applying the Cauchy–Schwarz inequality we deduce

$$\|\nabla u_{\gamma_0, f}\|_{L^2(D)} \le C \frac{M_0}{m} \|f\|_{H^{1/2}(\partial D)/\mathbb{C}}. \tag{30}$$

Overall we have shown

$$\begin{aligned}
\big\|(\Lambda_\gamma - \Lambda_{\gamma_0}) f\big\|_{H^{-1/2}(\partial D)} &\equiv \Big\| \frac{\partial w}{\partial \nu} \Big\|_{H^{-1/2}(\partial D)} \\
&\le C \big(1 + \tfrac{M_1}{m}\big) \tfrac{M_0}{m} \|\gamma - \gamma_0\|_\infty \|f\|_{H^{1/2}(\partial D)/\mathbb{C}}.
\end{aligned}$$

Taking the supremum over all f with $H^{1/2}(\partial D)/\mathbb{C}$ norm equal to 1, (27) follows.   □

Proof of Lemma 7. When $d \ge 3$, Theorem 1 in Alessandrini [2] states that there exist constants $\delta = \delta(d, \beta)$ and $C = C(M, m, D, D', \beta)$ such that there is a (monotone) function ω satisfying

$$\|\gamma - \gamma_0\|_\infty \le C \omega\big(\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_*\big), \quad \omega(t) \le \log(1/t)^{-\delta} \text{ for } t < e^{-1}. \tag{31}$$

Likewise, when $d = 2$, we can use Theorem 1.1 in Novikov & Santacesaria [37] in conjunction with the usual reduction of the Calderón problem to a suitable Schrödinger equation (e.g. see Lemma 4.8 and the proof of Theorem 4.1 in [43]) to show that (31) still holds in this case. (In [37], an a priori bound for $\|\gamma\|_{C^2}$ is assumed, which is derived here from the bound M on $\|\gamma\|_{H^\beta}$ using a Sobolev embedding.)

To proceed, we appeal to Lemma 5, noting that M upper bounds $\|\gamma\|_\infty$ and $\|\gamma_0\|_\infty$ (up to a multiplicative constant) by a Sobolev embedding. Thus, for a constant C' depending on M, m, D, D', β and r we have

$$\begin{aligned}
\omega\big(\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_*\big) &\le \omega\big(C' \|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r}^{1/2}\big) \\
&\le \big(\tfrac{1}{4} \log\big(\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r}^{-1}\big)\big)^{-\delta},
\end{aligned}$$

provided $\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r} < \min(e^{-2}(C')^{-2}, (C')^{-4}, 1)$. The result follows.   □

5.3. Tests and prior support properties. In this section we prove two main auxiliary results. First, we prove the existence of certain statistical test functions required in the contraction theorem given in Section 5.4. Instead of using robust 'Hellinger-distance'-based testing as in [30, 48], it is more convenient in the present setting (in part to deal with necessary boundedness restrictions on γ) to deduce the existence of tests from the existence of certain estimators with sufficiently good concentration properties (following ideas in [15]).

Recall Γ_{m_1, D_1} is a superset of Γ_{m_0, D_0} on which the prior on $\gamma = \Phi \circ \theta$ concentrates its mass.

Lemma 8. *Let $\gamma_0 \in \Gamma_{m_0,D_0}$ be bounded by M_0 on D. Let $\eta_\varepsilon > 0$ satisfy*

$$\eta_\varepsilon \varepsilon^{-(1-a)} \to \infty \quad as \ \varepsilon \to 0$$

for some $0 < a < 1$. For any $\kappa > 0$ and $M_1 > 1$, there exists a constant $C = C(\kappa, m_1, D, D_1, M_1, M_0)$ and tests (i.e. $\{0,1\}$-valued measurable functions of the data) $\psi = \psi_\varepsilon(Y)$ with $Y \sim P_\varepsilon^\gamma$ from (12) such that for all ε small enough

$$\max\left(E_\varepsilon^{\gamma_0}\psi, \sup\{E_\varepsilon^\gamma[1-\psi] : \gamma \in \Gamma_{m_1,D_1}, \ \|\gamma\|_\infty \leq M_1, \ \|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r} \geq C\eta_\varepsilon\}\right)$$
$$\leq e^{-\kappa(\eta_\varepsilon/\varepsilon)^2}.$$

Proof. We prove the existence of an estimator $\widehat{\Lambda}$ satisfying that for any $\kappa > 0$, there exists a constant $C' = C'(\kappa, m_1, D_1, M_0, M_1, D)$ such that for all ε small enough,

$$\sup\{P_\varepsilon^\gamma\left(\|\widehat{\Lambda} - \widetilde{\Lambda}_\gamma\|_{\mathbb{H}_r} > C'\eta_\varepsilon\right) : \gamma \in \Gamma_{m_1,D_1}, \ \|\gamma\|_\infty \leq \max(M_0,M_1)\}$$
$$\leq e^{-\kappa(\eta_\varepsilon/\varepsilon)^2}. \quad (32)$$

Then, setting $C = 2C'$ and $\psi_\varepsilon(Y) = \mathbb{1}\{\|\widehat{\Lambda} - \widetilde{\Lambda}_{\gamma_0}\|_{\mathbb{H}_r} > \frac{1}{2}C\eta_\varepsilon\}$, the result follows from an application of the triangle inequality (see, e.g., the proof of Proposition 6.2.2 in [16]).

Define an estimator $\widehat{\Lambda}$ by $\widehat{\Lambda} = \sum_{j,k \leq J} \widehat{\Lambda}_{jk} b_{jk}^{(r)}$, where $J = J_\varepsilon = \lfloor \eta_\varepsilon/\varepsilon \rfloor$ and

$$\widehat{\Lambda}_{jk} = \langle Y, b_{jk}^{(r)} \rangle_{\mathbb{H}_r} = \langle \widetilde{\Lambda}_\gamma \phi_j^{(r)}, \phi_k^{(0)} \rangle_{L^2(\partial D)} + \varepsilon g_{jk}, \ Y \sim P_\varepsilon^\gamma, \quad (33)$$

where we note $g_{jk} = \langle \mathbb{W}, b_{jk}^{(r)} \rangle_{\mathbb{H}_r} \overset{iid}{\sim} N(0,1)$. Then we have the bias-variance decomposition

$$P_\varepsilon^\gamma\left(\|\widehat{\Lambda} - \widetilde{\Lambda}_\gamma\|_{\mathbb{H}_r} > C'\eta_\varepsilon\right) \leq \mathbb{1}\{\|\widetilde{\Lambda}_\gamma - \pi_{JJ}\widetilde{\Lambda}_\gamma\|_{\mathbb{H}_r} > \frac{1}{2}C'\eta_\varepsilon\}$$
$$+ P_\varepsilon^\gamma\left(\|\widehat{\Lambda} - \pi_{JJ}\widetilde{\Lambda}_\gamma\|_{\mathbb{H}_r} > \frac{1}{2}C'\eta_\varepsilon\right). \quad (34)$$

Recall, by Lemma 4, for any $\nu > 0$ there is a constant $C_1 = C_1(\nu, r, M_0, M_1, m_1, D_1, D)$ such that

$$\|\widetilde{\Lambda}_\gamma - \pi_{JJ}\widetilde{\Lambda}_\gamma\|_{\mathbb{H}_r} \leq C_1 J^{-\nu}, \quad (35)$$

hence the indicator in (34) is bounded by $\mathbb{1}\{C_1 J^{-\nu} > \frac{1}{2}C'\eta_\varepsilon\}$. Choosing $\nu > (1-a)/a$, one finds that the assumption $\eta_\varepsilon \varepsilon^{-(1-a)} \to \infty$ ensures this term vanishes for ε small enough.

For the variance term in (34), observe that by Parseval's identity

$$\|\widehat{\Lambda} - \pi_{JJ}\widetilde{\Lambda}_\gamma\|_{\mathbb{H}_r}^2 = \sum_{j,k \leq J} \left(\widehat{\Lambda}_{jk} - \langle \widetilde{\Lambda}_\gamma \phi_j^{(r)}, \phi_k^{(0)} \rangle_{L^2(\partial D)}\right)^2 = \varepsilon^2 \sum_{j,k \leq J} g_{jk}^2.$$

One now applies a standard tail inequality (e.g. [16, Theorem 3.1.9]) to the effect that

$$\Pr\left(\sum_{j,k \leq J} g_{jk}^2 \geq J^2 + 2J\sqrt{x} + 2x \right) \leq e^{-x}. \tag{36}$$

For a constant $\kappa > 0$, taking $x = \kappa(\eta_\varepsilon/\varepsilon)^2$, and for our choice $J = \lfloor \eta_\varepsilon/\varepsilon \rfloor$, we see that for C' large enough depending only on κ we have

$$P_\varepsilon^\gamma\left(\|\widehat{\Lambda} - \pi_{JJ}\widetilde{\Lambda}_\gamma\|_{\mathbb{H}_r}^2 > \frac{1}{2}C'\eta_\varepsilon^2 \right) \leq e^{-\kappa(\eta_\varepsilon/\varepsilon)^2},$$

hence the result. $\qquad\square$

To proceed define $K(p,q) = E_{X \sim p} \log \frac{p}{q}(X)$ to be the Kullback–Leibler divergence between distributions with densities p and q, and recall the definition of the probability densities p_ε^γ from (13). Also denote by Var_γ the variance operator associated to the probability measure P_ε^γ. The following is then a standard result for a white noise model on a Hilbert space.

Lemma 9. *Let $\gamma_0, \gamma_1 \in \Gamma_{m_1, D_1}$. Then*

$$K(p_\varepsilon^{\gamma_0}, p_\varepsilon^{\gamma_1}) = \frac{1}{2}\varepsilon^{-2}\|\Lambda_{\gamma_0} - \Lambda_{\gamma_1}\|_{\mathbb{H}_r}^2,$$

and

$$\mathrm{Var}_{\gamma_0}\left(\log \frac{p_\varepsilon^{\gamma_0}}{p_\varepsilon^{\gamma_1}} \right) = \varepsilon^{-2}\|\Lambda_{\gamma_0} - \Lambda_{\gamma_1}\|_{\mathbb{H}_r}^2.$$

Proof. Using the explicit formula (13) for the log-likelihoods, we see that under γ_0,

$$\begin{aligned}
\ell(\gamma_0) - \ell(\gamma_1) &= \varepsilon^{-2}\langle Y, \widetilde{\Lambda}_{\gamma_0} - \widetilde{\Lambda}_{\gamma_1}\rangle_{\mathbb{H}_r} - \tfrac{1}{2}\varepsilon^{-2}\|\widetilde{\Lambda}_{\gamma_0}\|_{\mathbb{H}_r}^2 + \tfrac{1}{2}\varepsilon^{-2}\|\widetilde{\Lambda}_{\gamma_1}\|_{\mathbb{H}_r}^2 \\
&= \tfrac{1}{2}\varepsilon^{-2}\|\widetilde{\Lambda}_{\gamma_0} - \widetilde{\Lambda}_{\gamma_1}\|_{\mathbb{H}_r}^2 + \varepsilon^{-1}\langle \mathbb{W}, \widetilde{\Lambda}_{\gamma_0} - \widetilde{\Lambda}_{\gamma_1}\rangle_{\mathbb{H}_r},
\end{aligned}$$

which is normally distributed with mean $\tfrac{1}{2}\varepsilon^{-2}\|\widetilde{\Lambda}_{\gamma_0} - \widetilde{\Lambda}_{\gamma_1}\|_{\mathbb{H}_r}^2$ and variance $\varepsilon^{-2}\|\widetilde{\Lambda}_{\gamma_0} - \widetilde{\Lambda}_{\gamma_1}\|_{\mathbb{H}_r}^2$. Noting that $\widetilde{\Lambda}_{\gamma_0} - \widetilde{\Lambda}_{\gamma_1} = \Lambda_{\gamma_0} - \Lambda_{\gamma_1}$, we deduce the result. $\quad\square$

We define 'balls' $B_{KL}^\varepsilon(\eta)$ around the true parameter $\theta_0 = \Phi^{-1} \circ \gamma_0$ by

$$\begin{aligned}
B_{KL}^\varepsilon(\eta) = \big\{\theta \in C_u(D) : K(p_\varepsilon^{\gamma_0}, p_\varepsilon^{\Phi \circ \theta}) \leq (\eta/\varepsilon)^2, \\
\mathrm{Var}_{\gamma_0}\big(\log(p_\varepsilon^{\gamma_0}/p_\varepsilon^{\Phi \circ \theta}) \big) \leq (\eta/\varepsilon)^2 \big\}. \tag{37}
\end{aligned}$$

Then the following is an immediate consequence of Lemma 9.

Corollary 10. *For any $\eta > 0$,*

$$\{\theta \in C_u(D) : \|\Lambda_{\Phi \circ \theta} - \Lambda_{\gamma_0}\|_{\mathbb{H}_r} \leq \eta\} \subseteq B_{KL}^\varepsilon(\eta).$$

With the preceding preparations, we can now prove the following support result for the prior Π from Theorem 3, using a result of Li & Linde [26].

Lemma 11. *Let $\eta_\varepsilon = \varepsilon^{\alpha/(\alpha+d)}$. Under the conditions of Theorem 3, there exists a constant $\omega = \omega(\alpha, m_1, M, D, D_1, \Phi, r) > 0$ such that $\Pi(B^\varepsilon_{KL}(\eta_\varepsilon)) \geq e^{-\omega(\eta_\varepsilon/\varepsilon)^2}$ for all ε small enough, uniformly in γ_0 in the set (21).*

Proof. By (16) and a Sobolev embedding, there is a constant M_0 depending only on Φ, α and M such that $\|\gamma_0\|_\infty \leq M_0$. By Lemma 6, we deduce, for a constant $C = C(r, D, D_1)$,

$$\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r} \leq C \frac{M_0\|\gamma\|_\infty}{m_1^2}\|\gamma - \gamma_0\|_\infty^{1/2}$$

provided $\|\gamma - \gamma_0\|_\infty$ is small enough. It follows from this calculation and Corollary 10 that for η_ε small enough and some constant $C' > 0$ we have

$$\{\theta : \|\Phi \circ \theta - \gamma_0\|_\infty \leq C'\eta_\varepsilon^2\} \subseteq \{\|\Lambda_{\Phi \circ \theta} - \Lambda_{\gamma_0}\|_{\mathbb{H}_r} \leq \eta_\varepsilon\} \subseteq B^\varepsilon_{KL}(\eta_\varepsilon).$$

Appealing again to (16), we further deduce that

$$\{\theta \in C_u(D) : \|\theta - \theta_0\|_\infty \leq A\eta_\varepsilon^2\} \subseteq B^\varepsilon_{KL}(\eta_\varepsilon),$$

for a constant $A = A(\alpha, M, m_1, D, D_1, r, \Phi)$, so that it remains to lower bound $\Pi(\|\theta - \theta_0\|_\infty \leq A\eta_\varepsilon^2)$. Note (recalling Assumption 1 and the definition of the prior (15)) that Π has RKHS $\mathcal{H}_\varepsilon = \{\zeta\theta' : \theta' \in \mathcal{H}\}$, with norm $\|\cdot\|_{\mathcal{H}_\varepsilon}$ satisfying the bound

$$\|\theta\|_{\mathcal{H}_\varepsilon} \leq \varepsilon^{-d/(\alpha+d)}\|\theta'\|_{\mathcal{H}} = (\eta_\varepsilon/\varepsilon)\|\theta'\|_{\mathcal{H}},$$

for any θ' such that $\zeta\theta' = \theta$. Because $\theta_0 = \Phi^{-1} \circ \gamma_0$ for some $\gamma_0 \in \Gamma_{m_0,D_0}$, we see from the definitions of Φ and ζ that $\theta_0 = \zeta\theta_0$, hence we deduce that

$$\|\theta_0\|_{\mathcal{H}_\varepsilon} \leq (\eta_\varepsilon/\varepsilon)\|\theta_0\|_{\mathcal{H}} \leq M\eta_\varepsilon/\varepsilon.$$

By Corollary 2.6.18 in [16], we then have

$$\Pi(\|\theta - \theta_0\|_\infty \leq A\eta_\varepsilon^2) \geq e^{-\frac{1}{2}\|\theta_0\|^2_{\mathcal{H}_\varepsilon}}\Pi(\|\theta\|_\infty \leq A\eta_\varepsilon^2)$$
$$\geq e^{-\frac{1}{2}M^2(\eta_\varepsilon/\varepsilon)^2}\Pi'(\|\theta'\|_\infty \leq A\frac{\eta_\varepsilon^3}{\varepsilon}).$$

Next, since $\mathcal{H}$ embeds continuously into $H^\alpha(I_d)$ for some large enough cube I_d (by a standard extension argument for Sobolev spaces), the unit ball $B_{\mathcal{H}}$ of $\mathcal{H}$ has covering numbers with respect to the supremum norm $N = N(B_{\mathcal{H}}, \|\cdot\|_\infty, \delta)$ (i.e. the smallest number of $\|\cdot\|_\infty$ balls of radius δ needed to cover $B_{\mathcal{H}}$) satisfying

$$\log N(B_{\mathcal{H}}, \|\cdot\|_\infty, \delta) \leq K\delta^{-d/\alpha} \tag{38}$$

for some constant $K = K(\alpha, D)$ (cf. after Corollary 4.3.38 in [16]). We can thus apply [26], Theorem 1.2, to see

$$\Pi'\big(\|\theta'\|_\infty \leq A\tfrac{\eta_\varepsilon^3}{\varepsilon}\big) \geq e^{-A'(\frac{\eta_\varepsilon^3}{\varepsilon})^{-s}},$$

for some constant $A' = A'(A, K)$, where s is such that $\frac{d}{\alpha} = \frac{2s}{2+s}$, i.e. $s = \frac{2d}{2\alpha-d}$.

Overall, we have shown

$$\Pi(B_{KL}^\varepsilon(\eta_\varepsilon)) \geq e^{-\frac{1}{2}M^2(\eta_\varepsilon/\varepsilon)^2} e^{-A'(\eta_\varepsilon^3/\varepsilon)^{-2d/(2\alpha-d)}},$$

where the constant A' depends only on $D, \alpha, M, m_1, D_1, r$ and Φ. For $\eta_\varepsilon = \varepsilon^{\alpha/(\alpha+d)}$ we find

$$(\eta_\varepsilon^3/\varepsilon)^{-2d/(2\alpha-d)} = (\eta_\varepsilon/\varepsilon)^2,$$

and the result follows. $\qquad\square$

5.4. Posterior asymptotics.

5.4.1. Posterior regularity and contraction about Λ_{γ_0}. The following two results follow ideas from Bayesian nonparametric statistics [14, 48] combined with the lemmas from the previous section. Together with the stability estimate Lemma 7, these two estimates allow us to proceed as in [30] to prove Theorem 3.

Lemma 12. *Let η_ε, ω be as in Lemma 11. Under the assumptions of Theorem 3 there exists $M' > 0$ such that*

$$\sup_{\gamma_0 \in \Gamma_{m_0, D_0} \cap \{\Phi \circ \theta : \theta \in \mathcal{H}, \|\theta\|_{\mathcal{H}} \leq M\}} P_\varepsilon^{\gamma_0}\big(\Pi\big(\|\gamma\|_{H^\beta(D)} > M' \mid Y\big) > e^{-(\omega+4)(\eta_\varepsilon/\varepsilon)^2}\big)$$
$$\to 0, \quad (39)$$

as $\varepsilon \to 0$. The bound (39) also holds with the supremum norm $\|\cdot\|_\infty$ in place of the $H^\beta(D)$ norm.

Theorem 13. *Let η_ε, ω be as in Lemma 11. Under the conditions of Theorem 3, there exists $C > 0$ such that*

$$\sup_{\gamma_0 \in \Gamma_{m_0, D_0} \cap \{\Phi \circ \theta : \theta \in \mathcal{H}, \|\theta\|_{\mathcal{H}} \leq M\}} P_\varepsilon^{\gamma_0}\big(\Pi\big(\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r} > C\eta_\varepsilon \mid Y\big) > 2e^{-(\omega+4)(\eta/\varepsilon)^2}\big)$$
$$\to 0, \quad (40)$$

as $\varepsilon \to 0$.

To prove the preceding results, we note that the posterior from (14) can be written as

$$\Pi(B \mid Y) = \frac{\int_B p_\varepsilon^\gamma(Y)/p_\varepsilon^{\gamma_0}(Y)\,d\Pi(\gamma)}{\int_{\Gamma_{m_1, D_1}} p_\varepsilon^\gamma(Y)/p_\varepsilon^{\gamma_0}(Y)\,d\Pi(\gamma)}, \quad B \subset \Gamma_{m_1, D_1} \text{ measurable.} \quad (41)$$

The following lemma controls the size of the denominator in the last expression.

Lemma 14. *Let η_ε, ω be as in Lemma 11. Introduce the event*

$$L = L_{\omega,\gamma_0} = \left\{ \int_{\Gamma_{m_1,D_1}} \frac{p_\varepsilon^\gamma}{p_\varepsilon^{\gamma_0}}(Y)\, d\Pi(\gamma) \geq e^{-(\omega+2)(\eta_\varepsilon/\varepsilon)^2} \right\}.$$

Then, under the assumptions of Theorem 3, we have

$$\sup_{\gamma_0 \in \Gamma_{m_0,D_0} \cap \{\Phi \circ \theta : \theta \in \mathcal{H}, \|\theta\|_{\mathcal{H}} \leq M\}} P_\varepsilon^{\gamma_0}(L^c) \to 0, \quad as\ \varepsilon \to 0. \tag{42}$$

Proof. Given Lemma 11, the proof is a standard argument based on Chebyshev's inequality, Jensen's inequality and Fubini's theorem (see e.g. [16, Lemma 7.3.4] for a proof in the setting of white noise on $L^2([0,1])$ which adapts straightforwardly to the current Hilbert space setting, with probability measure ν there equal to the renormalised restriction of $d\Pi(\gamma)$ to $B_{KL}^\varepsilon(\eta_\varepsilon)$). $\qquad\square$

Proof of Lemma 12. Define L as in Lemma 14. By (41), using Fubini's theorem and the fact that $E_\varepsilon^{\gamma_0} \frac{p_\varepsilon^\gamma}{p_\varepsilon^{\gamma_0}}(Y) = 1$, we see that for every Borel set B

$$E_\varepsilon^{\gamma_0}\left[\mathbb{1}_L \Pi(B \mid Y)\right] \leq e^{(\omega+2)(\eta_\varepsilon/\varepsilon)^2} E_\varepsilon^{\gamma_0} \int_B \frac{p_\varepsilon^\gamma}{p_\varepsilon^{\gamma_0}}(Y)\, d\Pi(\gamma) = e^{(\omega+2)(\eta_\varepsilon/\varepsilon)^2} \Pi(B).$$

Then, setting $B = \{\|\gamma\|_{H^\beta(D)} > M'\}$, an application of Markov's inequality yields

$$P_\varepsilon^{\gamma_0}\left(\Pi\left(\|\gamma\|_{H^\beta(D)} > M' \mid Y\right) > e^{-(\omega+4)(\eta_\varepsilon/\varepsilon)^2}\right)$$
$$\leq P_\varepsilon^{\gamma_0}(L^c) + e^{(2\omega+6)(\eta_\varepsilon/\varepsilon)^2} \Pi\left(\|\gamma\|_{H^\beta(D)} > M'\right).$$

The first term on the right vanishes asymptotically, uniformly across γ_0 in the given set, by Lemma 14. For the second, we recall (18) and the definition (15) of the prior. In conjunction with the facts that $\varepsilon^{-d/(\alpha+d)} = \eta_\varepsilon/\varepsilon$ for our choice of η_ε and that

$$\|\zeta\theta'\|_{H^\beta(D)} \leq C \|\zeta\|_{H^\beta(D)} \|\theta'\|_{H^\beta(D)}$$

for some constant C (since $\beta > d/2$), these allow us to deduce that

$$\Pi\left(\|\gamma\|_{H^\beta(D)} > M'\right) \leq \Pi'\left(\|\theta'\|_{H^\beta(D)} > \tfrac{\eta_\varepsilon}{\varepsilon}(C\|\zeta\|_{H^\beta(D)})^{-1}(M'/C' - 1)^{1/\beta}\right).$$

Since $\eta_\varepsilon/\varepsilon \to \infty$ and since $\Pi'(H^\beta(D)) = 1$ by hypothesis, we can apply a version of Fernique's theorem, more specifically Theorem 2.1.20 in [16] (see also Example 2.1.6 therein) to deduce that for any $c > 0$ there exists a $M' = M'(\beta, c, C', \zeta)$ such that the last probability does not exceed $e^{-c(\eta_\varepsilon/\varepsilon)^2}$. Taking $c > 2\omega + 6$ concludes the proof for the $H^\beta(D)$ norm, and the result for the supremum norm follows by a Sobolev embedding and adjusting the constant M'. $\qquad\square$

Proof of Theorem 13. We decompose in a standard way: writing

$$S = \{\gamma \in \Gamma_{m_1, D_1} : \|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r} > C\eta_\varepsilon, \|\gamma\|_\infty \le M'\},$$

for M' the constant of Lemma 12 and C a large enough constant (to be chosen below), we have, for ψ the test given by Lemma 8 and L as in Lemma 14,

$$\Pi(\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r} > C\eta_\varepsilon \mid Y) \le \Pi(S \mid Y) + \Pi(\|\gamma\|_\infty > M' \mid Y),$$
$$\Pi(S \mid Y) \le \mathbb{1}_{L^c} + \psi + \Pi(S \mid Y)\mathbb{1}_L[1 - \psi]. \tag{43}$$

Hence, denoting by $\mathcal{C}$ the event $\{\Pi(\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r} > C\eta_\varepsilon \mid Y) > 2e^{-(\omega+4)(\eta_\varepsilon/\varepsilon)^2}\}$, we have

$$P_\varepsilon^{\gamma_0}(\mathcal{C}) \le P_\varepsilon^{\gamma_0}\big(\Pi(\|\gamma\|_\infty > M' \mid Y) > e^{-(\omega+4)(\eta_\varepsilon/\varepsilon)^2}\big)$$
$$+ P_\varepsilon^{\gamma_0}\big(\Pi(S \mid Y) > e^{-(\omega+4)(\eta_\varepsilon/\varepsilon)^2}\big),$$

$$P_\varepsilon^{\gamma_0}\big(\Pi(S \mid Y) > e^{-(\omega+4)(\eta_\varepsilon/\varepsilon)^2}\big) \le P_\varepsilon^{\gamma_0}(L^c) + E_\varepsilon^{\gamma_0}\psi$$
$$+ P_\varepsilon^{\gamma_0}\big(\Pi(S \mid Y)\mathbb{1}_L[1 - \psi] > e^{-(\omega+4)(\eta_\varepsilon/\varepsilon)^2}\big).$$

In view of Lemmas 8, 12, and 14, it suffices to show that there exists a constant C such that

$$P_\varepsilon^{\gamma_0}\big(\Pi(S \mid Y)\mathbb{1}_L[1 - \psi] > e^{-(\omega+4)(\eta_\varepsilon/\varepsilon)^2}\big) \to 0,$$

uniformly across γ_0 in the given set, and we note that $\|\gamma_0\|_\infty$ is uniformly bounded in the set by a Sobolev embedding. Appealing to (41), Fubini's theorem and again Lemma 8 we have for every $\kappa > 0$ and for C large enough (depending on κ),

$$E_\varepsilon^{\gamma_0}\big[\Pi(S \mid Y)\mathbb{1}_L(1 - \psi)\big] \le e^{(\omega+2)(\eta_\varepsilon/\varepsilon)^2} E_\varepsilon^{\gamma_0} \int_S \frac{p_\varepsilon^\gamma}{p_\varepsilon^{\gamma_0}}(Y)(1 - \psi)(Y)\, \mathrm{d}\Pi(\gamma)$$
$$\le e^{(\omega+2)(\eta_\varepsilon/\varepsilon)^2} \int_S E_\varepsilon^\gamma\big[(1 - \psi)(Y)\big] \mathrm{d}\Pi(\gamma)$$
$$\le e^{(\omega+2-\kappa)(\eta_\varepsilon/\varepsilon)^2},$$

hence by Markov's inequality, the probability in question is bounded by $e^{(2\omega+6-\kappa)(\eta_\varepsilon/\varepsilon)^2}$. This tends to zero, uniformly in γ_0, if C is large enough to permit $\kappa > 2\omega + 6$. $\square$

5.4.2. Proof of Theorem 3. Recall Lemma 7 to the effect that

$$\|\gamma - \gamma_0\|_\infty \le C'|\log\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r}|^{-\delta},$$

at least for $\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r}$ small enough, for some $\delta = \delta(d, \beta) > 0$ and a constant C' depending only on M, D, D_1, α, β, r, and an upper bound for $\|\gamma\|_{H^\beta(D)}$, where we have also used $\|\gamma_0\|_{H^\beta(D)} \le \|\gamma_0\|_{H^\alpha(D)}$, eq. (18) and the hypothesis on γ_0. Thus,

for $M' > 0$ the constant of Lemma 12 and $C_1 > 0$ large enough that Theorem 13 applies with constant C_1, we have

$$\left(\|\gamma\|_{H^\beta(D)} \le M'\right) \wedge \left(\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r} \le C_1\eta_\varepsilon\right) \implies \|\gamma - \gamma_0\|_\infty \le C'\xi_{\varepsilon,\delta}.$$

In view of (17), for ε small enough (noting that since $C'\xi_{\varepsilon,\delta} \to 0$ it is eventually smaller than $m_0 - m_1$), we further deduce for a constant C that

$$\Pi\left(\|\theta - \theta_0\|_\infty > C\xi_{\varepsilon,\delta} \mid Y\right) \le \Pi\left(\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r} > C_1\eta_\varepsilon \mid Y\right) \\ + \Pi\left(\|\gamma\|_{H^\beta(D)} > M' \mid Y\right). \quad (44)$$

Then Theorem 13 and Lemma 12 imply the contraction rate (22).

To prove consistency of the posterior mean $E[\theta \mid Y]$, introduce, for ω, L as in Lemma 14, the event

$$\mathcal{A} = L \cap \left\{\Pi\left(\|\theta - \theta_0\|_\infty > C\xi_{\varepsilon,\delta} \mid Y\right) \le 3e^{-(\omega+4)(\eta_\varepsilon/\varepsilon)^2}\right\},$$

and note that

$$P_\varepsilon^{\gamma_0}\left(\|E^\Pi[\theta \mid Y] - \theta_0\|_\infty > K\xi_{\varepsilon,\delta}\right) \\ \le P_\varepsilon^{\gamma_0}(\mathcal{A}^c) + P_\varepsilon^{\gamma_0}\left(\|E^\Pi[\theta - \theta_0 \mid Y]\|_\infty \mathbb{1}_\mathcal{A} > K\xi_{\varepsilon,\delta}\right). \quad (45)$$

In view of (44), observe that $\mathcal{A}^c$ is contained in

$$L^c \cup \left\{\Pi\left(\|\Lambda_\gamma - \Lambda_{\gamma_0}\|_{\mathbb{H}_r} > C_1\eta_\varepsilon \mid Y\right) > 2e^{-(\omega+4)(\eta_\varepsilon/\varepsilon)^2}\right\} \\ \cup \left\{\Pi\left(\|\gamma\|_{H^\beta(D)} > M' \mid Y\right) > e^{-(\omega+4)(\eta_\varepsilon/\varepsilon)^2}\right\}, \quad (46)$$

hence, by Lemmas 12 and 14, and Theorem 13, $P_\varepsilon^{\gamma_0}(\mathcal{A}^c) \to 0$ as $\varepsilon \to 0$, uniformly in γ_0. Now for the second term in (45), by Jensen's inequality and the Cauchy–Schwarz inequality, we have

$$\|E^\Pi[\theta - \theta_0 \mid Y]\|_\infty \mathbb{1}_\mathcal{A} \\ \le C\xi_{\varepsilon,\delta} + E^\Pi\left[\|\theta - \theta_0\|_\infty \mathbb{1}\{\|\theta - \theta_0\|_\infty > C\xi_{\varepsilon,\delta}\} \mid Y\right]\mathbb{1}_\mathcal{A} \\ \le C\xi_{\varepsilon,\delta} + \left(E^\Pi\left[\|\theta - \theta_0\|_\infty^2 \mid Y\right]\right)^{1/2}\Pi\left(\|\theta - \theta_0\|_\infty > C\xi_{\varepsilon,\delta} \mid Y\right)^{1/2}\mathbb{1}_\mathcal{A},$$

and by Markov's inequality, it follows for $K > C$ that

$$P_\varepsilon^{\gamma_0}\left(\|E^\Pi[\theta - \theta_0 \mid Y]\|_\infty \mathbb{1}_\mathcal{A} > K\xi_{\varepsilon,\delta}\right) \\ \le \frac{1}{(K-C)\xi_{\varepsilon,\delta}} E_\varepsilon^{\gamma_0}\left[\left(E^\Pi\left[\|\theta - \theta_0\|_\infty^2 \mid Y\right]\right)^{1/2}\left(\Pi\{\|\theta - \theta_0\|_\infty > C\xi_{\varepsilon,\delta} \mid Y\}\right)^{1/2}\mathbb{1}_\mathcal{A}\right]. \quad (47)$$

Again applying the Cauchy–Schwarz inequality the last expected value is bounded by

$$E_\varepsilon^{\gamma_0}\big[E^\Pi\big[\|\theta - \theta_0\|_\infty^2 \mid Y\big]\mathbb{1}_{\mathcal{A}}\big]^{1/2} E_\varepsilon^{\gamma_0}\big[\Pi\big(\|\theta - \theta_0\|_\infty > C\xi_{\varepsilon,\delta} \mid Y\big)\mathbb{1}_{\mathcal{A}}\big]^{1/2}.$$

From the definition of the event $\mathcal{A}$ it is immediate that

$$E_\varepsilon^{\gamma_0}\big[\Pi\big(\|\theta - \theta_0\|_\infty > C\xi_{\varepsilon,\delta} \mid Y\big)\mathbb{1}_{\mathcal{A}}\big] \le 3e^{-(\omega+4)(\eta_\varepsilon/\varepsilon)^2},$$

while, applying (41) and Fubini's theorem and since $E_\varepsilon^{\gamma_0}\frac{p_\varepsilon^{\Phi\circ\theta}}{p_\varepsilon^{\gamma_0}}(Y) = 1$, we have

$$E_\varepsilon^{\gamma_0}\big[E^\Pi\big[\|\theta - \theta_0\|_\infty^2 \mid Y\big]\mathbb{1}_{\mathcal{A}}\big]$$
$$\le e^{(\omega+2)(\eta_\varepsilon/\varepsilon)^2} E_\varepsilon^{\gamma_0}\bigg[\int_{C_u(D)} \|\theta - \theta_0\|_\infty^2 \frac{p_\varepsilon^{\Phi\circ\theta}}{p_\varepsilon^{\gamma_0}}(Y)\,\mathrm{d}\Pi(\theta)\bigg]$$
$$\le e^{(\omega+2)(\eta_\varepsilon/\varepsilon)^2} E^\Pi\|\theta - \theta_0\|_\infty^2.$$

Plugging this back into (47), we see

$$P_\varepsilon^{\gamma_0}\big(\|E^\Pi[\theta - \theta_0 \mid Y]\|_\infty\mathbb{1}_{\mathcal{A}} > K\xi_{\varepsilon,\delta}\big)$$
$$\le \frac{\sqrt{3}}{(K-C)\xi_{\varepsilon,\delta}}\big(e^{(\omega+2)(\eta_\varepsilon/\varepsilon)^2} E^\Pi\|\theta - \theta_0\|_\infty^2 e^{-(\omega+4)(\eta_\varepsilon/\varepsilon)^2}\big)^{1/2}.$$

Note that, since $E^\Pi\|\theta - \theta_0\|_\infty^2 \le 2(\|\theta_0\|_\infty^2 + E^\Pi\|\theta\|_\infty^2)$ and $E^{\Pi'}\|\theta'\|_\infty^2$ is finite ([16, Exercise 2.1.2 and Theorem 2.1.20a]), a Sobolev embedding combined with the prior definition (15) implies that $E^\Pi\|\theta - \theta_0\|_\infty^2$ is bounded uniformly across the specified θ_0's. Since $e^{-(\eta_\varepsilon/\varepsilon)^2}/\xi_{\varepsilon,\delta} \to 0$, we see, returning to (45), that the result follows.

5.5. Proof of the lower bound Theorem 2. Recall the shorthand (20) and also the definition of $K(p, q)$ from before Lemma 9. It is enough to find $\gamma_0, \gamma_1 \in \Gamma_{m_0, D_0}^\alpha(M)$ (both allowed to depend on ε) such that, for some μ small enough (to be chosen),

i. $\|\gamma_1 - \gamma_0\|_\infty \ge \xi_{\varepsilon,\delta'} \equiv \log(1/\varepsilon)^{-\delta'}$,

ii. $K(p_\varepsilon^{\gamma_1}, p_\varepsilon^{\gamma_0}) \le \mu$.

Indeed, the standard proof of this reduction (for example as in [16, Theorem 6.3.2] or [46, Chapter 2]) is as follows. Under condition (i), noting that

$$\psi = \mathbb{1}\big\{\|\widetilde{\gamma} - \gamma_1\|_\infty < \|\widetilde{\gamma} - \gamma_0\|_\infty\big\}$$

yields a test of $H_0: \gamma = \gamma_0$ against $H_1: \gamma = \gamma_1$, we see

$$\inf_{\widetilde{\gamma}} \sup_{\gamma \in \Gamma_{m_0, D_0}^\alpha(M)} P_\varepsilon^\gamma\big(\|\widetilde{\gamma} - \gamma\|_\infty \ge \tfrac{1}{2}\xi_{\varepsilon,\delta'}\big) \ge \inf_\psi \max\big(P_\varepsilon^{\gamma_0}(\psi \ne 0), P_\varepsilon^{\gamma_1}(\psi \ne 1)\big),$$

where the latter infimum is over all tests ψ. Introducing the event $A = \left\{ \frac{p_\varepsilon^{\gamma_0}}{p_\varepsilon^{\gamma_1}} \geq 1/2 \right\}$, we see

$$P_\varepsilon^{\gamma_0}(\psi \neq 0) \geq E_\varepsilon^{\gamma_1}\left[\frac{p_\varepsilon^{\gamma_0}}{p_\varepsilon^{\gamma_1}} \mathbb{1}_A \psi \right] \geq \tfrac{1}{2}\left[P_\varepsilon^{\gamma_1}(\psi = 1) - P_\varepsilon^{\gamma_1}(A^c) \right].$$

Thus, writing $p_1 = P_\varepsilon^{\gamma_1}(\psi = 1)$, we see

$$\max\left(P_\varepsilon^{\gamma_0}(\psi \neq 0), P_\varepsilon^{\gamma_1}(\psi \neq 1) \right) \geq \max\left(\tfrac{1}{2}(p_1 - P_\varepsilon^{\gamma_1}(A^c)), 1 - p_1 \right)$$
$$\geq \inf_{p \in [0,1]} \max\left(\tfrac{1}{2}(p - P_\varepsilon^{\gamma_1}(A^c)), 1 - p \right).$$

The infimum is attained when $\tfrac{1}{2}(p - P_\varepsilon^{\gamma_1}(A^c)) = 1 - p$ and takes the value $\tfrac{1}{3} P_\varepsilon^{\gamma_1}(A)$ so that

$$\inf_{\widetilde{\gamma}} \sup_{\gamma \in \Gamma_{m_0, D_0}^\alpha(M)} P_\varepsilon^\gamma\left(\|\widetilde{\gamma} - \gamma\|_\infty \geq \tfrac{1}{2}\xi_{\varepsilon,\delta'} \right) \geq \tfrac{1}{3} P_\varepsilon^{\gamma_1}(A). \qquad (48)$$

Next observe

$$P_\varepsilon^{\gamma_1}(A) = P_\varepsilon^{\gamma_1}\left[\frac{p_\varepsilon^{\gamma_1}}{p_\varepsilon^{\gamma_0}} \leq 2 \right]$$
$$= 1 - P_\varepsilon^{\gamma_1}\left[\log\left(\frac{p_\varepsilon^{\gamma_1}}{p_\varepsilon^{\gamma_0}} \right) > \log 2 \right]$$
$$\geq 1 - P_\varepsilon^{\gamma_1}\left[\left| \log\left(\frac{p_\varepsilon^{\gamma_1}}{p_\varepsilon^{\gamma_0}} \right) \right| > \log 2 \right]$$
$$\geq 1 - (\log 2)^{-1} E_\varepsilon^{\gamma_1}\left| \log\left(\frac{p_\varepsilon^{\gamma_1}}{p_\varepsilon^{\gamma_0}} \right) \right|,$$

where we have used Markov's inequality to attain the final expression. By the second Pinsker inequality ([16, Proposition 6.1.7b]), using condition (ii) we can continue the chain of inequalities to see

$$P_\varepsilon^{\gamma_1}(A) \geq 1 - (\log 2)^{-1}\left[K(p_\varepsilon^{\gamma_1}, p_\varepsilon^{\gamma_0}) + \sqrt{2K(p_\varepsilon^{\gamma_1}, p_\varepsilon^{\gamma_0})} \right] \geq 1 - (\log 2)^{-1}(\mu + \sqrt{2\mu}).$$

Choosing μ small enough, we can thus lower bound (48) by

$$\frac{1}{3}\left(1 - \frac{\mu + \sqrt{2\mu}}{\log 2} \right) > \frac{1}{4},$$

so that Theorem 2 will follow.

Now we prove the existence of $\gamma_0, \gamma_1 \in \Gamma_{m_0, D_0}^\alpha(M)$ satisfying conditions (i) and (ii). We appeal to Corollary 1 in [28], which says that for any integer $k \geq 2$, any $q \geq 0$, some $B > 0$ and any $\xi \in (0, 1)$ there exist γ_0, γ_1 such that

$$\mathrm{supp}(\gamma_j - 1) \subseteq D_0, \quad \gamma_j \geq 1 \text{ on } D \text{ for } j = 0, 1,$$

and

a. $\|\gamma_1 - \gamma_0\|_\infty \geq \xi$,

b. $\|\Lambda_{\gamma_1} - \Lambda_{\gamma_0}\|_{H^{-q}(\partial D)/\mathbb{C} \to H^q_\diamond(\partial D)} \leq \exp\left(-\xi^{-\frac{d}{(2d-1)k}}\right)$,

c. $\max\left(\|\gamma_1\|_{C^k(D)}, \|\gamma_0\|_{C^k(D)}\right) \leq B$,

where the norm in (b) is the operator norm and where $\|\cdot\|_{C^k(D)}$ denotes the usual norm on the space $C^k(D)$ of functions with bounded continuous partial derivatives up to order k. (Note that [28] states this with full space $H^{-q}(\partial D)$ in place of the quotient space, but since Λ_{γ_j} maps constant functions to 0 for $j = 0, 1$, the two operator norms coincide.) For M sufficiently large, we may take $k = \alpha$ to deduce (noting that $C^\alpha(D)$ continuously embeds into $H^\alpha(D)$) that there exist $\gamma_0, \gamma_1 \in \Gamma^\alpha_{m_0, D_0}(M)$ satisfying (a) and (b). Taking $\xi = \xi_{\varepsilon, \delta'}$ we note that (i) holds by definition.

For (ii), applying Lemma 18 with

$$s = 0, \quad p = \min(d - 1, r),$$
$$\text{and} \quad q = (d - 1 - r)_+ \equiv \max(d - 1 - r, 0) = d - 1 - p,$$

we see that, for a constant $C = C(d, r)$,

$$\|\Lambda_{\gamma_1} - \Lambda_{\gamma_0}\|_{\mathbb{H}_r} \leq C \|\Lambda_{\gamma_1} - \Lambda_{\gamma_0}\|_{H^{-q}(\partial D)/\mathbb{C} \to H^q_\diamond(\partial D)}.$$

Thus, appealing to Lemma 9, we can bound the KL-divergences $K(p_\varepsilon^{\gamma_1}, p_\varepsilon^{\gamma_0})$ by

$$\varepsilon^{-2} \|\Lambda_{\gamma_1} - \Lambda_{\gamma_0}\|_{\mathbb{H}_r}^2 \leq C^2 \varepsilon^{-2} \|\Lambda_{\gamma_1} - \Lambda_{\gamma_0}\|_{H^{-q} \to H^q}^2$$
$$\leq C^2 \exp\left[2\log(1/\varepsilon) - (\log(1/\varepsilon))^{-\frac{\delta' d}{(2d-1)\alpha}}\right].$$

Since $\delta' > \alpha(2d - 1)/d$ by assumption, the final expression tends to zero as $\varepsilon \to 0$ and in particular is less than the μ of condition (ii) for ε small enough.

A. Laplace–Beltrami eigenfunctions and Sobolev spaces

In this appendix, we define the Sobolev spaces $H^r(D)$, $H^r(\partial D)$ for a bounded domain $D \subset \mathbb{R}^d$, with smooth boundary ∂D.

Definition ($H^r(D)$). We follow [27] to define these Sobolev spaces: see Chapter I, Sections 1.1, 9.1, and 12.1 (pp. 1, 40, and 70, respectively) for details. For $r \in \mathbb{N} \cup \{0\}$ we set

$$H^r(D) = \left\{ f \in L^2(D) : \|f\|_{H^r(D)}^2 = \sum_{|\alpha| \leq r} \|D^\alpha f\|_{L^2(D)}^2 < \infty \right\},$$

where for a multi-index $\alpha = (\alpha_1, \ldots, \alpha_d)$, $\alpha_j \in \mathbb{N} \cup \{0\}$ for $j \leq d$, of order $|\alpha| = \sum_{j \leq d} \alpha_j$,

$$D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \ldots \partial x_d^{\alpha_d}},$$

with partial derivatives defined in a weak sense. $L^2(D)$ is defined with respect to the Lebesgue measure on D. Note $H^r(D)$ is a Hilbert space, with inner product

$$\langle f, g \rangle_{H^r(D)} = \sum_{|\alpha| \leq r} \int D^\alpha f \cdot \overline{D^\alpha g}.$$

For $r \in \mathbb{R}$, $r \geq 0$ we then define $H^r(D)$ via interpolation. Finally, defining

$$H_0^r(D) = \left\{ f \in H^r(D) : \frac{\partial^j f}{\partial \nu^j}\big|_{\partial D} = 0, 0 \leq j < r - 1/2 \right\},$$

with the normal boundary derivatives defined in a trace sense, we define $H^r(D)$, $r < 0$, as the topological dual space $(H_0^{|r|}(D))^*$, equipped with the dual norm. (cf. Chapter I, Sections 11.1, 11.4, and 12.1 on pp. 55, 62, and 70 of [27].)

For $C_c^\infty(D)$ the space of smooth functions compactly supported in D, $H_{\text{loc}}^1(D)$ is defined as

$$H_{\text{loc}}^1(D) = \left\{ f \text{ locally integrable} : f\phi \in H^1(D) \text{ for all } \phi \in C_c^\infty(D) \right\}$$

(see [27, Chapter II, Section 3.2 on p. 125]), or, equivalently,

$$H_{\text{loc}}^1(D) = \{ f \text{ locally integrable} :$$
$$f|_U \in H^1(U) \text{ for all domains } U \text{ satisfying } \bar{U} \subset D \}.$$

To define the Sobolev space $H^r(\partial D)$ for the compact boundary manifold ∂D, let $(\phi_k = \phi_k^{(0)} : k \in \mathbb{N} \cup \{0\})$ be an orthonormal basis of $L^2(\partial D)$ consisting of real-valued eigenfunctions of the Laplace–Beltrami operator $\Delta_{\partial D}$. The basic properties of such a basis are gathered, for example, in Chavel [7], Chapter I. Let $\lambda_k \geq 0$ be the corresponding eigenvalues:

$$-\Delta_{\partial D} \phi_k = \lambda_k \phi_k.$$

(By convention these are assumed to have been sorted in increasing order.)

Definition ($H^r(\partial D)$). For $r \geq 0$, we define

$$H^r(\partial D) = \left\{ f \in L^2(\partial D) \text{ s.t. } \sum_{k=0}^\infty (1 + \lambda_k)^r |\langle f, \phi_k \rangle_{L^2(\partial D)}|^2 =: \|f\|_{H^r(\partial D)}^2 < \infty \right\},$$

where the space $L^2(\partial D)$ is defined relative to the surface element on ∂D. For $r < 0$, we define $H^r(\partial D)$ as the completion of $L^2(\partial D)$ with respect to the norm $\|\cdot\|_{H^r(\partial D)}$ (defined as in the above display).

Remarks. i. It is immediate that $\{\phi_k : k \in \mathbb{N} \cup \{0\}\}$ is an orthogonal spanning set of $H^r(\partial D)$, and that setting $\phi_k^{(r)} = (1 + \lambda_k)^{-r/2} \phi_k$ yields an orthonormal basis of $H^r(\partial D)$.

ii. This definition of $H^r(\partial D)$ coincides with other possible definitions. For example, for $r = 1$ the calculation

$$\int_{\partial D} \nabla \phi_k \cdot \overline{\nabla \phi_l} = -\int_{\partial D} \phi_k \Delta_{\partial D} \phi_l = \lambda_l \int_{\partial D} \phi_k \phi_l = \lambda_l \delta_{kl},$$

derived via the divergence theorem on a closed manifold (e.g. see [7, eq. (35)]; note that the manifold ∂D is compact) implies that our definition of $\|\cdot\|_{H^1(\partial D)}$ is equivalent to the standard definition

$$\|f\|_{H^1(\partial D)}^2 = \|f\|_{L^2(\partial D)}^2 + \|\nabla f\|_{L^2(\partial D)}^2,$$

and inductively the same is true for $H^r(\partial D)$, $r \in \mathbb{N}$.

For the equivalence of this definition with some other definitions of negative or non-integer Sobolev spaces, see [27, Chapter I, Section 7.3 on pp. 34–37]. In particular note that $H^{-s}(\partial D)$ is the topological dual space of $H^s(\partial D)$ for any $s \in \mathbb{R}$.

iii. Note that ϕ_0 is a constant function, hence the $H^r(\partial D)/\mathbb{C}$ norm, defined by

$$\|[f]\|_{H^r(\partial D)/\mathbb{C}} = \inf_{z \in \mathbb{C}} \|f - z\|$$

for $[f]$ the equivalence class over $\mathbb{C}$ of a function $f \in H^r(\partial D)$, can also be characterised as

$$\|[f]\|_{H^r(\partial D)/\mathbb{C}}^2 = \sum_{k=1}^{\infty} (1 + \lambda_k)^r |\langle f, \phi_k \rangle_{L^2(\partial D)}|^2. \tag{49}$$

Recall also

$$H_\diamond^s(\partial D) = \left\{ g \in H^s(\partial D) : \langle g, 1 \rangle_{L^2(\partial D)} = 0 \right\}$$

(see (5)). Note that each $[f] \in H^s(\partial D)/\mathbb{C}$ has a representative $g \in H_\diamond^s(\partial D)$, and

$$\|[f]\|_{H^s(\partial D)/\mathbb{C}} = \|g\|_{H^s(\partial D)}.$$

We thus use the norm (49) on spaces $H^s(\partial D)/\mathbb{C}$ and on $H_\diamond^s(\partial D)$ without further mention. We also typically write f for the equivalence class $[f]$ and only comment further on this where necessary.

This 'spectral' definition of $H^r(\partial D)$ is useful particularly because Weyl's law explicitly specifies the scaling of λ_k as $k \to \infty$.

Lemma 15 (Weyl's law on a compact closed manifold, e.g. [7, eq. (49)]). *Suppose M is a closed compact manifold of dimension d. Then*

$$N(\lambda) = (2\pi)^{-d} \lambda^{d/2} \omega_d \operatorname{Vol}(M) + o(\lambda^{d/2}),$$

where $N(\lambda)$ is the number of eigenvalues (counted with multiplicity) no bigger than λ and ω_d is the volume of a unit disc in $\mathbb{R}^d$.

Corollary 16. *The eigenvalues of the Laplace–Beltrami operator $\Delta_{\partial D}$ satisfy*

$$C_1 k^{2/(d-1)} \leq \lambda_k \leq C_2 k^{2/(d-1)}$$

for constants C_1, C_2 depending only on D. Hence, the eigenfunctions satisfy

$$C_3\big(1 + k^{\frac{1}{d-1}}\big)^{s-r} \leq \|\phi_k^{(r)}\|_{H^s(\partial D)} \leq C_4\big(1 + k^{\frac{1}{d-1}}\big)^{s-r}, \quad s, r \in \mathbb{R}, \qquad (50)$$

for constants C_3 and C_4 depending only on ∂D and on the difference $s - r$. For $k > 0$ the same expression holds with the quotient norm $\|\phi_k^{(r)}\|_{H^s(\partial D)/\mathbb{C}}$ in place of $\|\phi_k^{(r)}\|_{H^s(\partial D)}$.

Proof. We apply Weyl's law on the manifold ∂D, which has dimension $d - 1$. Writing $N(\lambda^-)$ for $\lim_{x \uparrow \lambda} N(x)$ and $N(\lambda^+)$ for $\lim_{x \downarrow \lambda} N(x)$, we thus have

$$N(\lambda_k^-) \leq k \leq N(\lambda_k^+).$$

It follows that

$$C\lambda_k^{(d-1)/2} + o\big(\lambda_k^{(d-1)/2}\big) \leq k \leq C\lambda_k^{(d-1)/2} + o\big(\lambda_k^{(d-1)/2}\big)$$

for the constant $C = C(D) = (2\pi)^{-(d-1)}\omega_{d-1}\,\mathrm{Area}(\partial D)$ and hence we deduce the scaling of the eigenvalues. Then (50) follows from the definition of $H^r(\partial D)$ and the remarks thereafter. $\qquad\square$

B. Comparison results for Hilbert–Schmidt operators

For separable Hilbert spaces A and B, we use the notation $\mathcal{L}(A, B)$ for the space of bounded linear maps $A \to B$ equipped with the operator norm $\|\cdot\|_{A \to B}$, and $\mathcal{L}_2(A, B)$ for the space of Hilbert–Schmidt operators $A \to B$ equipped with inner product $\langle \cdot, \cdot \rangle_{\mathcal{L}_2(A,B)}$ (e.g. see [4, Chapter 12]). Define the orthonormal basis $(b_{jk}^{(p,q)})$ of $\mathcal{L}_2(H^p, H^q)$ by

$$b_{jk}^{(p,q)}(f) = \phi_j^{(p)} \otimes \phi_k^{(q)}(f) = \langle f, \phi_j^{(p)} \rangle_{H^p} \phi_k^{(q)}, \quad j, k \in \mathbb{N},$$

(in this appendix we omit explicit reference to the domain, writing H^p for either $H^p(\partial D)/\mathbb{C}$ or $H_\diamond^p(\partial D)$; as in remark (iii) in Appendix A, both spaces can be identified with $\overline{\mathrm{span}}\{\phi_k^{(p)} : k \geq 1\}$, hence the omission should not cause confusion). The compatibility between our bases of $H^\alpha(\partial D)$ for different $\alpha \in \mathbb{R}$ means that the subspaces spanned by $(b_{jk}^{(p,q)})_{j \leq J, k \leq K}$ coincide for all p and q, and the $\mathcal{L}_2(H^p, H^q)$-orthogonal projections onto this subspace coincide with π_{JK} (as

defined in (24)). Corollary 16 implies the following lemmas controlling Hilbert–Schmidt norms for different domains and codomains in terms of each other, and in terms of operator norms.

Lemma 17. *Let $p, q, r, s \in \mathbb{R}$ and let $T \in \mathrm{span}\{b_{jk}^{(p,q)} : 1 \leq j \leq J, \ 1 \leq k \leq K\}$. Then there is a constant C depending only on D and on the differences $r - p$, $s - q$ such that*

$$\|T\|_{\mathcal{L}_2(H^r, H^s)} \leq C\left(1 + J^{1/(d-1)}\right)^{(p-r)_+}\left(1 + K^{1/(d-1)}\right)^{(s-q)_+}\|T\|_{\mathcal{L}_2(H^p, H^q)},$$

where $x_+ = \max(x, 0)$ for $x \in \mathbb{R}$.

Proof. The coefficients $a_{jk}^{(r,s)}$ of T with respect to the basis $(b_{jk}^{(r,s)})$ are given by

$$a_{jk}^{(r,s)} = \langle T, b_{jk}^{(r,s)}\rangle_{\mathcal{L}_2(H^r, H^s)} \equiv \sum_p \langle T\phi_p^{(r)}, b_{jk}^{(r,s)}(\phi_p^{(r)})\rangle_{H^s} = \langle T\phi_j^{(r)}, \phi_k^{(s)}\rangle_{H^s}$$

and we see from Corollary 16 that

$$|a_{jk}^{(r,s)}| \leq C\left(1 + j^{1/(d-1)}\right)^{p-r}\left(1 + k^{1/(d-1)}\right)^{s-q}|a_{jk}^{(p,q)}| \tag{51}$$

for a constant C depending only on D and the differences $r - p$, $s - q$.

Upper bounding

$$\left(1 + j^{1/(d-1)}\right)^{(p-r)} \leq \left(1 + J^{1/(d-1)}\right)^{(p-r)_+}$$

for $j \leq J$, and similarly for k, we find that

$$\begin{aligned}
\|T\|_{\mathcal{L}_2(H^r, H^s)}^2 &= \sum_{j \leq J, k \leq K} |a_{jk}^{(r,s)}|^2 \\
&\leq C\left(1 + J^{1/(d-1)}\right)^{2(p-r)_+}\left(1 + K^{1/(d-1)}\right)^{2(s-q)_+}\|T\|_{\mathcal{L}_2(H^p, H^q)}^2,
\end{aligned}$$

hence the result. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Lemma 18. *For $p, q, r, s \in \mathbb{R}$ satisfying $p \leq r$ and $q \geq s$, let $T \in \mathcal{L}(H^{p-(d-1)}, H^q)$. Then we have $T \in \mathcal{L}_2(H^r, H^s)$ with, for constant C depending only on D and the differences $r - p$, $q - s$,*

$$\|T\|_{\mathcal{L}_2(H^r, H^s)} \leq C\|T\|_{H^{p-(d-1)} \to H^q},$$

and moreover

$$\begin{aligned}
\|T - \pi_{JK}T\|_{\mathcal{L}_2(H^r, H^s)} & \\
&\leq C\|T\|_{H^{p-(d-1)} \to H^q} \max\left((1 + J^{1/(d-1)})^{p-r}, (1 + K^{1/(d-1)})^{s-q}\right).
\end{aligned}$$

Proof. Firstly, as a consequence of Corollary 16, we have, for a constant $C = C(D)$,

$$\|T\phi_j^{(p)}\|_{H^q}^2 \le \|T\|_{H^{p-(d-1)} \to H^q}^2 \|\phi_j^{(p)}\|_{H^{p-(d-1)}}^2$$
$$\le C\|T\|_{H^{p-(d-1)} \to H^q}^2 \left(1 + j^{1/(d-1)}\right)^{-2(d-1)}, \qquad (52)$$

which is summable over j, hence by definition of the space of Hilbert–Schmidt operators,

$$T \in \mathcal{L}_2(H^p, H^q) \quad \text{and} \quad \|T\|_{\mathcal{L}_2(H^p, H^q)} \le C'\|T\|_{H^{p-(d-1)} \to H^q}.$$

That T lies in $\mathcal{L}_2(H^r, H^s)$ and satisfies the the specified bound follows by monotonicity of H^α norms.

Since the $\mathcal{L}_2(H^r, H^s)$-orthogonal projection maps coincide for all r and s, next defining $a_{jk}^{(r,s)}$ as in the previous proof, we have from (51) that for a constant C,

$$\|T - \pi_{JK}T\|_{\mathcal{L}_2(H^r, H^s)}^2 = \sum_{j > J \text{ or } k > K} |a_{jk}^{(r,s)}|^2$$
$$\le C \sum_{j > J \text{ or } k > K} \left(1 + j^{1/(d-1)}\right)^{2(p-r)} \left(1 + k^{1/(d-1)}\right)^{2(s-q)} |a_{jk}^{(p,q)}|^2.$$

Since $p \le r$ and $q \ge s$, we see that

$$\sum_{j > J} \sum_k \left(1 + j^{1/(d-1)}\right)^{2(p-r)} \left(1 + k^{1/(d-1)}\right)^{2(s-q)} |a_{jk}^{(p,q)}|^2$$
$$\le \left(1 + J^{1/(d-1)}\right)^{2(p-r)} \sum_{j > J} \sum_k |a_{jk}^{(p,q)}|^2$$
$$\le \left(1 + J^{1/(d-1)}\right)^{2(p-r)} \|T\|_{\mathcal{L}_2(H^p, H^q)}^2.$$

Arguing similarly for the sum over all j and over $k > K$, we deduce that

$$\|T - \pi_{JK}T\|_{\mathcal{L}_2(H^r, H^s)}^2$$
$$\le 2C\|T\|_{\mathcal{L}_2(H^p, H^q)}^2 \max\left(\left(1 + J^{1/(d-1)}\right)^{2(p-r)}, \left(1 + K^{1/(d-1)}\right)^{2(s-q)}\right).$$

The result follows. $\qquad \square$

C. Mapping properties of Λ_γ and $\widetilde{\Lambda}_\gamma$

In this appendix we prove the following mapping properties of Λ_γ and $\widetilde{\Lambda}_\gamma$ which were used throughout the main body of the paper. The proofs rely on basic theory for elliptic PDEs, which can be found in the many books on the subject (we shall refer mostly to [27]).

Lemma 19. *Let $s \in \mathbb{R}$, let $m \in (0,1)$ and let D' be a domain compactly contained in D. For $\gamma \in \Gamma_{m,D'}$ and $f \in H^{s+1}(\partial D)/\mathbb{C}$, there is a unique weak solution*

$$u_{\gamma,f} \in \mathcal{H}_s = \left(H^{\min\{1,s+3/2\}}(D) \cap H^1_{\mathrm{loc}}(D) \right)/\mathbb{C}$$

to the Dirichlet problem (1). *Moreover, if $u_{1,f}$ is the unique solution when $\gamma = 1$, then for any other $\gamma \in \Gamma_{m,D'}$ bounded by M on D, $u_{\gamma,f} - u_{1,f}$ lies in $H^1_0(D)/\mathbb{C}$ and satisfies the estimate*

$$\|u_{\gamma,f} - u_{1,f}\|_{H^1(D)/\mathbb{C}} \le C \frac{M}{m} \|f\|_{H^{s+1}(\partial D)/\mathbb{C}}, \tag{53}$$

for some constant $C = C(D, D', s)$.

Lemma 19 is neither novel nor necessarily sharp, but we require explicit bounds controlling how the constants depend on γ for our proofs, which is why the result is given here.

Lemma 20. *For some $m \in (0,1)$ and some domain D' compactly contained in D, let $\gamma \in \Gamma_{m,D'}$. For each $s \in \mathbb{R}$, Λ_γ is a continuous linear map from $H^{s+1}(\partial D)/\mathbb{C}$ to $H^s_\diamond(\partial D)$, and it is continuously invertible. For each $s, t \in \mathbb{R}$, the shifted operator $\widetilde{\Lambda}_\gamma = \Lambda_\gamma - \Lambda_1$ is a continuous map from $H^s(\partial D)/\mathbb{C}$ to $H^t_\diamond(\partial D)$.*

Moreover, if γ also satisfies the bound $\|\gamma\|_\infty \le M$, then we have the explicit bounds

$$\|\Lambda_\gamma\|_{H^{s+1}(\partial D)/\mathbb{C} \to H^s(\partial D)} \le C_1 \frac{M}{m}, \tag{54}$$

$$\|\widetilde{\Lambda}_\gamma\|_{H^s(\partial D)/\mathbb{C} \to H^t(\partial D)} \le C_2 \frac{M}{m}, \tag{55}$$

for constants $C_1 = C(D, D', s)$ and $C_2 = C_2(D, D', s, t)$.

Given Lemma 20, the following is an immediate consequence of Lemma 18 (recall $\mathbb{H}_r$ is defined in (10), and note that the proofs of the previous lemmas imply that $\widetilde{\Lambda}_\gamma$ maps real-valued functions to real-valued functions).

Lemma 21. *For any $r \in \mathbb{R}$ and any $\gamma \in \Gamma_{m,D'}$, $\widetilde{\Lambda}_\gamma \in \mathbb{H}_r$.*

A key to proving Lemmas 19 and 20 is the following basic fact about harmonic functions. For convenience of the reader, we include a proof (following Lemma A.1 in [19]). Recall that a function h is *harmonic* on some domain Ω if $\Delta h = 0$ on Ω, where Δ denotes the Laplacian. Note that as we have assumed $\gamma = 1$ on a neighbourhood $D \setminus D'$ of ∂D, our solutions $u_{\gamma,f}$ (and in particular, the *differences* $u_{\gamma,f} - u_{\gamma_0,f}$ between solutions) are harmonic on this neighbourhood.

Lemma 22 (Interior smoothness of harmonic functions). *Let U_0, U be bounded domains such that $\bar{U} \subseteq U_0$. Then for any $s, t \in \mathbb{R}$, there is a constant $C = C(s, t, U, U_0)$ such that for any harmonic function $v \in H^s(U)$,*

$$\|v\|_{H^s(U)/\mathbb{C}} \le C \|v\|_{H^t(U_0)/\mathbb{C}}.$$

Proof. By monotonicity of H^t norms it suffices to prove the result for $s = t + k$ for $k \in \mathbb{N}$. Let $v \in H^t(U_0)$ represent the equivalence class and choose a domain U_1 such that

$$\bar{U} \subseteq U_1 \subseteq \bar{U}_1 \subseteq U_0.$$

Let ϕ be a smooth cutoff function, identically one on U_1 and compactly supported in U_0. For $z \in \mathbb{C}$ we observe that $\tilde{v} := (v - z)\phi$ satisfies

$$\Delta \tilde{v} = F \quad \text{in } U_0,$$
$$\tilde{v} = 0 \quad \text{on } \partial U_0,$$

where $F = 2 \nabla \phi \cdot \nabla v + (v - z)\Delta \phi$. Then

$$\|v\|_{H^{t+1}(U_1)/\mathbb{C}} \leq \|v - z\|_{H^{t+1}(U_1)} \leq \|\tilde{v}\|_{H^{t+1}(U_0)} \leq C\|F\|_{H^{t-1}(U_0)},$$

by standard elliptic boundary value regularity results (e.g. [27, Chapter II, Remark 7.2 on p. 188] with $N = \{0\}$ there as we are considering the standard Laplace equation; in the case $t < 1$ Remark 7.2 gives the result with a different norm in place of the $H^{t-1}(U_0)$ norm on the right, but the two norms are equivalent when restricted to functions of compact support in U_0). Note

$$\|F\|_{H^{t-1}(U_0)} \leq C(\phi)\big(\|v\|_{H^t(U_0)/\mathbb{C}} + \|v - z\|_{H^{t-1}(U_0)}\big),$$

and optimising across $z \in \mathbb{C}$ yields

$$\|v\|_{H^{t+1}(U_1)/\mathbb{C}} \leq C\|v\|_{H^t(U_0)/\mathbb{C}}. \tag{56}$$

Finally, we choose a finite sequence of domains $(U_j)_{1 \leq j \leq k}$ such that $U_k = U$ and, for $1 \leq j \leq k$, $\bar{U}_j \subset U_{j-1}$; applying (56) successively on each pair (U_j, U_{j-1}), we deduce the result. $\qquad\square$

Proof of Lemma 19. We adapt the proof of Theorem A.2 from Hanke et al. [19] to the Dirichlet setting here. From standard theory for the Laplacian, for $\gamma = 1$ and $f \in H^{s+1}(\partial D)/\mathbb{C}$ there exists a solution $u_{1,f} \in H^{s+3/2}(D)/\mathbb{C}$ to the Dirichlet problem (1), and this solution satisfies

$$\|u_{1,f}\|_{H^{s+3/2}(D)/\mathbb{C}} \leq C\|f\|_{H^{s+1}(\partial D)/\mathbb{C}} \tag{57}$$

for a constant $C = C(D, s)$ (e.g. see [27, Chapter II, Remark 7.2 on p. 188] with $N = \{0\}$). Also note that, as a harmonic function, $u_{1,f} \in H^1_{\text{loc}}(D)/\mathbb{C}$ by Lemma 22.

Define the sesquilinear operator B_γ and the conjugate linear operator A,

$$B_\gamma(w, v) = \int_D \gamma \nabla w \cdot \nabla \bar{v}, \quad \text{and} \quad A(v) = -\int_D \gamma \nabla u_{1,f} \cdot \nabla \bar{v},$$

and consider the equation

$$B_\gamma(w, v) = A(v) \quad \forall v \in H^1_0(D). \tag{58}$$

Observe that if $w \in H_0^1(D)$ solves (58), then $u_{\gamma,f} = w + u_{1,f}$ solves the weak Dirichlet problem (4); note that $u_{\gamma,f}$ so defined lies in $\mathcal{H}_s$. We will use Lax–Milgram theory to show the existence and uniqueness of such a w.

By definition any $\gamma \in \Gamma_{m,D'}$ is bounded; consistently with the second half of the lemma let M be an upper bound for $\|\gamma\|_\infty$. Let $v, w \in H_0^1(D)$. By the Cauchy–Schwarz inequality,

$$|B_\gamma(w,v)| \leq M \|w\|_{H^1(D)} \|v\|_{H^1(D)},$$

and we also note

$$B_\gamma(v,v) \geq m \|\nabla v\|_{L^2(D)}^2 \geq cm \|v\|_{H^1(D)}^2,$$

where the latter inequality, with constant $c = c(D)$, is the Poincaré inequality (e.g. [1, Corollary 6.31] applied to $v \in H_0^1(D)$). In other words, B_γ is bounded and coercive. Next, since $\gamma = 1$ on $D \setminus D'$, for $v \in H_0^1(D)$, an application of the divergence theorem yields

$$-\int_{D\setminus D'} \gamma \, \nabla u_{1,f} \cdot \nabla \overline{v} = -\int_D \nabla u_{1,f} \cdot \nabla \overline{v} + \int_{D'} \nabla u_{1,f} \cdot \nabla \overline{v}$$

$$= \int_D \overline{v} \Delta u_{1,f} - \int_{\partial D} \overline{v} \frac{\partial u_{1,f}}{\partial \nu} + \int_{D'} \nabla u_{1,f} \cdot \nabla \overline{v}$$

$$= \int_{D'} \nabla u_{1,f} \cdot \nabla \overline{v}.$$

It follows, since $\|1 - \gamma\|_\infty \leq \|\gamma\|_\infty \leq M$, that

$$|A(v)| = \left| \int_{D'} -\gamma \, \nabla u_{1,f} \cdot \nabla \overline{v} - \int_{D\setminus D'} \gamma \, \nabla u_{1,f} \cdot \nabla \overline{v} \right|$$

$$= \left| \int_{D'} (1 - \gamma) \, \nabla u_{1,f} \cdot \nabla \overline{v} \right|$$

$$\leq M \|v\|_{H^1(D)} \|\nabla u_{1,f}\|_{L^2(D')}.$$

By Lemma 22 and recalling (57), there are constants C and C' depending only on D, D' and s such that

$$\|\nabla u_{1,f}\|_{L^2(D')} \leq \|u_{1,f}\|_{H^1(D')/\mathbb{C}}$$
$$\leq C' \|u_{1,f}\|_{H^{s+3/2}(D)/\mathbb{C}} \leq C \|f\|_{H^{s+1}(\partial D)/\mathbb{C}}.$$

Thus,

$$|A(v)| \leq CM \|f\|_{H^{s+1}(\partial D)} \|v\|_{H^1(D)}.$$

We deduce from the Lax–Milgram theorem (e.g. [11, Theorem 1 in Section 6.2.1]; note while the theorem there is stated for real scalars and bilinear maps, the same

proof works for complex scalars and sesquilinear maps) that (58) has a unique solution $w \in H_0^1(D)$. Moreover, the equation $B_\gamma(w,w) = A(w)$ shows that $\|w\|_{H^1(D)}$ is upper bounded the operator norm of A divided by the coercivity constant of B_γ, yielding (53).

It remains to show that the (equivalence class of) function(s) $u_{\gamma,f}$ so constructed is the unique solution in $\mathcal{H}_s$ to (4). Since we have shown uniqueness of w, it is enough to show that the difference h between two $\mathcal{H}_s$ solutions lies in $H_0^1(D)$, as then it must be the zero function. (We are considering h as a function, rather than an equivalence class of functions, which we can do by for example choosing a representative with average zero.) This is clear for $s \geq -1/2$ as then $\mathcal{H}_s = H^1(D)/\mathbb{C}$, and can be shown also for $s < -1/2$ as in [19], Theorem A.2. $\square$

Proof of Lemma 20. We first remark, for $\gamma \in \Gamma_{m,D'}$, that by the divergence theorem,

$$\langle \tfrac{\partial u}{\partial \nu}, 1 \rangle_{L^2(\partial D)} = \int_{\partial D} \gamma \tfrac{\partial u}{\partial \nu} = \int_D \nabla \cdot (\gamma \nabla u) = 0$$

for a solution u to the Dirichlet problem (1), so that it suffices to prove (54), (55), and the continuity of $\Lambda_\gamma^{-1} \colon H_\diamond^s(\partial D) \to H^{s+1}(\partial D)/\mathbb{C}$.

We first prove (55), by adapting the proof of Theorem A.3 from [19] and tracking the constants. Given $f \in H^s(\partial D)$ let $u_{\gamma,f} \in \mathcal{H}_{s-1}$ be the unique solution to the Dirichlet problem (1) and let $w \in H_0^1$ be a representative of the function class $u_{\gamma,f} - u_{1,f}$. Choose a domain Ω with smooth boundary $\partial\Omega$, satisfying

$$\bar{D}' \subset \Omega \subset \bar{\Omega} \subset D.$$

Choose also domains U, U_0 with smooth boundaries such that

$$\partial\Omega \subseteq U \subset \bar{U} \subset U_0 \subset \bar{U}_0 \subset D \setminus D'.$$

We can apply an appropriate trace theorem ([27] Chapter I, Theorem 9.4 for $t > 0$, Chapter II, Theorem 6.5 (and Remark 6.4) for $t \leq -3/2$, or Chapter II, Theorem 7.3 for $-3/2 < t < 1/2$; note in the latter two cases we use that w is harmonic on $D \setminus \bar{\Omega}$) to $w - z$ and optimise across $z \in \mathbb{C}$ to see

$$\|\partial w/\partial \nu\|_{H^t(\partial D)} \leq C \|w\|_{H^{t+3/2}(D\setminus\bar{\Omega})/\mathbb{C}}. \tag{59}$$

Applying [27, Chapter II, Remark 7.2 on p. 188] with $N = \{0\}$ we see

$$\|w\|_{H^{t+3/2}(D\setminus\bar{\Omega})/\mathbb{C}} \leq C \big(\|\mathrm{tr}\, w\|_{H^{t+1}(\partial D)/\mathbb{C}} + \|\mathrm{tr}\, w\|_{H^{t+1}(\partial\Omega)/\mathbb{C}} \big)$$
$$= C \|\mathrm{tr}\, w\|_{H^{t+1}(\partial\Omega)/\mathbb{C}}.$$

Again applying an appropriate trace theorem, this time on a subset of U bounded on one side by $\partial\Omega$, and applying Lemma 22, we see

$$\|\mathrm{tr}\, w\|_{H^{t+1}(\partial\Omega)/\mathbb{C}} \leq C \|w\|_{H^{t+3/2}(U)/\mathbb{C}} \leq C' \|w\|_{H^1(U_0)/\mathbb{C}}.$$

The constants in the above depend on D and (via Ω, U and U_0) on D', but are otherwise independent of γ. Recalling (53), which, because the smoothness of f here is s, tells us

$$\|w\|_{H^1(U_0)/\mathbb{C}} \leq C\left(\frac{M}{m}\right)\|f\|_{H^s(D)/\mathbb{C}},$$

we overall have

$$\|\partial w/\partial \nu\|_{H^t(\partial D)} \leq C\left(\frac{M}{m}\right)\|f\|_{H^s(\partial D)/\mathbb{C}},$$

so that we have proved (55).

Now we prove (54); given (55), it suffices to show

$$\|\Lambda_1\|_{H^{s+1}(\partial D)/\mathbb{C} \to H^s(\partial D)} \leq C\frac{M}{m}$$

for an appropriate constant C. Since $u_{1,f}$ is harmonic on D, applying the trace theorems from [27] as for (59) yields

$$\|\partial u_{1,f}/\partial \nu\|_{H^s(\partial D)} \leq C\|u_{1,f}\|_{H^{s+3/2}(D)/\mathbb{C}}$$

for a constant $C = C(D, s)$. Then [27, Chapter II, Remark 7.2] with $N = \{0\}$ yields

$$\|u_{1,f}\|_{H^{s+3/2}(D)/\mathbb{C}} \leq C'\|f\|_{H^{s+1}(\partial D)/\mathbb{C}}$$

for a constant $C' = C'(D, s)$, and (54) follows.

Finally we remark that the same arguments (see [19, Theorem A.3]) applied to the inverse of Λ_γ, which is the *Neumann-to-Dirichlet map*, show that this is continuous from $H^s_\diamond(\partial D)$ to $H^{s+1}(\partial D)/\mathbb{C}$ as claimed. □

D. Asymptotic comparison of noise models

In this appendix we formulate and prove the asymptotic comparison results discussed informally in Section 2.

D.1. A brief overview of the Le Cam distance. We first define the Le Cam deficiency and the Le Cam distance, providing the sense in which the models will be shown to be asymptotically related to each other. The concepts throughout this appendix are drawn from Le Cam's 1986 monograph [25]. We refer to the expository paper of Mariucci [29] for a gentler introduction to the area.

Definitions. Statistical experiment. A *statistical experiment*, or just *experiment*, is the triple $(\mathcal{X}, \mathcal{F}, \{P_\theta\}_{\theta \in \Theta})$, where for each θ in the parameter space Θ, P_θ is a probability measure on the measurable space $(\mathcal{X}, \mathcal{F})$.

Markov Kernel. For measurable spaces $(\mathcal{X}_i, \mathcal{F}_i)$, $i = 1, 2$, a *Markov kernel with source $(\mathcal{X}_1, \mathcal{F}_1)$ and target $(\mathcal{X}_2, \mathcal{F}_2)$* is a map $T: \mathcal{X}_1 \times \mathcal{F}_2 \to [0, 1]$ such that $T(x, \cdot)$ is a probability measure for each $x \in \mathcal{X}_1$, and $T(\cdot, A)$ is measurable for each $A \in \mathcal{F}_2$.

Given a (deterministic) measurable function $F: \mathcal{X}_1 \to \mathcal{X}_2$ we will denote by T_F the Markov kernel

$$T_F(x, A) = \mathbb{1}\{F(x) \in A\}. \tag{60}$$

Le Cam deficiency. The *Le Cam deficiency* between experiments $\mathcal{E}_1$ and $\mathcal{E}_2$, where $\mathcal{E}_i = (\mathcal{X}_i, \mathcal{F}_i, \{P_{i,\theta}\}_{\theta \in \Theta})$ for $i = 1, 2$, for a common parameter space Θ, is

$$\delta(\mathcal{E}_1, \mathcal{E}_2) = \inf_T \sup_{\theta \in \Theta} \|TP_{1,\theta} - P_{2,\theta}\|_{\mathrm{TV}},$$

where the infimum is over all Markov kernels with source $(\mathcal{X}_1, \mathcal{F}_1)$ and target $(\mathcal{X}_2, \mathcal{F}_2)$. The measure $TP_{1,\theta}$ is defined as

$$TP_{1,\theta}(A) = \int_{\mathcal{X}_1} T(x, A) \, dP_{1,\theta}(x),$$

and $\|\cdot\|_{\mathrm{TV}}$ denotes the total variation norm on signed measures,

$$\|\nu\|_{\mathrm{TV}} = \sup_A |\nu(A)|.$$

The Le Cam deficiency satisfies the triangle inequality, but is not symmetric.

Le Cam distance. The *Le Cam distance* between experiments $\mathcal{E}_1$ and $\mathcal{E}_2$ on a common parameter space Θ is

$$\Delta(\mathcal{E}_1, \mathcal{E}_2) = \max(\delta(\mathcal{E}_1, \mathcal{E}_2), \delta(\mathcal{E}_2, \mathcal{E}_1)).$$

If we identify experiments whose Le Cam distance is zero, this defines a proper metric.

Remark. Given any action set A, any loss function $L: \Theta \times A \to [0, 1]$, and any decision rule $\rho_2: \mathcal{X}_2 \to A$, there exists a (possibly randomised) decision rule $\rho_1: \mathcal{X}_1 \to A$ such that, denoting the risk functions by

$$R_j(\rho_j, \theta) = E_{X \sim P_{j,\theta}} L(\theta, \rho_j(X)), \quad j = 1, 2,$$

we have (see [29, Theorem 2.7])

$$R_1(\rho_1, \theta) \leq R_2(\rho_2, \theta) + \delta(\mathcal{E}_1, \mathcal{E}_2), \quad \forall \theta \in \Theta. \tag{61}$$

We may further take the supremum over θ and then the infimum over ρ_2 (note that the infimum over decision rules ρ_1 associated as above to decision rules ρ_2 upper bounds the infimum over all decision rules ρ_1) to see that the minimax risks

$$R_{j,\mathrm{minimax}} = \inf_{\rho_j} \sup_\theta R_j(\rho_j, \theta), \quad j = 1, 2$$

satisfy

$$R_{1,\text{minimax}} \le R_{2,\text{minimax}} + \delta(\mathcal{E}_1, \mathcal{E}_2). \tag{62}$$

These equations capture the intuitive definition that the Le Cam deficiency is the worst-case error we incur when reconstructing a decision rule in $\mathcal{E}_2$ using data from $\mathcal{E}_1$ (see also Theorem 27 where this intuition is made concrete for the models considered here).

We gather the key tools we will use to control Le Cam deficiencies in the following lemma. Recall that $K(p, q)$ denotes the Kullback–Leibler divergence between distributions with densities p and q; in an abuse of notation we will in this section also write $K(P, Q)$ for the Kullback–Leibler divergence between distributions P and Q.

Lemma 23. *Let $\mathcal{E}_1$ and $\mathcal{E}_2$ be experiments with a common parameter set Θ: write $\mathcal{E}_j = (\mathcal{X}_j, \mathcal{F}_j, \{P_{j,\theta}\}_{\theta \in \Theta})$.*

a. *Suppose further that the experiments are defined on a common probability space, i.e. that $\mathcal{X}_1 = \mathcal{X}_2$ and $\mathcal{F}_1 = \mathcal{F}_2$. Then*

$$\Delta(\mathcal{E}_1, \mathcal{E}_2) \le \sup_{\theta \in \Theta} \| P_{1,\theta} - P_{2,\theta} \|_{\text{TV}} \le \sup_{\theta \in \Theta} \sqrt{K(P_{1,\theta}, P_{2,\theta})/2}. \tag{63}$$

b. *Let $F: \mathcal{X}_1 \to \mathcal{X}_2$ be any (deterministic) measurable map. Then*

$$\delta(\mathcal{E}_1, \mathcal{E}_2) \le \sup_{\theta \in \Theta} \| P_{1,\theta} \circ F^{-1} - P_{2,\theta} \|_{\text{TV}}. \tag{64}$$

c. *Let $F: \mathcal{X}_1 \to \mathcal{X}_2$ be a measurable map. Suppose that $P_{1,\theta} \circ F^{-1} = P_{2,\theta}$ for each $\theta \in \Theta$ and suppose that $F(X)$ is a sufficient statistic for $X \sim P_{1,\theta}$. Then $\Delta(\mathcal{E}_1, \mathcal{E}_2) = 0$.*

Proof. a. The first inequality is immediate from the definition, since the Markov kernel T_{Id} corresponding to the identity map $\text{Id}: \mathcal{X}_1 \to \mathcal{X}_2 = \mathcal{X}_1$ satisfies $T_{\text{Id}} P = P$ for all probability measures P on $(\mathcal{X}_1, \mathcal{F}_1)$. The second inequality is Pinsker's inequality (e.g. [16, Proposition 6.1.7a]).

b. Observe that $T_F P_{1,\theta}(A) = P_{1,\theta}(F(X) \in A) = P_{1,\theta} \circ F^{-1}(A)$. The result follows.

c. See [29, Property 3.12]. $\qquad \square$

D.2. Le Cam deficiency bounds for noisy Calderón problems. First we recall the noise models under consideration. The following displays are labelled with the notation we will use for experiments with the specified data, in each case taking the parameter space to be

$$\{\gamma \in \Gamma_{m,D'} : \|\gamma\|_\infty \le M\}$$

for some constants $m \in (0, 1), M > 1$ and some domain D' compactly contained in D.

For noise level $\varepsilon > 0$ and some $P \in \mathbb{N}$, in the 'electrode' model (7) we are given data

$$\mathcal{E}_0 : Y_{p,q} = \langle \widetilde{\Lambda}_\gamma[\psi_p], \psi_q \rangle_{L^2(\partial D)} + \varepsilon g_{p,q}, \quad p, q \leq P, \; g_{p,q} \overset{iid}{\sim} N(0, 1),$$

where $\psi_p = c_p \mathbb{1}_{I_p}$ for some disjoint (measurable) sets $I_p = I_{p,P} \subseteq \partial D$, with constants c_p chosen such that the ψ_p are $L^2(\partial D)$ orthonormal. For some $\varepsilon > 0$, $r \in \mathbb{R}$, and $J, K \in \mathbb{N}$, in the 'discrete spectral' model (8) we are given data

$$\mathcal{E}_1^{(r)} : Y_{j,k} = \langle \widetilde{\Lambda}_\gamma[\phi_j^{(r)}], \phi_k^{(0)} \rangle_{L^2(\partial D)} + \varepsilon g_{j,k}, \quad j \leq J, k \leq K, \; g_{j,k} \overset{iid}{\sim} N(0, 1),$$

for a (real-valued) Laplace–Beltrami basis $(\phi_k^{(r)} : k \in \mathbb{N})$ of $H^r(\partial D)/\mathbb{C}$. For $\varepsilon > 0$ and $r \in \mathbb{R}$, in the 'continuous' model (12) (see also the equivalent (11)) we are given data

$$\mathcal{E}_2^{(r)} : Y = \widetilde{\Lambda}_\gamma + \varepsilon \mathbb{W}, \quad \mathbb{W} \text{ a Gaussian white noise indexed by } \mathbb{H}_r.$$

We start with the following result which shows how to approximate the discrete spectral model using the electrode model.

Theorem 24. *Suppose* $\cup_{p \leq P} I_p = \partial D$ *and* $\mathrm{diam}(I_p) \leq (A/P)^{1/(d-1)}$ *for a constant A independent of P, where* $\mathrm{diam}\, S = \sup_{x,y \in S} |x - y|$ *denotes the Euclidean diameter of a subset* $S \subset \mathbb{R}^d$. *Then the one-way Le Cam deficiency* $\delta(\mathcal{E}_0, \mathcal{E}_1^{(0)})$ *satisfies*

$$\delta(\mathcal{E}_0, \mathcal{E}_1^{(0)}) \leq C \left(\max(J, K)^{(5d-2)/(2d-2)} + \varepsilon^{-1} \max(J, K)^{3d/(2d-2)} \right) P^{-1/(d-1)},$$

for some constant $C = C(A, D', D, M, m)$, *and hence vanishes asymptotically if P is large enough compared to* ε^{-1}, *J and K.*

Remarks. i. The conditions on $(I_p)_{p \leq P}$ are only used to prove that we can approximate any Laplace–Beltrami eigenfunction using a linear combination of the $(\psi_p)_{p \leq P}$ with $L^2(\partial D)$-norm approximation error proportional to $P^{-1/(d-1)}$ (Lemma 25). If $(\psi_p)_{p \leq P}$ are such that we can approximate Laplace–Beltrami eigenfunctions at a rate $f(P)$ then we achieve the result with $f(P)$ in place of $P^{-1/(d-1)}$.

ii. The given conditions are naturally satisfied by 'evenly spaced' sets $(I_p)_{p \leq P}$ partitioning the boundary ∂D, with a constant A depending only on the domain D. This can be seen by considering the covering numbers $N(\partial D, d_{\partial D}, \delta)$ (the smallest number of $d_{\partial D}$ balls of radius δ needed to cover ∂D) for $d_{\partial D}$ the geodesic distance. Theorem 4.5 in Geller & Pensenson [13] applied to the current setting yields

$$N(\partial D, d_{\partial D}, \delta) \leq A \delta^{-(d-1)}$$

for a constant $A = A(D)$, for any $\delta > 0$. Taking $\delta = 2(A/P)^{1/(d-1)}$ we deduce that there exist P balls of $(d_{\partial D})$-radius $\delta/2$ covering D. To construct P disjoint subsets of diameter at most δ, we simply assign each $x \in \partial D$ to exactly one of the balls containing it (note the Euclidean diameter is upper bounded by the geodesic diameter, because any geodesic path on the surface ∂D is also a path in Euclidean space).

Proof. Let $(Y_{p,q})$ be the data from experiment $\mathcal{E}_0$. Let ϕ_j^P denote the L^2-orthogonal projection of $\phi_j^{(0)}$ onto $\mathrm{span}\{\psi_p : p \leq P\}$, and write

$$a_{jp} = \langle \phi_j^{(0)}, \psi_p \rangle_{L^2(\partial D)} \in \mathbb{R},$$

so that $\phi_j^P = \sum_{p=1}^P a_{jp}\psi_p$. Define $F: \mathbb{R}^{P \times P} \to \mathbb{R}^{J \times K}$ via

$$F((u_{pq})_{p,q \leq P})_{jk} = \sum_{p,q \leq P} a_{jp}a_{kq}u_{pq}.$$

For parameter set $\{\gamma \in \Gamma_{m,D'} : \|\gamma\|_\infty \leq M\}$, let $\mathcal{E}_0'$ denote the experiment with data

$$\mathcal{E}_0': Y_{j,k}' = F((Y_{p,q})_{p,q \leq P})_{jk} = \langle \widetilde{\Lambda}_\gamma \phi_j^P, \phi_k^P \rangle_{L^2(\partial D)} + \varepsilon g_{j,k}', \tag{65}$$

where we define $g_{j,k}' = \sum_{p,q} a_{jp}a_{kq}g_{p,q}$, and let $\mathcal{E}_1'$ denote the experiment with data (65) but for i.i.d. Gaussian noise. By Lemma 23b. we see that $\delta(\mathcal{E}_0, \mathcal{E}_0') = 0$, so by the triangle inequality we deduce

$$\delta(\mathcal{E}_0, \mathcal{E}_1^{(0)}) \leq \delta(\mathcal{E}_0, \mathcal{E}_0') + \delta(\mathcal{E}_0', \mathcal{E}_1') + \delta(\mathcal{E}_1', \mathcal{E}_1^{(0)}) \leq \Delta(\mathcal{E}_0', \mathcal{E}_1') + \Delta(\mathcal{E}_1', \mathcal{E}_1^{(0)}).$$

We control the terms on the right.

$\mathbf{\Delta(\mathcal{E}_0', \mathcal{E}_1').}$ The covariance of $(g_{j,k}')$ is given by

$$\mathrm{Cov}(g_{j,k}', g_{l,m}') = \langle \phi_j^P, \phi_l^P \rangle_{L^2(\partial D)} \langle \phi_k^P, \phi_m^P \rangle_{L^2(\partial D)}.$$

Writing

$$\langle \phi_j^P, \phi_l^P \rangle_{L^2(\partial D)} = \langle \phi_j^{(0)}, \phi_l^{(0)} \rangle_{L^2(\partial D)} + \langle \phi_j^{(0)}, \phi_l^P - \phi_l^{(0)} \rangle_{L^2(\partial D)}$$
$$+ \langle \phi_j^P - \phi_j^{(0)}, \phi_l^P \rangle_{L^2(\partial D)},$$

and applying the Cauchy–Schwarz inequality (note also that $\|\phi_l^P\|_{L^2(\partial D)} \leq \|\phi_l^{(0)}\|_{L^2(\partial D)} = 1$) we see that

$$|\langle \phi_j^P, \phi_l^P \rangle_{L^2(\partial D)} - \delta_{jl}| \leq \|\phi_j^P - \phi_j^{(0)}\|_{L^2(\partial D)} + \|\phi_l^P - \phi_l^{(0)}\|_{L^2(\partial D)}.$$

A similar decomposition further yields

$$|\langle \phi_j^P, \phi_l^P \rangle_{L^2(\partial D)} \langle \phi_k^P, \phi_m^P \rangle - \delta_{jl}\delta_{km}| \leq 4 \max_{i \leq \max(J,K)} \|\phi_i^P - \phi_i^{(0)}\|_{L^2(\partial D)}.$$

Lemma 25 controls this L^2 approximation error, yielding that for a constant $C = C(A, D)$,

$$|\mathrm{Cov}(g_{j,k}', g_{l,m}') - \delta_{jl}\delta_{km}| \leq C \max(J, K)^{(1+d/2)/(d-1)} P^{-1/(d-1)}.$$

Thus, controlling the Le Cam distance between Gaussian experiments with equal means by $\sqrt{2}$ times the Frobenius distance between the covariance matrices (e.g. as in the proof of Theorem 3.1 in [41]) yields, for constants C', c',

$$\begin{aligned}
\Delta(\mathcal{E}_0', \mathcal{E}_1') &\leq \sqrt{2}\Big(\sum_{j,l \leq J, \, k,m \leq K} \big(\mathrm{Cov}(g_{j,k}', g_{l,m}') - \delta_{jl}\delta_{km} \big)^2 \Big)^{1/2} \\
&\leq c' JK \max(J, K)^{(1+d/2)/(d-1)} P^{-1/(d-1)} \\
&\leq C' \max(J, K)^{(5d-2)/(2d-2)} P^{-1/(d-1)}.
\end{aligned}$$

$\boldsymbol{\Delta(\mathcal{E}_1', \mathcal{E}_1^{(0)})}.$ Explicitly calculating the Kullback–Leibler divergence between multivariate normals with the same covariance matrix (cf. the similar calculation in Lemma 9) and using Lemma 23a. yields

$$\Delta(\mathcal{E}_1', \mathcal{E}_1^{(0)}) \leq \tfrac{1}{2}\varepsilon^{-1}$$
$$\times \sup_{\gamma \in \Gamma_{m,D'}: \|\gamma\|_\infty \leq M} \big\| \big(\langle \widetilde{\Lambda}_\gamma \phi_j^{(0)}, \phi_k^{(0)} \rangle_{L^2(\partial D)} - \langle \widetilde{\Lambda}_\gamma \phi_j^P, \phi_k^P \rangle_{L^2(\partial D)} \big)_{j \leq J, k \leq K} \big\|_{\mathbb{R}^{J \times K}},$$

where the norm on the right is the usual Frobenius or Hilbert–Schmidt norm on the space of $J \times K$ matrices. By Lemma 20, $\|\widetilde{\Lambda}_\gamma\|_{L^2(\partial D) \to L^2(\partial D)}$ is bounded by a constant $C = C(D, D', M, m)$, hence applying also Lemma 25 (as well as the Cauchy–Schwarz inequality) we have for a different constant $C' = C'(A, D, D', M, m)$,

$$\begin{aligned}
\big(\langle \widetilde{\Lambda}_\gamma \phi_j^{(0)}, \phi_k^{(0)} \rangle_{L^2(\partial D)} &- \langle \widetilde{\Lambda}_\gamma \phi_j^P, \phi_k^P \rangle_{L^2(\partial D)} \big)^2 \\
&= \big(\langle \widetilde{\Lambda}_\gamma (\phi_j^{(0)} - \phi_j^P), \phi_k^{(0)} \rangle_{L^2(\partial D)} + \langle \widetilde{\Lambda}_\gamma \phi_j^P, \phi_k^{(0)} - \phi_k^P \rangle_{L^2(\partial D)} \big)^2 \\
&\leq C^2 \big(\|\phi_j^{(0)} - \phi_j^P\|_{L^2(\partial D)} + \|\phi_k^{(0)} - \phi_k^P\|_{L^2(\partial D)} \big)^2 \\
&\leq C' \max(J, K)^{(2+d)/(d-1)} P^{-2/(d-1)}.
\end{aligned}$$

Summing over j and k we deduce

$$\Delta(\mathcal{E}_1', \mathcal{E}_1^{(0)}) \leq C'\varepsilon^{-1} \max(J, K)^{3d/(2d-2)} P^{-1/(d-1)},$$

concluding the proof. $\qquad\square$

Lemma 25. *Under the hypotheses of Theorem 24, let ϕ_j^P denote the L^2-orthogonal projection of $\phi_j^{(0)}$ onto $\mathrm{span}\{\psi_p : p \leq P\}$. Then there is a constant C depending only on the constant A of Theorem 24 and on D such that, for $j \leq \max(J, K)$,*

$$\|\phi_j^{(0)} - \phi_j^P\|_{L^2(\partial D)}^2 \leq C \max(J, K)^{(2+d)/(d-1)} P^{-2/(d-1)}. \tag{66}$$

Proof. Since ϕ_j^P as the L^2-orthogonal projection minimises the L^2 distance to $\phi_j^{(0)}$ of any function in $\mathrm{span}\{\psi_p : p \leq P\}$, for any points $x_p \in I_p$ we see

$$\|\phi_j^{(0)} - \phi_j^P\|_{L^2(\partial D)}^2 \leq \|\phi_j^{(0)} - \sum_{p=1}^P \phi_j^{(0)}(x_p)\mathbb{1}_{I_p}\|_{L^2(\partial D)}^2$$

$$\leq \max_{p \leq P}(\mathrm{diam}(I_p)^2) \sum_{p=1}^P \int_{I_p} \frac{|\phi_j^{(0)}(x) - \phi_j^{(0)}(x_p)|^2}{|x - x_p|^2} \, dx$$

$$\leq (A/P)^{2/(d-1)}\|\phi_j^{(0)}\|_{\mathrm{Lip}}^2 \, \mathrm{Area}(\partial D).$$

Using a Sobolev embedding for the compact manifold ∂D, we may estimating the Lipschitz constant of $\phi_j^{(0)}$ by a constant times $\|\phi_j^{(0)}\|_{H^\kappa(\partial D)}$ for any $\kappa > 1+(d-1)/2$. In particular, taking $\kappa = 1 + d/2$, we see that the final expression is bounded by

$$C \max(J, K)^{(2+d)/(d-1)} P^{-2/(d-1)}$$

for some $C = C(A, D)$ by Corollary 16. $\qquad\square$

The following theorem shows that the 'discrete spectral' and the continuous measurement models are very close to each other.

Theorem 26. *For any $r \in \mathbb{R}$ and any $\nu > 0$ there is a constant $C = C(\nu, r, D, D_0, M, m)$ such that the Le Cam distance $\Delta(\mathcal{E}_1^{(r)}, \mathcal{E}_2^{(r)})$ satisfies*

$$\Delta(\mathcal{E}_1^{(r)}, \mathcal{E}_2^{(r)}) \leq C\varepsilon^{-1} \min(J, K)^{-\nu}.$$

Proof. We introduce the experiments $\mathcal{E}_i^{(r)}$, $i = 3, 4$ with parameter space

$$\{\gamma \in \Gamma_{m,D'} : \|\gamma\|_\infty \leq M\}$$

corresponding to observations

$$\mathcal{E}_3^{(r)}: \left(\pi_{JK}\widetilde{\Lambda}_\gamma + \varepsilon\mathbb{W}\right)(U)_{U \in \mathbb{H}_r}$$
$$= \left(\langle \pi_{JK}\widetilde{\Lambda}_\gamma, U\rangle_{\mathbb{H}_r} + \varepsilon \sum_{j,k} g_{jk}\langle U\phi_j^{(r)}, \phi_k^{(0)}\rangle_{L^2(\partial D)}\right)_{U \in \mathbb{H}_r},$$

$$\mathcal{E}_4^{(r)}: \left(\pi_{JK}\widetilde{\Lambda}_\gamma + \varepsilon\mathbb{W}\right)(\pi_{JK}U)_{U \in \mathbb{H}_r}$$
$$= \left(\langle \pi_{JK}\widetilde{\Lambda}_\gamma, \pi_{JK}U\rangle_{\mathbb{H}_r} + \varepsilon \sum_{j \leq J, k \leq K} g_{jk}\langle U\phi_j^{(r)}, \phi_k^{(0)}\rangle_{L^2(\partial D)}\right)_{U \in \mathbb{H}_r},$$

where we recall the projection π_{JK} was defined in (24), and

$$g_{jk} = \mathbb{W}(b_{jk}^{(r)}) \stackrel{iid}{\sim} N(0, 1).$$

By the triangle inequality, we decompose

$$\Delta(\mathcal{E}_1^{(r)}, \mathcal{E}_2^{(r)}) \leq \Delta(\mathcal{E}_2^{(r)}, \mathcal{E}_3^{(r)}) + \Delta(\mathcal{E}_3^{(r)}, \mathcal{E}_4^{(r)}) + \Delta(\mathcal{E}_4^{(r)}, \mathcal{E}_1^{(r)}).$$

We control each of the terms on the right.

$\Delta(\mathcal{E}_2^{(r)}, \mathcal{E}_3^{(r)})$. Lemmas 23a. and 4, and the proof of Lemma 9 yield

$$\Delta(\mathcal{E}_2^{(r)}, \mathcal{E}_3^{(r)}) \leq \tfrac{1}{2}\varepsilon^{-1} \times \sup_{\gamma \in \Gamma_{m,D'} : \|\gamma\|_\infty \leq M} \|\widetilde{\Lambda}_\gamma - \pi_{JK}\widetilde{\Lambda}_\gamma\|_{\mathbb{H}_r}$$

$$\leq C\varepsilon^{-1} \min(J, K)^{-\nu},$$

for a constant $C = C(\nu, r, D, D', M, m)$.

$\Delta(\mathcal{E}_3^{(r)}, \mathcal{E}_4^{(r)})$. We note that $(\pi_{JK}\widetilde{\Lambda}_\gamma + \varepsilon\mathbb{W})(\pi_{JK}U)_{U \in \mathbb{H}_r}$ is a sufficient statistic for $(\pi_{JK}\widetilde{\Lambda}_\gamma + \varepsilon\mathbb{W})(U)_{U \in \mathbb{H}_r}$ by independence of $(g_{jk})_{j \leq J, k \leq K}$ from $(g_{jk} : j > J$ or $k > K)$, so that $\Delta(\mathcal{E}_3^{(r)}, \mathcal{E}_4^{(r)}) = 0$ by Lemma 23c.

$\Delta(\mathcal{E}_4^{(r)}, \mathcal{E}_1^{(r)})$. Using Lemma 23b. as in the proof of Theorem 24, it is clear that the experiment $\mathcal{E}_1^{(r)}$ is equivalent to observing

$$\left(\sum_{j \leq J, k \leq K} \left(\langle \widetilde{\Lambda}_\gamma \phi_j^{(r)}, \phi_k^{(0)}\rangle_{L^2(\partial D)} u_{jk} + \varepsilon g_{j,k} u_{jk}\right)\right)_{u \in \ell^2},$$

where ℓ_2 denotes the space of square-summable real sequences. Since

$$(g_{jk}) := \left(\mathbb{W}(b_{jk}^{(r)})\right) \stackrel{d}{=} (g_{j,k})$$

(the latter being the noise in $\mathcal{E}_1^{(r)}$) and, for $U = \sum_{j,k} u_{jk} b_{jk}^{(r)}$,

$$\sum_{j \leq J, k \leq K} \langle \widetilde{\Lambda}_\gamma \phi_j^{(r)}, \phi_k^{(0)}\rangle_{L^2(\partial D)} u_{jk} = \langle \pi_{JK}\widetilde{\Lambda}_\gamma, \pi_{JK}U\rangle_{\mathbb{H}_r},$$

we deduce $\Delta(\mathcal{E}_4^{(r)}, \mathcal{E}_1^{(r)}) = 0$. $\square$

Complementing the two previous theorems, let us remark that in fact

$$\delta\left(\mathcal{E}_2^{(r)}, \mathcal{E}_1^{(r)}\right) = 0$$

and similarly

$$\delta\left(\mathcal{E}_2^{(0)}, \mathcal{E}_0\right) = 0,$$

as is shown in the proof of the lower bound in the following theorem.

In our final result we prove that the rates obtained in Theorems 1 and 2 hold with P_ε^γ replaced by the law $P_{\varepsilon,P}^\gamma$ of the data $Y = (Y_{p,q})_{p,q \leq P}$ in the electrode model (7), for an appropriate number of electrodes P and for appropriate sets $(I_p)_{p \leq P}$. The same conclusion holds as well in the 'discrete spectral' model (8) — the proof of this fact is similar (in fact simpler) and omitted.

Theorem 27. *Let $P_{\varepsilon,P}^\gamma$ denote the law of the data $Y = (Y_{p,q})_{p,q \leq P}$ in the electrode model (7). Recall the parameter set $\Gamma_{m_0,D_0}^\alpha(M)$ defined in (3).*

A. *Suppose m_0, D_0, α, M satisfy the conditions of Theorem 1. Suppose that the sets $(I_p)_{p \leq P}$ satisfy the conditions of Theorem 24 and that $P \geq \varepsilon^{-\mu(d-1)}$ for some $\mu > 1$. Then there exists a measurable function $\widehat{\gamma}'$ of the data $Y \sim P_{\varepsilon,P}^\gamma$ such that, for C, δ the same constants as in Theorem 1,*

$$\sup_{\gamma \in \Gamma_{m_0,D_0}^\alpha(M)} P_{\varepsilon,P}^\gamma\big(\|\widehat{\gamma}' - \gamma\|_\infty > C\log(1/\varepsilon)^{-\delta}\big) \to 0, \quad as \ \varepsilon \to 0. \quad (67)$$

B. *Suppose $D_0, D, m_0, \alpha, M, c, \delta'$ are as in Theorem 2. Let $P \in \mathbb{N}$ be arbitrary, and let $(I_p)_{p \leq P}$ be arbitrary disjoint measurable sets. Then for all ε small enough,*

$$\inf_{\widetilde{\gamma}'} \sup_{\gamma \in \Gamma_{m_0,D_0}^\alpha(M)} P_{\varepsilon,P}^\gamma\big(\|\widetilde{\gamma}' - \gamma\|_\infty > c\log(1/\varepsilon)^{-\delta'}\big) > 1/4, \quad (68)$$

where the infimum extends over all measurable functions $\widetilde{\gamma}' = \widetilde{\gamma}'(Y)$ of the data $Y \sim P_{\varepsilon,P}^\gamma$.

Proof. We consider statistical experiments denoted $\mathcal{E}_0, \mathcal{E}_1^{(r)}, \mathcal{E}_2^{(r)}$ defined as at the start of Section D.2, but here for the smaller parameter space $\Gamma_{m_0,D_0}^\alpha(M)$. The definition of the Le Cam deficiency ensures that it cannot increase upon considering a smaller parameter space, so that Theorems 24 and 26 continue to hold for these experiments.

A. Consider the loss function

$$L(\gamma, \rho) = \mathbb{1}\big\{\|\gamma - \rho\|_\infty > C\log(1/\varepsilon)^{-\delta}\big\}.$$

The associated risk R of the decision rule $\widehat{\gamma}$ in the continuous model (12) is

$$P_\varepsilon^\gamma\big(\|\widehat{\gamma} - \gamma\|_\infty > C\log(1/\varepsilon)^{-\delta}\big),$$

so that $\sup_\gamma R \to 0$ as $\varepsilon \to 0$, for any $r \in \mathbb{R}$, by Theorem 1.

In view of (61), there exists a measurable function $\widehat{\gamma}'$ of the data in model (7) whose risk R' satisfies $R' \leq R + \delta(\mathcal{E}_0, \mathcal{E}_2^{(0)})$. (In fact, as the proof shows, all Markov kernels involved in bounding $\delta(\mathcal{E}_0, \mathcal{E}_2^{(0)})$ arise from deterministic maps, and hence $\widehat{\gamma}'$

can be taken to be non-randomised.) The limit (67) (which is $\sup_\gamma R' \to 0$) will follow from showing that $\delta(\mathcal{E}_0, \mathcal{E}_2^{(0)}) \to 0$. The triangle inequality gives that

$$\delta(\mathcal{E}_0, \mathcal{E}_2^{(0)}) \le \delta(\mathcal{E}_0, \mathcal{E}_1^{(0)}) + \delta(\mathcal{E}_1^{(0)}, \mathcal{E}_2^{(0)}),$$

where we may freely choose the parameters J, K of the intermediate model. For some positive constant

$$\mu' < \min\{(2d-2)(\mu-1)/(3d), (2d-2)\mu/(5d-2)\},$$

take $\nu > 1/\mu'$ and choose $J = K$ of order $\varepsilon^{-\mu'}$. Then, by Theorems 24 and 26, and the triangle inequality, we have for a constant C

$$\delta(\mathcal{E}_0, \mathcal{E}_2^{(0)}) \le C\left(\varepsilon^{\mu-1} J^{3d/(2d-2)} + J^{(5d-2)/(2d-2)}\varepsilon^\mu + \varepsilon^{-1}J^{-\nu}\right) \to 0,$$

as required.

B. Consider the loss function

$$\widetilde{L}(\gamma, \rho) = \mathbb{1}\{\|\gamma - \rho\|_\infty > c\log(1/\varepsilon)^{-\delta'}\}.$$

In view of Theorem 2, the associated minimax risk in the continuous model,

$$\widetilde{R}_{\mathrm{minimax}} = \inf_\rho \sup_\gamma P_\varepsilon^\gamma\left(\|\gamma - \rho\|_\infty > c\log(1/\varepsilon)^{-\delta'}\right)$$

is greater than $1/4$, at least for ε small enough. Then (62) implies that the minimax risk $\widetilde{R}'_{\mathrm{minimax}}$ in model (7) satisfies

$$\widetilde{R}'_{\mathrm{minimax}} > 1/4 - \delta\left(\mathcal{E}_2^{(r)}, \mathcal{E}_0\right).$$

The lower bound (68) will follow from showing that $\delta(\mathcal{E}_2^{(0)}, \mathcal{E}_0) = 0$. By Lemma 23b., it suffices to show that the data in model (7) has law matching that of some subset of the data observed in model (12) with $r = 0$. Such a subset, recalling the concrete interpretation (11) of model (12), is given by $(Y(T_{pq}))_{p,q \le P}$, where

$$T_{pq} = \psi_p \otimes \psi_q = \langle \cdot, \psi_p \rangle_{L^2(\partial D)} \psi_q.$$

To see this, observe that

$$\langle T_{pq}(\phi_j^{(0)}), \phi_k^{(0)} \rangle_{L^2(\partial D)} = \langle \phi_j^{(0)}, \psi_p \rangle_{L^2(\partial D)} \langle \psi_q, \phi_k^{(0)} \rangle_{L^2(\partial D)},$$

hence

$$\langle \widetilde{\Lambda}_\gamma, T_{pq} \rangle_{\mathbb{H}_0} = \sum_{j,k} \langle \widetilde{\Lambda}_\gamma \phi_j^{(0)}, \phi_k^{(0)} \rangle_{L^2(\partial D)} \langle \phi_j^{(0)}, \psi_p \rangle_{L^2(\partial D)} \langle \psi_q, \phi_k^{(0)} \rangle_{L^2(\partial D)}$$

$$= \langle \widetilde{\Lambda}_\gamma \psi_p, \psi_q \rangle_{L^2(\partial D)},$$

$$\mathbb{W}(T_{pq}) = \sum_{j,k} g_{jk} \langle \phi_j^{(0)}, \psi_p \rangle_{L^2(\partial D)} \langle \psi_q, \phi_k^{(0)} \rangle_{L^2(\partial D)}.$$

The noise variables $g'_{p,q} := \langle \mathbb{W}, T_{pq} \rangle_{\mathbb{H}_0}$ are jointly normally distributed with mean zero and covariances

$$
\begin{aligned}
&\mathrm{Cov}\left(g'_{p,q}, g'_{l,m}\right) \\
&\quad = \sum_{j,k} \langle \phi_j^{(0)}, \psi_p \rangle_{L^2(\partial D)} \langle \phi_j^{(0)}, \psi_l \rangle_{L^2(\partial D)} \langle \psi_q, \phi_k^{(0)} \rangle_{L^2(\partial D)} \langle \psi_m, \phi_k^{(0)} \rangle_{L^2(\partial D)} \\
&\quad = \langle \psi_p, \psi_l \rangle_{L^2(\partial D)} \langle \psi_q, \psi_m \rangle_{L^2(\partial D)} = \delta_{pl} \delta_{qm},
\end{aligned}
$$

so that indeed $(Y(T_{pq}))_{p,q \leq P} \overset{d}{=} (Y_{p,q})_{p,q \leq P}$ as claimed. $\qquad\square$

E. Notation index

$D \subseteq \mathbb{R}^d$, $d \geq 2$. A bounded domain, which is taken to mean a connected open set with smooth boundary ∂D.

$\Gamma_{m,D'} = \{\gamma \in C(D) : \inf_{x \in D} \gamma(x) \geq m, \ \gamma = 1 \text{ on } D \setminus D'\}$, some $m \in (0,1)$ and some domain D' compactly contained in D (i.e. the closure $\overline{D}'$ is a subset of D). $C(D)$ denotes the real-valued bounded continuous functions from D to $\mathbb{R}$. $C_u(D)$ denotes the real-valued uniformly continuous functions on D.

$\Gamma^\alpha_{m,D'}(M) = \{\gamma \in \Gamma_{m,D'} : \|\gamma\|_{H^\alpha(D)} \leq M\}$.

$\gamma \in \Gamma_{m,D'}$. A conductivity function, γ_0 its 'true' value for some statistical theorems.

$\theta_0 = \Phi^{-1} \circ \gamma_0$ for Φ described in Section 3.1. $\theta \in C_u(D)$ a function of the form $\Phi^{-1} \circ \gamma$ for $\gamma \in \Gamma_{m,D'}$ (usually denoting a generic draw from the prior Π of (15)).

m_0, D_0. A lower bound and support set for the 'true' γ_0; m_1, D_1 a lower bound and support set for any draw $\gamma = \Phi \circ \theta$ from the prior Π of Section 3.1.

$\|\cdot\|_\infty$. The usual supremum norm on $C(D)$ or $C(\mathbb{R})$.

$\overline{v}$. The usual complex conjugate of a number or function ($\overline{v}(x) := \overline{v(x)}$).

$u_{\gamma, f}$. The (weak) solution to the Dirichlet problem (1) ($\nabla \cdot (\gamma \nabla u) = 0$ in D, $u = f$ on ∂D).

H^s. An L^2-Sobolev space of complex-valued functions (carefully defined in Appendix A); $H_0^1(D)$ the traceless subset of $H^1(D)$. Inner products here are linear in the first argument and conjugate-linear in the second.

$\mathcal{H}_s = (H^{\min\{1,s+3/2\}}(D) \cap H^1_{\mathrm{loc}}(D))/\mathbb{C}$. H^1_{loc} is defined in Appendix A, and $/\mathbb{C}$ means we identify functions f, $f + c$ which are equal up to a scalar $c \in \mathbb{C}$.

$H^s_\diamond(\partial D) = \{g \in H^s(\partial D) : \langle g, 1 \rangle_{L^2(\partial D)} = 0\}$, $L^2_\diamond(\partial D) = H^0_\diamond(\partial D)$.

$(\phi_k^{(r)})_{k \in \mathbb{N} \cup \{0\}}$. An orthonormal basis of $H^r(\partial D)$ consisting of real-valued eigenfunctions of the Laplace–Beltrami operator on ∂D, with corresponding eigenvalues (sorted so they increase with k) $\lambda_k \geq 0$. More details in Appendix A.

$\psi_p = c_p \mathbb{1}_{I_p}$. Indicator functions of some disjoint measurable subsets $(I_p)_{p \leq P}$ of ∂D, scaled to be orthonormal.

$\frac{\partial}{\partial \nu}$. The outward normal derivative at the boundary of a domain (i.e. usually on ∂D) defined in a trace sense. tr denotes the usual trace map, taking $V \colon D \to \mathbb{C}$ to its 'boundary values' $\mathrm{tr}\, V = V|_{\partial D} \colon \partial D \to \mathbb{C}$.

$\Lambda_\gamma \colon H^{s+1}(\partial D)/\mathbb{C} \to H^s_\diamond(\partial D)$. The Dirichlet-to-Neumann map, taking f to $\gamma \frac{\partial u_{\gamma, f}}{\partial \nu}|_{\partial D}$.

$\widetilde{\Lambda}_\gamma = \Lambda_\gamma - \Lambda_1$.

$\|\cdot\|_* = \|\cdot\|_{H^{1/2}(\partial D)/\mathbb{C} \to H^{-1/2}(\partial D)}$, where $\|\cdot\|_{A \to B}$ denotes the operator norm between Banach spaces A and B.

$\mathcal{L}(A, B) = \{T : A \to B \text{ linear s.t. } \|T\|_{A \to B} < \infty\}$.

$\mathcal{L}_2(A, B) = \{T \in \mathcal{L}(A, B) : \|T\|^2_{\mathcal{L}_2(A,B)} := \sum_k \|Te_k^{(A)}\|^2_B < \infty\}$. The space of Hilbert–Schmidt operators from A to B for separable Hilbert spaces A and B, with $(e_k^{(A)})_{k \in \mathbb{N}}$ an orthonormal basis of A.

$b_{jk}^{(r)}(f) \equiv (\phi_j^{(r)}) \otimes \phi_k^{(0)}(f) = \langle f, \phi_j^{(r)} \rangle_{H^r(\partial D)} \phi_k^{(0)}$, $j, k \in \mathbb{N}$. These form an orthonormal basis of $\mathbb{H}_r$ (considering real linear combinations) and of $\mathcal{L}_2(H^r(\partial D)/\mathbb{C}, L^2_\diamond(\partial D))$ (considering complex linear combinations).

$\mathbb{H}_r = \{T \colon H^r(\partial D) \to L^2(\partial D), \ T = \sum_{j,k=1}^\infty t_{jk} b_{jk}^{(r)} : t_{jk} \in \mathbb{R}, \sum_{j,k=1}^\infty t_{jk}^2 < \infty\}$.

A (real) Hilbert space for the inner product

$$\langle S, T \rangle_{\mathbb{H}_r} = \sum_{j,k=1}^\infty s_{jk} t_{jk} = \sum_{j,k=1}^\infty \langle S\phi_j^{(r)}, \phi_k^{(0)} \rangle_{L^2(\partial D)} \langle T\phi_j^{(r)}, \phi_k^{(0)} \rangle_{L^2(\partial D)}.$$

We can also view $\mathbb{H}_r$ as the subset (closed under vector addition and *real* scalar multiplication) of $\mathcal{L}_2(H^r(\partial D)/\mathbb{C}, L^2_\diamond(\partial D))$ comprising those operators which map real-valued functions to real-valued functions.

$Y = \widetilde{\Lambda}_\gamma + \varepsilon \mathbb{W}$. The observed data in model (12), where $\mathbb{W}$ is a Gaussian white noise indexed by the Hilbert space $\mathbb{H}_r$, and ε is a noise level which tends to zero for our asymptotic results. Concrete interpretation of this model given in (11).

$P_\varepsilon^\gamma = P_{\varepsilon, r}^\gamma$. The law of Y. E_ε^γ the corresponding expectation operator, Var_γ the corresponding variance operator.

$p_\varepsilon^\gamma(Y) = \exp(\frac{1}{\varepsilon^2} \langle Y, \widetilde{\Lambda}_\gamma \rangle_{\mathbb{H}_r} - \frac{1}{2\varepsilon^2} \|\widetilde{\Lambda}_\gamma\|^2_{\mathbb{H}_r})$. The probability density of the law of Y w.r.t. the law of $\varepsilon \mathbb{W}$.

$\ell(\gamma) = \log p_\varepsilon^\gamma$. The log-likelihood function.

$\xi_{\varepsilon, \delta} = (\log(\varepsilon^{-1}))^{-\delta}$ for $\varepsilon, \delta > 0$.

Π. A prior for $\theta \in C_u(D)$ described in Section 3.1. Π also denotes the induced prior on $\gamma = \Phi \circ \theta \in \Gamma_{m_1, D_1}$ and the induced prior on Λ_γ.

$\Pi(\,\cdot\, | \, Y)$. The corresponding posterior. $E^\Pi[\,\cdot\, | \, Y]$ the posterior expectation.

$(\mathcal{H}, \|\cdot\|_{\mathcal{H}}) \subseteq (H^\alpha(D), \|\cdot\|_{H^\alpha(D)})$. The RKHS of a base prior Π' from which Π is constructed.

Φ. A 'regular link function' used in the prior construction (see Section 3.1).

$\zeta \colon D \to [0, 1]$. A smooth cutoff function used in the prior construction.

π_{JK}. The $\mathbb{H}_r$-orthogonal projection map onto $\mathrm{span}\{b_{jk}^{(r)} : j \leq J, k \leq K\}$ (see (24)).

$K(P, Q) = K(p, q) = E_{X \sim p} \log((p/q)(X))$, for distributions P, Q with densities p, q (the *Kullback–Leibler*, or just KL, divergence)

$B_{KL}^\varepsilon(\eta) = \{\theta \in C_u(D) : K(p_\varepsilon^{\theta_0}, p_\varepsilon^{\Phi \circ \theta}) \leq (\eta/\varepsilon)^2, \mathrm{Var}_{\gamma_0}(\log(p_\varepsilon^{\theta_0}/p_\varepsilon^{\Phi \circ \theta})) \leq (\eta/\varepsilon)^2\}$.

$N(S, \rho, \delta)$. The covering numbers of the set S for metric ρ, i.e. the smallest number of ρ-balls of radius δ needed to cover S.

References

[1] R. A. Adams and J. J. F. Fournier, *Sobolev spaces*. Second edition, Pure and Applied Mathematics (Amsterdam), 140, Elsevier/Academic Press, Amsterdam, 2003. Zbl 1098.46001 MR 2424078

[2] G. Alessandrini, Stable determination of conductivity by boundary measurements, *Appl. Anal.*, **27** (1988), no. 1-3, 153–172. Zbl 0616.35082 MR 922775

[3] G. Alessandrini and S. Vessella, Lipschitz stability for the inverse conductivity problem, *Adv. in Appl. Math.*, **35** (2005), no. 2, 207–241. Zbl 1095.35058 MR 2152888

[4] J. Aubin, *Applied functional analysis*. Translated from the French by Carole Labrousse. With exercises by Bernard Cornet and Jean-Michel Lasry, John Wiley & Sons, New York-Chichester-Brisbane, 1979. Zbl 0424.46001 MR 549483

[5] A.-P. Calderón, On an inverse boundary value problem, in *Seminar on Numerical Analysis and its Applications to Continuum Physics (Rio de Janeiro, 1980)*, 65–73, Soc. Brasil. Mat., Rio de Janeiro, 1980. Zbl 1182.35230 MR 590275

[6] P. Caro and A. Garcia, The Calderón problem with corrupted data, *Inverse Problems*, **33** (2017), no. 8, 085001, 17pp. Zbl 06774350 MR 3663121

[7] I. Chavel, *Eigenvalues in Riemannian geometry*. Including a chapter by Burton Randol. With an appendix by Jozef Dodziuk, Pure and Applied Mathematics, 115, Academic Press, Inc., Orlando, FL, 1984. Zbl 0551.53001 MR 768584

[8] S. L. Cotter, G. O. Roberts, A. M. Stuart, and D. White, MCMC methods for functions: modifying old algorithms to make them faster, *Statist. Sci.*, **28** (2013), no. 3, 424–446. Zbl 1331.62132 MR 3135540

[9] P. Diaconis, Bayesian numerical analysis, in *Statistical decision theory and related topics. IV. Vol. 1 (West Lafayette, Ind., 1986)*, 163–175, Springer, New York, 1988. Zbl 0671.65117 MR 927099

[10] M. M. Dunlop and A. M. Stuart, The Bayesian formulation of EIT: analysis and algorithms, *Inverse Probl. Imaging*, **10** (2016), no. 4, 1007–1036. Zbl 1348.62104 MR 3610749

[11] L. C. Evans, *Partial differential equations*, Graduate Studies in Mathematics, 19, American Mathematical Society, Providence, RI, 1998. Zbl 0902.35002 MR 1625845

[12] M. Gehre and B. Jin, Expectation propagation for nonlinear inverse problems – with an application to electrical impedance tomography, *J. Comput. Phys.*, **259** (2014), 513–535. Zbl 1349.78046

[13] D. Geller and I. Z. Pesenson, Band-limited localized Parseval frames and Besov spaces on compact homogeneous manifolds, *J. Geom. Anal.*, **21** (2011), no. 2, 334–371. Zbl 1216.43004 MR 2772076

[14] S. Ghosal and A. van der Vaart, *Fundamentals of nonparametric Bayesian inference*, Cambridge Series in Statistical and Probabilistic Mathematics, 44, Cambridge University Press, Cambridge, 2017. Zbl 1376.62004 MR 3587782

[15] E. Giné and R. Nickl, Rates of contraction for posterior distributions in L^r-metrics, $1 \leq r \leq \infty$, *Ann. Statist.*, **39** (2011), no. 6, 2883–2911. Zbl 1246.62095 MR 3012395

[16] E. Giné and R. Nickl, *Mathematical foundations of infinite-dimensional statistical models*, Cambridge Series in Statistical and Probabilistic Mathematics, 40, Cambridge University Press, New York, 2016. Zbl 1358.62014 MR 3588285

[17] M. Giordano and R. Nickl, Consistency of Bayesian inference with Gaussian process priors in an elliptic inverse problem, to appear in *Inverse Problems*. arXiv:1910.07343

[18] M. Hairer, A. M. Stuart, and S. J. Vollmer, Spectral gaps for a Metropolis–Hastings algorithm in infinite dimensions, *Ann. Appl. Probab.*, **24** (2014), no. 6, 2455–2490. Zbl 1307.65002 MR 3262508

[19] M. Hanke, N. Hyvönen, and S. Reusswig, Convex backscattering support in electric impedance tomography, *Numer. Math.*, **117** (2011), no. 2, 373–396. Zbl 1215.78011 MR 2754855

[20] M. Hinze, B. Kaltenbacher, and T. N. T. Quyen, Identifying conductivity in electrical impedance tomography with total variation regularization, *Numer. Math.*, **138** (2018), no. 3, 723–765. Zbl 1397.65238 MR 3767699

[21] B. Jin and P. Maass, An analysis of electrical impedance tomography with applications to Tikhonov regularization, *ESAIM Control Optim. Calc. Var.*, **18** (2012), no. 4, 1027–1048. Zbl 1259.49056 MR 3019471

[22] J. P. Kaipio, V. Kolehmainen, E. Somersalo, and M. Vauhkonen, Statistical inversion and Monte Carlo sampling methods in electrical impedance tomography, *Inverse Problems*, **16** (2000), no. 5, 1487–1522. Zbl 1044.78513 MR 1800606

[23] R. V. Kohn and M. Vogelius, Relaxation of a variational method for impedance computed tomography, *Comm. Pure Appl. Math.*, **40** (1987), no. 6, 745–777. Zbl 0659.49009 MR 910952

[24] V. Kolehmainen, M. Lassas, P. Ola, and S. Siltanen, Recovering boundary shape and conductivity in electrical impedance tomography, *Inverse Probl. Imaging*, **7** (2013), no. 1, 217–242. Zbl 1267.35254 MR 3031845

[25] L. Le Cam, *Asymptotic methods in statistical decision theory*, Springer Series in Statistics, Springer-Verlag, New York, 1986. Zbl 0605.62002 MR 856411

[26] W. V. Li and W. Linde, Approximation, metric entropy and small ball estimates for Gaussian measures, *Ann. Probab.*, **27** (1999), no. 3, 1556–1578. Zbl 0983.60026 MR 1733160

[27] J. L. Lions and E. Magenes, *Non-homogeneous boundary value problems and applications. Vol. II*. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 182, Springer-Verlag, New York-Heidelberg, 1972. Zbl 0227.35001 MR 350178

[28] N. Mandache, Exponential instability in an inverse problem for the Schrödinger equation, *Inverse Problems*, **17** (2001), no. 5, 1435–1444. Zbl 0985.35110 MR 1862200

[29] E. Mariucci, Le Cam theory on the comparison of statistical models, *Grad. J. Math.*, **1** (2016), no. 2, 81–91. MR 3850766

[30] F. Monard, R. Nickl, and G. P. Paternain, Consistent inversion of noisy non-Abelian X-ray transforms, to appear in *Comm. Pure Appl. Math.* arXiv:1905.00860

[31] A. I. Nachman, Reconstructions from boundary measurements, *Ann. of Math. (2)*, **128** (1988), no. 3, 531–576. Zbl 0675.35084 MR 970610

[32] A. I. Nachman, Global uniqueness for a two-dimensional inverse boundary value problem, *Ann. of Math. (2)*, **143** (1996), no. 1, 71–96. Zbl 0857.35135 MR 1370758

[33] R. Nickl, Bernstein–von Mises theorems for statistical inverse problems I: Schrödinger equation, *J. Eur. Math. Soc. (JEMS)*, **22** (2020), no. 8, 2697–2750. MR 4118619

[34] R. Nickl and J. Söhl, Nonparametric Bayesian posterior contraction rates for discretely observed scalar diffusions, *Ann. Statist.*, **45** (2017), no. 4, 1664–1693. Zbl 1411.62087 MR 3670192

[35] R. Nickl and J. Söhl, Bernstein–von Mises theorems for statistical inverse problems II: compound Poisson processes, *Electron. J. Stat.*, **13** (2019), no. 2, 3513–3571. Zbl 1429.62168 MR 4013745

[36] R. Nickl, S. van de Geer, and S. Wang, Convergence Rates for Penalized Least Squares Estimators in PDE Constrained Regression Problems, *SIAM/ASA J. Uncertain. Quantif.*, **8** (2020), no. 1, 374–413. Zbl 07193416 MR 4074017

[37] R. Novikov and M. Santacesaria, A global stability estimate for the Gel'fand-Calderón inverse problem in two dimensions, *J. Inverse Ill-Posed Probl.*, **18** (2010), no. 7, 765–785. Zbl 1279.35120 MR 2772570

[38] R. G. Novikov, A multidimensional inverse spectral problem for the equation $-\Delta\psi + (v(x) - Eu(x))\psi = 0$, *Funktsional. Anal. i Prilozhen.*, **22** (1988), no. 4, 11–22, 96; translation in *Funct. Anal. Appl.*, **22** (1988), no. 4, 263–272 (1989). Zbl 0697.35108 MR 976992

[39] R. G. Novikov, New global stability estimates for the Gel'fand–Calderon inverse problem, *Inverse Problems*, **27** (2011), no. 1, 015001, 21pp. Zbl 1208.35173 MR 2746404

[40] H. Poincaré, *Calcul des probabilités*. Rédaction de A. Quiquet. 2e édition, Gauthier-Villars, Paris, 1912. Zbl 43.0308.04

[41] M. Reiß, Asymptotic equivalence for nonparametric regression with multivariate and random design, *Ann. Statist.*, **36** (2008), no. 4, 1957–1982. Zbl 1142.62023 MR 2435461

[42] L. Roininen, J. M. J. Huttunen, and S. Lasanen, Whittle–Matérn priors for Bayesian statistical inversion with applications in electrical impedance tomography, *Inverse Probl. Imaging*, **8** (2014), no. 2, 561–586. Zbl 1302.65245 MR 3209311

[43] M. Salo, Calderón problem, lecture notes, 2008. Available at: http://users.jyu.fi/~salomi/lecturenotes/calderon_lectures.pdf

[44] A. M. Stuart, Inverse problems: a Bayesian perspective, *Acta Numer.*, **19** (2010), 451–559. Zbl 1242.65142 MR 2652785

[45] J. Sylvester and G. Uhlmann, A global uniqueness theorem for an inverse boundary value problem, *Ann. of Math. (2)*, **125** (1987), no. 1, 153–169. Zbl 0625.35078 MR 873380

[46] A. B. Tsybakov, *Introduction to nonparametric estimation*. Revised and extended from the 2004 French original. Translated by Vladimir Zaiats, Springer Series in Statistics, Springer, New York, 2009. Zbl 1176.62032 MR 2724359

[47] G. Uhlmann, Electrical impedance tomography and Calderón's problem, *Inverse Problems*, **25** (2009), no. 12, 123011, 39pp. Zbl 1181.35339 MR 3460047

[48] A. W. van der Vaart and J. H. van Zanten, Rates of contraction of posterior distributions based on Gaussian process priors, *Ann. Statist.*, **36** (2008), no. 3, 1435–1463. Zbl 1141.60018 MR 2418663

[49] S. J. Vollmer, Posterior consistency for Bayesian inverse problems through stability and regression results, *Inverse Problems*, **29** (2013), no. 12, 125011, 32pp. Zbl 1284.35064 MR 3141858

K. Abraham, Université Paris-Saclay, CNRS,
Laboratoire de mathématiques d'Orsay, 91405 Orsay, France
E-mail: kweku.abraham@universite-paris-saclay.fr

R. Nickl, Department of Pure Mathematics and Mathematical Statistics,
University of Cambridge, Wilberforce Road, Cambridge CB3 0WB, UK
E-mail: r.nickl@statslab.cam.ac.uk