The Annals of Statistics
2022, Vol. 50, No. 2, 1038–1065
https://doi.org/10.1214/21-AOS2138

NECESSARY AND SUFFICIENT CONDITIONS FOR ASYMPTOTICALLY OPTIMAL LINEAR PREDICTION OF RANDOM FIELDS ON COMPACT METRIC SPACES

BY KRISTIN KIRCHNER[1,a] AND DAVID BOLIN[2,b]

[1]*Delft Institute of Applied Mathematics, Delft University of Technology,* [a]*k.kirchner@tudelft.nl*
[2]*CEMSE Division, King Abdullah University of Science and Technology,* [b]*david.bolin@kaust.edu.sa*

Optimal linear prediction (aka. kriging) of a random field $\{Z(x)\}_{x\in\mathcal{X}}$ indexed by a compact metric space $(\mathcal{X}, d_{\mathcal{X}})$ can be obtained if the mean value function $m\colon \mathcal{X} \to \mathbb{R}$ and the covariance function $\varrho\colon \mathcal{X}\times\mathcal{X} \to \mathbb{R}$ of Z are known. We consider the problem of predicting the value of $Z(x^*)$ at some location $x^* \in \mathcal{X}$ based on observations at locations $\{x_j\}_{j=1}^n$, which accumulate at x^* as $n \to \infty$ (or, more generally, predicting $\varphi(Z)$ based on $\{\varphi_j(Z)\}_{j=1}^n$ for linear functionals $\varphi, \varphi_1, \dots, \varphi_n$). Our main result characterizes the asymptotic performance of linear predictors (as n increases) based on an incorrect second-order structure $(\widetilde{m}, \widetilde{\varrho})$, without any restrictive assumptions on $\varrho, \widetilde{\varrho}$ such as stationarity. We, for the first time, provide necessary and sufficient conditions on $(\widetilde{m}, \widetilde{\varrho})$ for asymptotic optimality of the corresponding linear predictor holding uniformly with respect to φ. These general results are illustrated by weakly stationary random fields on $\mathcal{X} \subset \mathbb{R}^d$ with Matérn or periodic covariance functions, and on the sphere $\mathcal{X} = \mathbb{S}^2$ for the case of two isotropic covariance functions.

1. Introduction. Optimal linear prediction of random fields, often also called kriging, is an important and widely used technique for interpolation of spatial data. Consider a random field $\{Z(x) : x \in \mathcal{X}\}$ on a compact topological space $\mathcal{X}$ such as a closed and bounded subset of $\mathbb{R}^d$. Assume that Z is almost surely continuous on $\mathcal{X}$ and that we want to predict its value at a location $x^* \in \mathcal{X}$ based on a set of observations $\{Z(x_j)\}_{j=1}^n$ for locations $x_1, \dots, x_n \in \mathcal{X}$ all distinct from x^*. The kriging predictor is the linear predictor $\widehat{Z}(x^*) = \alpha_0 + \sum_{j=1}^n \alpha_j Z(x_j)$ of $Z(x^*)$ based on the observations, where the coefficients $\alpha_0, \dots, \alpha_n \in \mathbb{R}$ are chosen such that the variance of the error $(\widehat{Z} - Z)(x^*)$ is minimized. By letting $m(\,\cdot\,)$ and $\varrho(\,\cdot\,,\,\cdot\,)$ denote the mean and the covariance function of Z, we can express $\widehat{Z}(x^*)$ as

$$\widehat{Z}(x^*) = m(x^*) + \mathbf{c}_n^\top \boldsymbol{\Sigma}_n^{-1}(\mathbf{Z}_n - \mathbf{m}_n), \tag{1.1}$$

where $\mathbf{Z}_n := (Z(x_1), \dots, Z(x_n))^\top$, $\mathbf{m}_n := (m(x_1), \dots, m(x_n))^\top$, $\boldsymbol{\Sigma}_n \in \mathbb{R}^{n\times n}$ has elements $[\boldsymbol{\Sigma}_n]_{ij} := \varrho(x_i, x_j)$ and $\mathbf{c}_n := (\varrho(x^*, x_1), \dots, \varrho(x^*, x_n))^\top$.

In applications, the mean and covariance functions are rarely known and, therefore, need to be estimated from data. It is thus of interest to study the effect, which a misspecification of the mean or the covariance function has on the efficiency of the linear predictor. Stein [20, 21] considered the situation that the sequence $\{x_j\}_{j\in\mathbb{N}} \subset \mathbb{R}^d$ has x^* as a limiting point and the predictor $\widehat{Z}$ is computed using misspecified mean and covariance functions, $\widetilde{m}$ and $\widetilde{\varrho}$. His main outcome was that the best linear predictor based on $(\widetilde{m}, \widetilde{\varrho})$ is asymptotically efficient, as $n \to \infty$, provided that the Gaussian measures corresponding to (m, ϱ) and $(\widetilde{m}, \widetilde{\varrho})$ are equivalent (see Appendix A). This result in fact holds uniformly with respect to x^* and, moreover, uniformly for each linear functional φ such that $\varphi(Z)$ has finite variance [21].

Received May 2020; revised September 2021.
MSC2020 subject classifications. Primary 62M20; secondary 60G12, 60G25, 60G60.
Key words and phrases. Kriging, approximation in Hilbert spaces, spatial statistics.

For stationary covariance functions, there exist simple conditions for verifying whether the corresponding Gaussian measures are equivalent [2, 4, 22], and thus if the linear predictions are asymptotically efficient. However, for any constant $c \in (0, \infty)$, the linear predictor based on $(m, c\varrho)$ is equal to that based on (m, ϱ), whereas the Gaussian measures corresponding to $(m, c\varrho)$ and (m, ϱ) are orthogonal for all $c \neq 1$. This shows that equivalence of the measures is a sufficient but not necessary condition for asymptotic efficiency. Less restrictive conditions have been derived for some specific cases such as periodic processes on $[0, 1]^d$ and weakly stationary random fields on $\mathbb{R}^d$ observed on a lattice [23, 24]. With these results in mind, an immediate question is if one can find *necessary and sufficient* conditions for uniform asymptotic efficiency of linear prediction using misspecified mean and covariance functions. The aim of this work is to show that this indeed is the case.

We derive necessary and sufficient conditions for general second-order structures (m, ϱ) and $(\widetilde{m}, \widetilde{\varrho})$, without any restrictive assumptions such as periodicity or stationarity. These conditions are weaker compared to those of the Feldman–Hájek theorem, and thus clearly exhibit their fulfillment in the case that the Gaussian measures corresponding to (m, ϱ) and $(\widetilde{m}, \widetilde{\varrho})$ are equivalent (see Remark 3.4). Furthermore, our results are formulated for random fields on general compact metric spaces, which include compact Euclidean domains in $\mathbb{R}^d$, but also more general domains such as the sphere $\mathbb{S}^2$ or metric graphs (see Example 2.1). Assuming compactness of the space is meaningful, since in applications the observed locations are always contained in some compact subset (e.g., a bounded and closed domain in $\mathbb{R}^d$).

This general setting is outlined in Section 2. Our main results are stated in Section 3 and proven in Section 4. Section 5 presents simplified necessary and sufficient conditions for the two important special cases when ϱ, $\widetilde{\varrho}$ induce the same eigenfunctions, or when ϱ, $\widetilde{\varrho}$ are translation invariant on $\mathbb{R}^d$ and have spectral densities. Section 6 verifies these conditions for weakly stationary random fields on $\mathcal{X} \subset \mathbb{R}^d$, where ϱ, $\widetilde{\varrho}$ are of Matérn type or periodic on $[0, 1]^d$. We also discuss an example on $\mathcal{X} = \mathbb{S}^2$ which, to the best of our knowledge, is the first result on asymptotically optimal linear prediction on the sphere. The Supplementary Material [9] contains three appendices (Appendix A, B and C) pertaining to this article.

2. Setting and problem formulation. We assume that we are given a square-integrable stochastic process $Z\colon \mathcal{X} \times \Omega \to \mathbb{R}$ defined on a complete probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and indexed by a connected, compact metric space $(\mathcal{X}, d_{\mathcal{X}})$ of infinite cardinality. In addition, we let $m\colon \mathcal{X} \to \mathbb{R}$ denote the mean value function of Z and assume that the covariance function,

$$\varrho\colon \mathcal{X} \times \mathcal{X} \to \mathbb{R}, \qquad \varrho(x, x') := \int_{\Omega} \big(Z(x, \omega) - m(x)\big)\big(Z(x', \omega) - m(x')\big)\, \mathrm{d}\mathbb{P}(\omega),$$

is (strictly) positive definite and continuous. Let $\nu_{\mathcal{X}}$ be a strictly positive and finite Borel measure on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$. Here and throughout, $\mathcal{B}(\mathcal{T})$ denotes the Borel σ-algebra on a topological space $\mathcal{T}$. As the symmetric covariance function ϱ is assumed to be positive definite and continuous, the corresponding covariance operator, defined by

$$\mathcal{C}\colon L_2(\mathcal{X}, \nu_{\mathcal{X}}) \to L_2(\mathcal{X}, \nu_{\mathcal{X}}), \qquad (\mathcal{C}w)(x) := \int_{\mathcal{X}} \varrho(x, x') w(x')\, \mathrm{d}\nu_{\mathcal{X}}(x'), \tag{2.1}$$

is self-adjoint, positive definite and compact on $L_2(\mathcal{X}, \nu_{\mathcal{X}})$. Since $(\mathcal{X}, d_{\mathcal{X}})$ is connected and compact, the set $\mathcal{X}$ is uncountable and $L_2(\mathcal{X}, \nu_{\mathcal{X}})$ is an infinite-dimensional separable Hilbert space. By compactness of $\mathcal{C}$, there exists a countable system of (equivalence classes of) eigenfunctions $\{e_j\}_{j\in\mathbb{N}}$ of $\mathcal{C}$, which can be chosen as an orthonormal basis for $L_2(\mathcal{X}, \nu_{\mathcal{X}})$.

Moreover, it can be shown that $\mathcal{C}$ maps into the space of continuous functions. For this reason, we may identify the eigenfunctions $\{e_j\}_{j\in\mathbb{N}}$ with their continuous representatives. We let $\{\gamma_j\}_{j\in\mathbb{N}}$ denote the positive eigenvalues corresponding to $\{e_j\}_{j\in\mathbb{N}}$. By Mercer's theorem (see, e.g., [15, 26]), the covariance function ϱ then admits the series representation

$$\varrho(x,x') = \sum_{j\in\mathbb{N}} \gamma_j e_j(x) e_j(x'), \quad x,x' \in \mathcal{X}, \tag{2.2}$$

where the convergence of this series is absolute and uniform. In addition, we can express the action of the covariance operator by a series,

$$\mathcal{C}w = \sum_{j\in\mathbb{N}} \gamma_j (w, e_j)_{L_2(\mathcal{X},\nu_{\mathcal{X}})} e_j, \quad w \in L_2(\mathcal{X}, \nu_{\mathcal{X}}),$$

converging pointwise—that is, for all $w \in L_2(\mathcal{X}, \nu_{\mathcal{X}})$—and uniformly, that is, in the operator norm. Finally, we note that square-integrability of the stochastic process implies that $\mathcal{C}$ has a finite trace on $L_2(\mathcal{X}, \nu_{\mathcal{X}})$, that is, $\operatorname{tr}(\mathcal{C}) = \sum_{j\in\mathbb{N}} \gamma_j < \infty$.

EXAMPLE 2.1. Examples of covariance functions on a compact metric space $(\mathcal{X}, d_{\mathcal{X}})$ are given by the Matérn class, where

$$\varrho(x,x') := \varrho_{\mathsf{M}}(d_{\mathcal{X}}(x,x')) \quad \text{with} \quad \varrho_{\mathsf{M}}(r) := \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)} (\kappa r)^{\nu} K_{\nu}(\kappa r), \quad r \geq 0.$$

More precisely, one may consider stochastic processes with Matérn covariance functions, indexed by one of the following compact metric spaces:

(a) $\mathcal{X} \subset \mathbb{R}^d$ is a *connected, compact Euclidean domain* equipped with the Euclidean metric for all parameters $\sigma, \nu, \kappa \in (0,\infty)$; see [14],

(b) $\mathcal{X} := \mathbb{S}^d = \{x \in \mathbb{R}^{d+1} : \|x\|_{\mathbb{R}^{d+1}} = 1\}$ is the *d-sphere* equipped with the great circle distance $d_{\mathbb{S}^d}(x,x') := \arccos((x,x')_{\mathbb{R}^{d+1}})$ for $\sigma, \kappa \in (0,\infty)$ and $\nu \in (0, 1/2]$; see [6], Section 4.5, Example 2,

(c) $\mathcal{X} \subset \mathbb{R}^D$ is a *d-dimensional connected, compact manifold* (e.g., the d-sphere $\mathbb{S}^d$), embedded in $\mathbb{R}^D$ for some $D > d$ and equipped with the Euclidean metric on $\mathbb{R}^D$ for any set of parameters $\sigma, \nu, \kappa \in (0,\infty)$; see, for example [7],

(d) $\mathcal{X}$ is a *graph with Euclidean edges* equipped with the resistance metric for the parameters $\sigma, \kappa \in (0,\infty)$ and $\nu \in (0, 1/2]$; see [1], Definition 1, Section 2.3 and Table 1.

We point out that, for $\nu \in (1/2, \infty)$, the function $(x,x') \mapsto \varrho_{\mathsf{M}}(d_{\mathcal{X}}(x,x'))$ in (b) and (d) is not (strictly) positive definite, and thus, not a valid covariance function for our setting. We furthermore emphasize that the Matérn covariance families in (a) and (b) are stationary on $\mathbb{R}^d$ and isotropic on $\mathbb{S}^d$, respectively, but we do not require ϱ to have these properties.

Since the kriging predictor in (1.1) only depends on the mean value function and the covariance function of the process Z, it is identical to the kriging predictor for a *Gaussian* process with the same first two moments. For ease of presentation, we therefore from now on assume that Z is a Gaussian process on $(\mathcal{X}, d_{\mathcal{X}})$ with mean value function $m \in L_2(\mathcal{X}, \nu_{\mathcal{X}})$, continuous, (strictly) positive definite covariance function ϱ and corresponding covariance operator $\mathcal{C}$. Note, however, that all our results extend to the case of non-Gaussian processes, as their proofs rely only on the first two statistical moments. We write $\mu = \mathsf{N}(m, \mathcal{C})$ for the Gaussian measure on the Hilbert space $L_2(\mathcal{X}, \nu_{\mathcal{X}})$ induced by the process Z, that is, for every Borel set $A \in \mathcal{B}(L_2(\mathcal{X}, \nu_{\mathcal{X}}))$ we have

$$\mu(A) = \mathbb{P}(\{\omega \in \Omega : Z(\cdot, \omega) \in A\}).$$

The operator $\mathsf{E}[\cdot]$ will denote the expectation operator under μ, that is, for an $L_2(\mathcal{X}, \nu_{\mathcal{X}})$-valued random variable Y with distribution μ and a real-valued, Borel measurable mapping

$\varphi\colon L_2(\mathcal{X}, \nu_{\mathcal{X}}) \to \mathbb{R}$, the expected values $\mathsf{E}[Y] \in L_2(\mathcal{X}, \nu_{\mathcal{X}})$ and $\mathsf{E}[\varphi(Y)] \in \mathbb{R}$ (when existent) are the Bochner integrals

$$\mathsf{E}[Y] = \int_{L_2(\mathcal{X},\nu_{\mathcal{X}})} y \,\mathrm{d}\mu(y) \quad \text{and} \quad \mathsf{E}\bigl[\varphi(Y)\bigr] = \int_{L_2(\mathcal{X},\nu_{\mathcal{X}})} \varphi(y) \,\mathrm{d}\mu(y);$$

cf. [11], Corollary 5.1. Furthermore, we use the following notation for the real-valued variance and covariance operators with respect to μ: If $\varphi, \varphi'\colon L_2(\mathcal{X}, \nu_{\mathcal{X}}) \to \mathbb{R}$ are Borel measurable and $g := \varphi(Y)$, $g' := \varphi'(Y)$, then

$$\mathsf{Var}[g] := \mathsf{E}\bigl[\bigl(g - \mathsf{E}[g]\bigr)^2\bigr], \qquad \mathsf{Cov}\bigl[g, g'\bigr] := \mathsf{E}\bigl[\bigl(g - \mathsf{E}[g]\bigr)\bigl(g' - \mathsf{E}\bigl[g'\bigr]\bigr)\bigr].$$

To present the theoretical setting of optimal linear prediction (kriging) as well as the necessary notation, we proceed in two steps: We first consider the centered case with $m = 0$, and then extend it to the general case.

2.1. *Kriging assuming zero mean.* Let $Z^0\colon \mathcal{X} \times \Omega \to \mathbb{R}$ be a *centered* Gaussian process. Under the assumption that it has a continuous covariance function ϱ, we may identify $Z^0\colon \mathcal{X} \to L_2(\Omega, \mathbb{P})$ with its continuous representative. In particular, for each $x \in \mathcal{X}$, the real-valued random variable $Z^0(x)$ is a well-defined element of $L_2(\Omega, \mathbb{P})$. Consider the vector space $\mathcal{Z}^0 \subset L_2(\Omega, \mathbb{P})$ of finite linear combinations of such random variables,

$$\mathcal{Z}^0 := \left\{ \sum_{j=1}^{K} \alpha_j Z^0(x_j) : K \in \mathbb{N}, \alpha_1, \ldots, \alpha_K \in \mathbb{R}, x_1, \ldots, x_K \in \mathcal{X} \right\}. \tag{2.3}$$

We then define the (Gaussian) Hilbert space $\mathcal{H}^0$ (cf. [8]) as the closure of $\mathcal{Z}^0$ with respect to the norm $\|\cdot\|_{\mathcal{H}^0}$ induced by the $L_2(\Omega, \mathbb{P})$ inner product,

$$\begin{aligned} &\left(\sum_{i=1}^{K} \alpha_i Z^0(x_i), \sum_{j=1}^{K'} \alpha'_j Z^0\bigl(x'_j\bigr) \right)_{\mathcal{H}^0} := \sum_{i=1}^{K} \sum_{j=1}^{K'} \alpha_i \alpha'_j \mathsf{E}\bigl[Z^0(x_i) Z^0\bigl(x'_j\bigr)\bigr], \\ &\mathcal{H}^0 := \Bigl\{ g \in L_2(\Omega, \mathbb{P}) \,\big|\, \exists \{g_j\}_{j\in\mathbb{N}} \subset \mathcal{Z}^0 : \lim_{j\to\infty} \|g - g_j\|_{L_2(\Omega,\mathbb{P})} = 0 \Bigr\}. \end{aligned} \tag{2.4}$$

Continuity of the covariance kernel ϱ on $\mathcal{X} \times \mathcal{X}$ implies separability of the Hilbert space $\mathcal{H}^0$; see [3], Theorem 32, and [16], Theorem 2C.

By definition (see, e.g., [25], Section 1.2), the kriging predictor h_n^0 of $h^0 \in \mathcal{H}^0$ based on a set of observations $\{y_{n1}^0, \ldots, y_{nn}^0\} \subset \mathcal{H}^0$ is the best linear predictor in $\mathcal{H}^0$ or, in other words, the $\mathcal{H}^0$-orthogonal projection of h^0 onto the linear space generated by $y_{n1}^0, \ldots, y_{nn}^0$. By recalling the inner product on $\mathcal{H}^0$ from (2.4), the kriging predictor h_n^0 is thus the unique element in the finite-dimensional subspace $\mathcal{H}_n^0 := \operatorname{span}\{y_{n1}^0, \ldots, y_{nn}^0\} \subset \mathcal{H}^0$ satisfying

$$h_n^0 \in \mathcal{H}_n^0 : \quad \bigl(h_n^0 - h^0, g_n^0\bigr)_{\mathcal{H}^0} = \mathsf{E}\bigl[\bigl(h_n^0 - h^0\bigr) g_n^0\bigr] = 0 \quad \forall g_n^0 \in \mathcal{H}_n^0. \tag{2.5}$$

Consequently, h_n^0 is the $\mathcal{H}^0$-best approximation of h^0 in $\mathcal{H}_n^0$, that is,

$$\bigl\|h_n^0 - h^0\bigr\|_{\mathcal{H}^0} = \inf_{g_n^0 \in \mathcal{H}_n^0} \bigl\|g_n^0 - h^0\bigr\|_{\mathcal{H}^0}.$$

2.2. *Kriging with general mean.* Let us next consider the case that the Gaussian process Z has a general mean value function $m\colon \mathcal{X} \to \mathbb{R}$ which, for now, we assume to be continuous. The analytical framework for kriging then needs to be adjusted, since the space of possible predictors has to contain functions of the form (1.1) including constants.

Evidently, every linear combination $h = \sum_{j=1}^{K} \alpha_j Z(x_j)$ has a representation $h = c + h^0$ with $c \in \mathbb{R}$ and $h^0 \in \mathcal{Z}^0 \subset \mathcal{H}^0$, where the vector spaces $\mathcal{Z}^0, \mathcal{H}^0 \subset L_2(\Omega, \mathbb{P})$ are generated by linear combinations of the centered process $Z^0 := Z - m$ as in (2.3) and (2.4); namely,

$$h = \sum_{j=1}^{K} \alpha_j Z(x_j) = \sum_{j=1}^{K} \alpha_j m(x_j) + \sum_{j=1}^{K} \alpha_j Z^0(x_j) =: c + h^0, \tag{2.6}$$

or, more generally, $h = \mathsf{E}[h] + (h - \mathsf{E}[h])$. Furthermore, note that zero is the only constant contained in $\mathcal{Z}^0, \mathcal{H}^0 \subset L_2(\Omega, \mathbb{P})$. This follows from the fact that elements in $\mathcal{Z}^0$ are linear combinations of the process Z^0 at locations in $\mathcal{X}$; see (2.3). A constant $c \neq 0$ in $\mathcal{Z}^0$ would thus imply that the corresponding linear combination $c = \sum_{j=1}^{K} \alpha_j Z^0(x_j)$ has zero variance, which contradicts the (strict) positive definiteness of the covariance function ϱ. For this reason, the decomposition in (2.6) is unique. This motivates to define the Hilbert space containing all possible observations and predictors for a general second-order structure (m, ϱ) as the (internal) direct sum of vector spaces, given by

$$\mathcal{H} := \mathbb{R} \oplus \mathcal{H}^0 = \{h \in L_2(\Omega, \mathbb{P}) : \exists c \in \mathbb{R}, \exists h^0 \in \mathcal{H}^0 \text{ with } h = c + h^0\}, \tag{2.7}$$

which is equipped with the graph norm,

$$\|h\|_{\mathcal{H}}^2 = |c|^2 + \|h^0\|_{\mathcal{H}^0}^2 \quad \text{if} \quad h = c + h^0 \in \mathbb{R} \oplus \mathcal{H}^0 = \mathcal{H}. \tag{2.8}$$

Note that, similarly as for $\mathcal{H}^0$, the inner product on $\mathcal{H}$ equals the inner product on $L_2(\Omega, \mathbb{P})$:

$$(g, h)_{\mathcal{H}} = (\mathsf{E}[g], \mathsf{E}[h])_{\mathbb{R}} + (g - \mathsf{E}[g], h - \mathsf{E}[h])_{\mathcal{H}^0} = \mathsf{E}[g]\mathsf{E}[h] + \mathsf{Cov}[g, h] = \mathsf{E}[gh].$$

Now suppose that we want to predict $h \in \mathcal{H}$ given a set of observations

$$y_{nj} = c_{nj} + y_{nj}^0 \in \mathcal{H}, \quad \text{where} \quad c_{nj} \in \mathbb{R}, \quad y_{nj}^0 \in \mathcal{H}^0, \qquad j \in \{1, \dots, n\}.$$

The kriging predictor of $h = c + h^0 \in \mathbb{R} \oplus \mathcal{H}^0 = \mathcal{H}$ based on the observations $\{y_{n1}, \dots, y_{nn}\}$ is then $h_n = c + h_n^0$, where h_n^0 is the kriging predictor of h^0 based on the centered observations $\{y_{n1}^0, \dots, y_{nn}^0\} \subset \mathcal{H}^0$, as defined in (2.5). The definition of the norm on $\mathcal{H}$ in (2.8) readily implies that

$$\|h_n - h\|_{\mathcal{H}}^2 = |c - c|^2 + \|h_n^0 - h^0\|_{\mathcal{H}^0}^2 = 0 + \inf_{g_n^0 \in \mathcal{H}_n^0} \|g_n^0 - h^0\|_{\mathcal{H}^0}^2.$$

Hence, if we, for $y_{n1}^0, \dots, y_{nn}^0 \in \mathcal{H}^0$, define the subspace $\mathcal{H}_n \subset \mathcal{H}$ by

$$\mathcal{H}_n := \mathbb{R} \oplus \mathcal{H}_n^0, \quad \text{where} \quad \mathcal{H}_n^0 := \operatorname{span}\{y_{n1}^0, \dots, y_{nn}^0\} \subset \mathcal{H}^0, \tag{2.9}$$

we have that in either case (centered and noncentered) the kriging predictor of $h \in \mathcal{H}$ based on the observations $\{y_{n1} = c_{n1} + y_{n1}^0, \dots, y_{nn} = c_{nn} + y_{nn}^0\}$ is given by the $\mathcal{H}$-orthogonal projection of h onto $\mathcal{H}_n$, that is,

$$\begin{aligned} &h_n \in \mathcal{H}_n : \quad (h_n - h, g_n)_{\mathcal{H}} = \mathsf{E}[(h_n - h)g_n] = 0 \quad \forall g_n \in \mathcal{H}_n, \\ &h_n \in \mathcal{H}_n : \qquad \|h_n - h\|_{\mathcal{H}} = \inf_{g_n \in \mathcal{H}_n} \|g_n - h\|_{\mathcal{H}}. \end{aligned} \tag{2.10}$$

For this reason, for every $h \in \mathcal{H}$, the kriging predictor h_n is fully determined by the subspace $\mathcal{H}_n$ and we also call h_n the kriging predictor (or best linear predictor) *based on* $\mathcal{H}_n$ (instead of *based on the set of observations* $\{y_{n1}, \dots, y_{nn}\}$).

Finally, since the definitions (2.7), (2.9) and (2.10) of the spaces $\mathcal{H}$, $\mathcal{H}_n$ and the kriging predictor h_n are meaningful even if the mean value function is not continuous, hereafter we only require that $m \in L_2(\mathcal{X}, \nu_{\mathcal{X}})$. Note, however, that then the point evaluation $Z(x^*)$, $x^* \in \mathcal{X}$, might not be an element of $\mathcal{H}$.

2.3. *Problem formulation.* We assume without loss of generality that the centered observations $y_{n1}^0, \dots, y_{nn}^0$ are linearly independent in $\mathcal{H}^0$ so that in (2.9) we have $\dim(\mathcal{H}_n^0) = n$. Furthermore, we suppose that the family of subspaces $\{\mathcal{H}_n\}_{n\in\mathbb{N}}$ generated by the observations (see (2.9)) is dense in $\mathcal{H}$. More specifically, we require that, for any $h \in \mathcal{H}$, the corresponding kriging predictors $\{h_n\}_{n\in\mathbb{N}}$ defined via (2.10) are consistent in the sense that

$$\lim_{n\to\infty} \mathsf{E}\big[(h_n - h)^2\big] = \lim_{n\to\infty} \|h_n - h\|_{\mathcal{H}}^2 = 0. \tag{2.11}$$

For future reference, we introduce the set $\mathcal{S}_{\text{adm}}^{\mu}$, which contains all admissible sequences $\{\mathcal{H}_n\}_{n\in\mathbb{N}}$ of subspaces of $\mathcal{H}$ generated by observations which provide μ-consistent kriging,

$$\begin{aligned}\mathcal{S}_{\text{adm}}^{\mu} := \big\{\{\mathcal{H}_n\}_{n\in\mathbb{N}} \,\big|\, &\forall n \in \mathbb{N} : \mathcal{H}_n \text{ is as in (2.9) with } \dim(\mathcal{H}_n^0) = n, \\ &\forall h \in \mathcal{H} : \{h_n\}_{n\in\mathbb{N}} \text{ as in (2.10) satisfy (2.11)}\big\}.\end{aligned} \tag{2.12}$$

Since $(\mathcal{X}, d_{\mathcal{X}})$ is connected and compact, we have $\mathcal{S}_{\text{adm}}^{\mu} \neq \varnothing$. Note that we do *not assume nestedness* of $\{\mathcal{H}_n\}_{n\in\mathbb{N}}$. Therefore, we cover situations when the observations are not part of a sequence and $\{y_{n1}, \dots, y_{nn}\} \not\subset \{y_{n+1,1}, \dots, y_{n+1,n+1}\}$.

EXAMPLE 2.2. Suppose that m and ϱ are continuous and that $\{x_j\}_{j\in\mathbb{N}}$ is a sequence in $(\mathcal{X}, d_{\mathcal{X}})$, which *accumulates at* $x^* \in \mathcal{X}$, that is, there exists a subsequence $\{\bar{x}_k\}_{k\in\mathbb{N}} \subseteq \{x_j\}_{j\in\mathbb{N}}$ such that $\lim_{k\to\infty} d_{\mathcal{X}}(\bar{x}_k, x^*) = 0$. Assume further that $\mathcal{H}_n \supseteq \mathbb{R} \oplus \operatorname{span}\{Z^0(x_j) : j \leq n\}$ for all n. Then the kriging predictors $\{h_n\}_{n\in\mathbb{N}}$ for $h := Z(x^*)$ are consistent: For n sufficiently large such that $\{x_j\}_{j=1}^n \cap \{\bar{x}_k\}_{k\in\mathbb{N}} = \{\bar{x}_k\}_{k=1}^{k_n^*}$ is not empty (i.e., $k_n^* \in \mathbb{N}$), we have

$$\begin{aligned}\mathsf{E}\big[(h_n - h)^2\big] &\leq \inf_{\alpha_0,\alpha_1,\dots,\alpha_n\in\mathbb{R}} \mathsf{E}\Bigg[\Bigg(Z(x^*) - \alpha_0 - \sum_{j=1}^n \alpha_j Z^0(x_j)\Bigg)^2\Bigg] \\ &= \inf_{\alpha_1,\dots,\alpha_n\in\mathbb{R}} \mathsf{E}\Bigg[\Bigg(Z^0(x^*) - \sum_{j=1}^n \alpha_j Z^0(x_j)\Bigg)^2\Bigg] \\ &\leq \mathsf{E}\big[\big(Z^0(x^*) - Z^0(\bar{x}_{k_n^*})\big)^2\big] \\ &= \varrho(x^*, x^*) + \varrho(\bar{x}_{k_n^*}, \bar{x}_{k_n^*}) - 2\varrho(\bar{x}_{k_n^*}, x^*) \to 0, \quad \text{as } n \to \infty.\end{aligned}$$

This shows (2.11) for $h = Z(x^*)$. Note that the kriging predictors based on the subspaces $\{\mathcal{H}_n\}_{n\in\mathbb{N}}$ have to be consistent for every $h \in \mathcal{H}$ so that the sequence is admissible, $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\text{adm}}^{\mu}$; see (2.12). Assuming that every $\mathcal{H}_n^0$ is generated by centered point observations $Z^0(x_1), Z^0(x_2), \dots$, the above argument shows that, for any $h = \sum_{\ell=1}^L c_\ell Z(x_\ell^*)$ in $\mathbb{R} \oplus \mathcal{Z}^0$, the kriging predictors $\{h_n\}_{n\in\mathbb{N}}$ based on $\{\mathcal{H}_n\}_{n\in\mathbb{N}}$, $\mathcal{H}_n := \mathbb{R} \oplus \mathcal{H}_n^0$, are consistent whenever the sequence of observation points $\{x_j\}_{j\in\mathbb{N}}$ *accumulates at any* $x^* \in \mathcal{X}$. Since $\mathcal{Z}^0$ is dense in the Hilbert space $\mathcal{H}^0$, the same is true for every $h \in \mathcal{H} = \mathbb{R} \oplus \mathcal{H}^0$.

Suppose that $\widetilde{\mu} = \mathsf{N}(\widetilde{m}, \widetilde{\mathcal{C}})$ is a second Gaussian measure on $L_2(\mathcal{X}, \nu_{\mathcal{X}})$ with mean value function $\widetilde{m} \in L_2(\mathcal{X}, \nu_{\mathcal{X}})$ and trace-class covariance operator $\widetilde{\mathcal{C}} : L_2(\mathcal{X}, \nu_{\mathcal{X}}) \to L_2(\mathcal{X}, \nu_{\mathcal{X}})$. Let $\widetilde{\mathsf{E}}[\cdot]$, $\widetilde{\mathsf{Var}}[\cdot]$ and $\widetilde{\mathsf{Cov}}[\cdot,\cdot]$ denote the real-valued expectation, variance and covariance operators under $\widetilde{\mu}$. We are now interested in the asymptotic behavior of the linear predictor based on $\widetilde{\mu}$. That is, what happens if, instead of the kriging predictor h_n, we use the linear predictor $\widetilde{h}_n$, which is the kriging predictor if $\widetilde{\mu}$ was the correct model?

3. General results on compact metric spaces. We first generalize the results of [21], Section 3, and [25], Chapter 4, Theorem 10 to our setting of Gaussian processes on a compact metric space $(\mathcal{X}, d_{\mathcal{X}})$. That is, uniformly asymptotically optimal linear prediction under the assumption that the two Gaussian measures μ and $\widetilde{\mu}$ are equivalent on $L_2(\mathcal{X}, \nu_{\mathcal{X}})$.

THEOREM 3.1. *Let $\mu = \mathsf{N}(m, \mathcal{C})$ and $\widetilde{\mu} = \mathsf{N}(\widetilde{m}, \widetilde{\mathcal{C}})$ be equivalent. Define $\mathcal{H}^0$ and $\widetilde{\mathcal{H}}^0$ as in (2.4) with respect to μ and $\widetilde{\mu}$, respectively. Then $\mathcal{H}^0$ and $\widetilde{\mathcal{H}}^0$ are norm equivalent, $\mathcal{S}^{\mu}_{\mathrm{adm}} = \mathcal{S}^{\widetilde{\mu}}_{\mathrm{adm}}$ (see (2.12)), and for all $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}^{\mu}_{\mathrm{adm}}$ the following hold*:

$$\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \frac{\mathsf{E}[(\widetilde{h}_n - h)^2]}{\mathsf{E}[(h_n - h)^2]} = \lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \frac{\widetilde{\mathsf{E}}[(h_n - h)^2]}{\widetilde{\mathsf{E}}[(\widetilde{h}_n - h)^2]} = 1, \tag{3.1}$$

$$\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \left| \frac{\widetilde{\mathsf{E}}[(h_n - h)^2]}{\mathsf{E}[(h_n - h)^2]} - 1 \right| = \lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \left| \frac{\mathsf{E}[(\widetilde{h}_n - h)^2]}{\widetilde{\mathsf{E}}[(\widetilde{h}_n - h)^2]} - 1 \right| = 0. \tag{3.2}$$

Here, $h_n, \widetilde{h}_n$ are the best linear predictors of $h \in \mathcal{H}$ based on $\mathcal{H}_n$ and the measures μ and $\widetilde{\mu}$, respectively. The set $\mathcal{H}_{-n} \subset \mathcal{H}$ contains all elements with $\mathsf{E}[(h_n - h)^2] > 0$.

The norm equivalence of $\mathcal{H}^0$ and $\widetilde{\mathcal{H}}^0$ guarantees that the best linear predictors $\{\widetilde{h}_n\}_{n\in\mathbb{N}}$ of $h \in \mathcal{H}$ based on $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}^{\mu}_{\mathrm{adm}}$ and $\widetilde{\mu}$ are well-defined. Furthermore, it corresponds to equivalence of $\mathsf{Var}[\cdot]$ and $\widetilde{\mathsf{Var}}[\cdot]$ on $\mathcal{H}$ (see Proposition 3.5 below) so that in combination with the restriction $h \in \mathcal{H}_{-n}$ it ensures that the case $0/0$ in (3.1) and (3.2) is evaded. Indeed, for every $h \in \mathcal{H}_{-n}$, we obtain

$$\widetilde{\mathsf{E}}[(\widetilde{h}_n - h)^2] = \widetilde{\mathsf{Var}}[\widetilde{h}_n - h] \geq c\,\mathsf{Var}[\widetilde{h}_n - h] \geq c\,\mathsf{Var}[h_n - h] = c\,\mathsf{E}[(h_n - h)^2] > 0,$$

with $c \in (0, \infty)$ independent of n and h.

Equivalence of the Gaussian measures $\mu = \mathsf{N}(m, \mathcal{C})$ and $\widetilde{\mu} = \mathsf{N}(\widetilde{m}, \widetilde{\mathcal{C}})$ implies that $m - \widetilde{m}$ is an element of the Cameron–Martin space

$$H^* := \mathcal{C}^{1/2}(L_2(\mathcal{X}, \nu_{\mathcal{X}})), \qquad (\cdot\,, \cdot)_{H^*} := \left(\mathcal{C}^{-1/2}\,\cdot\,, \mathcal{C}^{-1/2}\,\cdot\right)_{L_2(\mathcal{X}, \nu_{\mathcal{X}})},$$

which also is a Hilbert space; see Appendix A in the Supplementary Material [9]. However, m and $\widetilde{m}$ are not necessarily both elements of H^*. Thus, Theorem 3.1 generalizes ([25], Chapter 4, Theorem 10) where $m = 0$ is assumed, even on Euclidean domains.

These results for equivalent measures μ and $\widetilde{\mu}$ also apply when considering the variances of the prediction errors. This is subject of the next corollary.

COROLLARY 3.2. *The statements of Theorem 3.1 remain true if we replace each second moment in (3.1) and (3.2) by the corresponding variance. That is, for all $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}^{\mu}_{\mathrm{adm}}$, the following hold*:

$$\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \frac{\mathsf{Var}[\widetilde{h}_n - h]}{\mathsf{Var}[h_n - h]} = \lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \frac{\widetilde{\mathsf{Var}}[h_n - h]}{\widetilde{\mathsf{Var}}[\widetilde{h}_n - h]} = 1, \tag{3.3}$$

$$\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \left| \frac{\widetilde{\mathsf{Var}}[h_n - h]}{\mathsf{Var}[h_n - h]} - 1 \right| = \lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \left| \frac{\mathsf{Var}[\widetilde{h}_n - h]}{\widetilde{\mathsf{Var}}[\widetilde{h}_n - h]} - 1 \right| = 0. \tag{3.4}$$

Theorem 3.1 shows that equivalence of μ and $\widetilde{\mu}$ is *sufficient* for uniformly asymptotically optimal linear prediction. The following (less restrictive) assumptions will subsequently be shown to be *necessary and sufficient*.

ASSUMPTION 3.3. Let $\varrho, \widetilde{\varrho} \colon \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ be two continuous, (strictly) positive definite covariance functions with corresponding covariance operators $\mathcal{C}$, $\widetilde{\mathcal{C}}$, defined on $L_2(\mathcal{X}, \nu_{\mathcal{X}})$ via (2.1). Assume that $\mathcal{C}, \widetilde{\mathcal{C}}$ and $m, \widetilde{m} \in L_2(\mathcal{X}, \nu_{\mathcal{X}})$ are such that:

I. The Cameron–Martin spaces $H^* = \mathcal{C}^{1/2}(L_2(\mathcal{X}, \nu_{\mathcal{X}}))$ and $\widetilde{H}^* = \widetilde{\mathcal{C}}^{1/2}(L_2(\mathcal{X}, \nu_{\mathcal{X}}))$ are norm equivalent Hilbert spaces.

II. The difference between the mean value functions $m, \widetilde{m} \in L_2(\mathcal{X}, \nu_{\mathcal{X}})$ is an element of the Cameron–Martin space, that is, $m - \widetilde{m} \in H^*$.

III. There exists a positive real number $a \in (0, \infty)$ such that the operator

$$T_a \colon L_2(\mathcal{X}, \nu_{\mathcal{X}}) \to L_2(\mathcal{X}, \nu_{\mathcal{X}}), \qquad T_a := \mathcal{C}^{-1/2}\widetilde{\mathcal{C}}\mathcal{C}^{-1/2} - a\mathcal{I} \tag{3.5}$$

is compact. Here and below, $\mathcal{I}$ denotes the identity on $L_2(\mathcal{X}, \nu_{\mathcal{X}})$.

REMARK 3.4. We briefly comment on the similarities and differences between Assumptions 3.3.I–III and the set of assumptions in the Feldman–Hájek theorem (Theorem A.1 for $E = L_2(\mathcal{X}, \nu_{\mathcal{X}})$; see Appendix A in the Supplementary Material [9]) that are necessary and sufficient for equivalence of two Gaussian measures $\mu = \mathsf{N}(m, \mathcal{C})$ and $\widetilde{\mu} = \mathsf{N}(\widetilde{m}, \widetilde{\mathcal{C}})$. Note that all assumptions are identical except for the third. For equivalence of the measures μ and $\widetilde{\mu}$, the operator $T_1 = \mathcal{C}^{-1/2}\widetilde{\mathcal{C}}\mathcal{C}^{-1/2} - \mathcal{I}$ has to be Hilbert–Schmidt on $L_2(\mathcal{X}, \nu_{\mathcal{X}})$. Since every Hilbert–Schmidt operator is compact, this in particular implies that Assumption 3.3.III holds for $a = 1$. This shows the greater generality of our Assumptions 3.3.I–III compared to the assumption that the two Gaussian measures μ and $\widetilde{\mu}$ are equivalent.

PROPOSITION 3.5. *Let $\mu = \mathsf{N}(m, \mathcal{C})$, $\widetilde{\mu} = \mathsf{N}(\widetilde{m}, \widetilde{\mathcal{C}})$, and define $\mathcal{H}^0$, $\widetilde{\mathcal{H}}^0$ as in* (2.4) *with respect to the measures μ and $\widetilde{\mu}$, respectively. The following are equivalent*:

(i) *Assumption* 3.3.I *is satisfied.*

(ii) *The linear operator $\widetilde{\mathcal{C}}^{1/2}\mathcal{C}^{-1/2} \colon L_2(\mathcal{X}, \nu_{\mathcal{X}}) \to L_2(\mathcal{X}, \nu_{\mathcal{X}})$ is an isomorphism, that is, it is bounded and has a bounded inverse.*

(iii) *The Hilbert spaces $\mathcal{H}^0$, $\widetilde{\mathcal{H}}^0$ are norm equivalent. In particular, there exist constants $k_0, k_1 \in (0, \infty)$ such that $k_0 \operatorname{Var}[h] \le \widetilde{\operatorname{Var}}[h] \le k_1 \operatorname{Var}[h]$, for all $h \in \mathcal{H}$, with $\mathcal{H}$ as in* (2.7).

(iv) *There exist constants $0 < k \le K < \infty$ such that, for all $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}^{\mu}_{\mathrm{adm}}$, any of the following fractions is bounded from below by $k > 0$ and from above by $K < \infty$, uniformly with respect to $n \in \mathbb{N}$ and $h \in \mathcal{H}_{-n}$*:

$$\frac{\widetilde{\operatorname{Var}}[h_n - h]}{\operatorname{Var}[h_n - h]}, \qquad \frac{\operatorname{Var}[\widetilde{h}_n - h]}{\widetilde{\operatorname{Var}}[\widetilde{h}_n - h]}, \qquad \frac{\operatorname{Var}[\widetilde{h}_n - h]}{\operatorname{Var}[h_n - h]}, \qquad \frac{\widetilde{\operatorname{Var}}[h_n - h]}{\widetilde{\operatorname{Var}}[\widetilde{h}_n - h]}. \tag{3.6}$$

Here, h_n, $\widetilde{h}_n$ are the best linear predictors of h based on $\mathcal{H}_n$ and μ, respectively, $\widetilde{\mu}$.

Proposition 3.5 elucidates the role of Assumption 3.3.I: As previously noted, the norm equivalence of the spaces $\mathcal{H}^0$ and $\widetilde{\mathcal{H}}^0$ in (iii) ensures that, for any $h \in \mathcal{H}$, the best linear predictors $\{\widetilde{h}_n\}_{n\in\mathbb{N}}$ based on $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}^{\mu}_{\mathrm{adm}}$ and the measure $\widetilde{\mu}$ are well-defined. Furthermore, uniform boundedness of the fractions in (iv) guarantees that the sequence $\{\widetilde{h}_n\}_{n\in\mathbb{N}}$ is μ-consistent,

$$\lim_{n\to\infty} \operatorname{Var}[\widetilde{h}_n - h] \le \sup_{\ell\in\mathbb{N}} \sup_{g\in\mathcal{H}_{-\ell}} \frac{\operatorname{Var}[\widetilde{g}_\ell - g]}{\operatorname{Var}[g_\ell - g]} \lim_{n\to\infty} \mathsf{E}\big[(h_n - h)^2\big] = 0,$$

which clearly is necessary for asymptotically optimal linear prediction.

Including one more assumption, namely Assumption 3.3.III, yields necessary and sufficient conditions for uniform asymptotic optimality of linear predictions, when the quality of the linear predictors is measured by the variance of the error. This result is formulated in the following theorem.

THEOREM 3.6. *Let $\mu = \mathsf{N}(m, \mathcal{C})$ and $\widetilde{\mu} = \mathsf{N}(\widetilde{m}, \widetilde{\mathcal{C}})$. In addition, let h_n, $\widetilde{h}_n$ denote the best linear predictors of h based on $\mathcal{H}_n$ and the measures μ and $\widetilde{\mu}$, respectively. Then, any of the assertions,*

$$\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \frac{\mathsf{Var}[\widetilde{h}_n - h]}{\mathsf{Var}[h_n - h]} = 1, \tag{3.7}$$

$$\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \frac{\widetilde{\mathsf{Var}}[h_n - h]}{\widetilde{\mathsf{Var}}[\widetilde{h}_n - h]} = 1, \tag{3.8}$$

$$\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \left| \frac{\widetilde{\mathsf{Var}}[h_n - h]}{\mathsf{Var}[h_n - h]} - a \right| = 0, \tag{3.9}$$

$$\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \left| \frac{\mathsf{Var}[\widetilde{h}_n - h]}{\widetilde{\mathsf{Var}}[\widetilde{h}_n - h]} - \frac{1}{a} \right| = 0, \tag{3.10}$$

holds for all $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}^{\mu}_{\mathrm{adm}}$ if and only if Assumptions 3.3.I and 3.3.III are fulfilled. The constant $a \in (0, \infty)$ in (3.9) and (3.10) is the same as that in (3.5) of Assumption 3.3.III.

REMARK 3.7. Theorem 3.6 shows, in particular, that either all of the four asymptotic statements (3.7)–(3.10) hold simultaneously or none of them are true.

Finally, when measuring the quality of the linear predictors in terms of the mean squared error, additionally the behavior of the difference $m - \widetilde{m}$ between the mean value functions matters, and all three of Assumptions 3.3.I–III are necessary and sufficient for uniform asymptotic optimality in this sense. This characterization is formulated in Theorem 3.8, which is our main result.

THEOREM 3.8. *Let $\mu = \mathsf{N}(m, \mathcal{C})$ and $\widetilde{\mu} = \mathsf{N}(\widetilde{m}, \widetilde{\mathcal{C}})$. In addition, let h_n, $\widetilde{h}_n$ denote the best linear predictors of h based on $\mathcal{H}_n$ and the measures μ and $\widetilde{\mu}$, respectively. Then any of the assertions,*

$$\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \frac{\mathsf{E}[(\widetilde{h}_n - h)^2]}{\mathsf{E}[(h_n - h)^2]} = 1, \tag{3.11}$$

$$\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \frac{\widetilde{\mathsf{E}}[(h_n - h)^2]}{\widetilde{\mathsf{E}}[(\widetilde{h}_n - h)^2]} = 1, \tag{3.12}$$

$$\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \left| \frac{\widetilde{\mathsf{E}}[(h_n - h)^2]}{\mathsf{E}[(h_n - h)^2]} - a \right| = 0, \tag{3.13}$$

$$\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \left| \frac{\mathsf{E}[(\widetilde{h}_n - h)^2]}{\widetilde{\mathsf{E}}[(\widetilde{h}_n - h)^2]} - \frac{1}{a} \right| = 0, \tag{3.14}$$

holds for all $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}^{\mu}_{\mathrm{adm}}$ if and only if Assumptions 3.3.I–III are satisfied. The constant $a \in (0, \infty)$ in (3.13) and (3.14) is the same as that in (3.5) of Assumption 3.3.III.

4. Proofs of the results. Throughout this section, we abbreviate $L_2(\mathcal{X}, \nu_{\mathcal{X}})$ by L_2, $\mathcal{L}(L_2)$ is the space of bounded linear operators on L_2 and the subspaces $\mathcal{K}(L_2) \subset \mathcal{L}(L_2)$ as well as $\mathcal{L}_2(L_2) \subset \mathcal{L}(L_2)$ contain all compact and Hilbert–Schmidt operators, respectively (see Appendix A in the Supplementary Material [9]).

Recall that $Z^0 = Z - m$, where Z is a Gaussian process on $(\mathcal{X}, d_{\mathcal{X}})$ with corresponding Gaussian measure $\mu = \mathsf{N}(m, \mathcal{C})$, and that $\{e_j\}_{j\in\mathbb{N}}$ is an orthonormal basis for L_2 consisting of (the continuous representatives of) eigenfunctions of the covariance operator $\mathcal{C}$, with corresponding positive eigenvalues $\{\gamma_j\}_{j\in\mathbb{N}}$. In the next lemma, a relation between the (dual of the) Cameron–Martin space for μ and the Hilbert space $\mathcal{H}^0$ in (2.4) is established, similarly as in [16], Theorem 5D, or [3], Theorem 35. This relation will be crucial for proving all results.

LEMMA 4.1. *For each $j \in \mathbb{N}$, define $v_j := \frac{1}{\sqrt{\gamma_j}} e_j$ as well as the real-valued random variable $z_j := (Z^0, v_j)_{L_2}$. Then the following hold*:

(i) *$\{v_j\}_{j\in\mathbb{N}}$ is an orthonormal basis for $H := \mathcal{C}^{-1/2}(L_2)$, which is the dual of the Cameron–Martin space $H^* = \mathcal{C}^{1/2}(L_2)$ with $(v, v')_H := (\mathcal{C}^{1/2}v, \mathcal{C}^{1/2}v')_{L_2}$.*

(ii) *$\{z_j\}_{j\in\mathbb{N}}$ is an orthonormal basis for the Hilbert space $\mathcal{H}^0$ equipped with the inner product $(\cdot,\cdot)_{\mathcal{H}^0} = \mathsf{E}[\cdot\,\cdot] = \mathsf{Cov}[\cdot,\cdot]$, see* (2.4).

(iii) *The linear operator*

$$\mathcal{J}: H \to \mathcal{H}^0, \quad \textit{satisfying} \quad \mathcal{J}v = (Z^0, v)_{L_2} \quad \forall v \in L_2 \subset H, \tag{4.1}$$

is a well-defined isometric isomorphism and, for all $v, v' \in H$, we have

$$(v, v')_H = (\mathcal{C}^{1/2}v, \mathcal{C}^{1/2}v')_{L_2} = \mathsf{Cov}[\mathcal{J}v, \mathcal{J}v'] = \mathsf{E}[\mathcal{J}v\mathcal{J}v'] = (\mathcal{J}v, \mathcal{J}v')_{\mathcal{H}^0}. \tag{4.2}$$

PROOF. Since $L_2 \subset H$ and $v_j = \mathcal{C}^{-1/2}e_j$, claim (i)—orthonormality and the basis property of $\{v_j\}_{j\in\mathbb{N}}$ in H—follows directly from the corresponding properties of $\{e_j\}_{j\in\mathbb{N}}$ in L_2.

To prove the second assertion (ii), we first note that $v_j : \mathcal{X} \to \mathbb{R}$ is continuous for every $j \in \mathbb{N}$. Thus, by Lemma B.3 (see Appendix B in the Supplementary Material [9]) we find that $z_j \in \mathcal{H}^0$ for all $j \in \mathbb{N}$. Next, we prove that $\{z_j\}_{j\in\mathbb{N}}$ constitutes an orthonormal basis for $\mathcal{H}^0$. For this, we need to show that $(z_i, z_j)_{\mathcal{H}^0} = \delta_{ij}$ (*orthonormality*) and

$$(h, z_j)_{\mathcal{H}^0} = 0 \quad \forall j \in \mathbb{N} \quad \Rightarrow \quad h = 0 \qquad (\textit{basis poperty}).$$

Due to the identities $\mathsf{E}[z_i z_j] = \mathsf{Cov}[(Z^0, v_i)_{L_2}, (Z^0, v_j)_{L_2}] = (\mathcal{C}v_i, v_j)_{L_2} = (v_i, v_j)_H$, orthonormality follows from (i). Now let $h \in \mathcal{H}^0$ be such that $(h, z_j)_{\mathcal{H}^0} = 0$ vanishes for all $j \in \mathbb{N}$. By Fubini's theorem, we then obtain that, for all $j \in \mathbb{N}$,

$$0 = \mathsf{E}[h(Z^0, e_j)_{L_2}] = \int_{\mathcal{X}} \mathsf{E}[hZ^0(x)]e_j(x)\,\mathrm{d}\nu_{\mathcal{X}}(x) = (\mathsf{E}[hZ^0(\cdot)], e_j)_{L_2}.$$

Since $\{e_j\}_{j\in\mathbb{N}}$ is an orthonormal basis for L_2 and the mapping $\mathcal{X} \ni x \mapsto \mathsf{E}[hZ^0(x)] \in \mathbb{R}$ is continuous, this shows that $\mathsf{E}[hZ^0(x)] = 0$ for all $x \in \mathcal{X}$, which implies (due to strict positive definiteness of ϱ) that $h \in \mathcal{H}^0$ has to vanish. We conclude (ii), $\{z_j\}_{j\in\mathbb{N}}$ is an orthonormal basis for $\mathcal{H}^0$.

It remains to prove (iii). Clearly, $\mathcal{J}v_j = z_j$ for all $j \in \mathbb{N}$. Thus, the linear mapping $\mathcal{J}: H \to \mathcal{H}^0$ is well-defined and an isometry, since by (i) and (ii) $\{v_j\}_{j\in\mathbb{N}}$ and $\{z_j\}_{j\in\mathbb{N}}$ are orthonormal bases for H and $\mathcal{H}^0$, respectively. Furthermore,

$$\begin{aligned}(v, v')_H = (\mathcal{C}v, v')_{L_2} = \mathsf{Cov}[(Z^0, v)_{L_2}, (Z^0, v')_{L_2}] &= \mathsf{Cov}[\mathcal{J}v, \mathcal{J}v'] \\ &= \mathsf{E}[\mathcal{J}v\mathcal{J}v'] = (\mathcal{J}v, \mathcal{J}v')_{\mathcal{H}^0}\end{aligned}$$

holds for all $v, v' \in L_2$, completing the proof of (iii) by density of L_2 in H. □

PROOF OF PROPOSITION 3.5. We first show (i) ⇒ (ii) ⇒ (iii) ⇒ (i), followed by the proof of the equivalence (iii) ⇔ (iv):

(i) ⇒ (ii): Under Assumption 3.3.I, the norms on H^* and $\widetilde{H}^*$ are equivalent, that is, there are $c_0, c_1 \in (0, \infty)$ such that $\|\mathcal{C}^{-1/2}u\|_{L_2} \le c_0\|\widetilde{\mathcal{C}}^{-1/2}u\|_{L_2}$ and $\|\widetilde{\mathcal{C}}^{-1/2}u\|_{L_2} \le c_1\|\mathcal{C}^{-1/2}u\|_{L_2}$

for all $u \in H^* = \widetilde{H}^*$. Thus, for any $w \in L_2$, $\widetilde{\mathcal{C}}^{1/2}w \in H^*$ with $\|\mathcal{C}^{-1/2}\widetilde{\mathcal{C}}^{1/2}w\|_{L_2} \le c_0\|w\|_{L_2}$ and, in addition, $\mathcal{C}^{1/2}w \in \widetilde{H}^*$ with $\|\widetilde{\mathcal{C}}^{-1/2}\mathcal{C}^{1/2}w\|_{L_2} \le c_1\|w\|_{L_2}$. This shows that $\mathcal{C}^{-1/2}\widetilde{\mathcal{C}}^{1/2}$, and thus, also its adjoint $\widetilde{\mathcal{C}}^{1/2}\mathcal{C}^{-1/2}$ are isomorphisms on L_2, and (ii) follows.

(ii) ⇒ (iii): Let the Hilbert space H and the isometry $\mathcal{J}\colon H \to \mathcal{H}^0$ be defined as in Lemma 4.1(iii). In addition, define $\widetilde{H} := \widetilde{\mathcal{C}}^{-1/2}(L_2)$ and the isometry $\widetilde{\mathcal{J}}\colon \widetilde{H} \to \widetilde{\mathcal{H}}^0$ in the obvious analogous way. If $c_0 := \|\widetilde{\mathcal{C}}^{1/2}\mathcal{C}^{-1/2}\|_{\mathcal{L}(L_2)}$ and $c_1 := \|\mathcal{C}^{1/2}\widetilde{\mathcal{C}}^{-1/2}\|_{\mathcal{L}(L_2)}$ are finite, then H and $\widetilde{H}$ are norm equivalent, that is, $c_1^{-1}\|v\|_H \le \|v\|_{\widetilde{H}} \le c_0\|v\|_H$ holds for all $v \in H = \widetilde{H}$, and the inclusion mapping $\mathcal{I}_{H\to\widetilde{H}}$ of H in $\widetilde{H}$ is continuous. Thus, we obtain $\|\widetilde{\mathcal{J}}\mathcal{I}_{H\to\widetilde{H}}\mathcal{J}^{-1}h^0\|_{\widetilde{\mathcal{H}}^0} = \|\mathcal{I}_{H\to\widetilde{H}}\mathcal{J}^{-1}h^0\|_{\widetilde{H}} \le c_0\|\mathcal{J}^{-1}h^0\|_H = c_0\|h^0\|_{\mathcal{H}^0}$ for every $h^0 \in \mathcal{H}^0$. Next, let $h^0 \in \mathcal{H}^0$ and set $v^h := \mathcal{J}^{-1}h^0 \in H$. Then we observe the identities $\|\widetilde{\mathcal{J}}\mathcal{I}_{H\to\widetilde{H}}\mathcal{J}^{-1}h^0\|^2_{\widetilde{\mathcal{H}}^0} = \|\widetilde{\mathcal{J}}v^h\|^2_{\widetilde{\mathcal{H}}^0} = \widetilde{\mathsf{Var}}[\widetilde{\mathcal{J}}v^h] = \widetilde{\mathsf{Var}}[\mathcal{J}v^h] = \widetilde{\mathsf{Var}}[h^0]$, which combined with the above show that $\widetilde{\mathsf{Var}}[h] = \widetilde{\mathsf{Var}}[h^0] \le c_0^2\|h^0\|^2_{\mathcal{H}^0} = c_0^2\,\mathsf{Var}[h]$ for all $h \in \mathcal{H}$, where we set $h^0 := h - \mathsf{E}[h] \in \mathcal{H}^0$. Similarly, we derive $\|\mathcal{J}\mathcal{I}_{\widetilde{H}\to H}\widetilde{\mathcal{J}}^{-1}\widetilde{h}^0\|_{\mathcal{H}^0} \le c_1\|\widetilde{h}^0\|_{\widetilde{\mathcal{H}}^0}$ for all $\widetilde{h}^0 \in \widetilde{\mathcal{H}}^0$, and we may change the roles of H and $\widetilde{H}$ (resp. of $\mathcal{H}^0$ and $\widetilde{\mathcal{H}}^0$) to conclude that also the relation $\mathsf{Var}[h] \le c_1^2\widetilde{\mathsf{Var}}[h]$ holds for all $h \in \widetilde{\mathcal{H}} = \mathbb{R} \oplus \widetilde{\mathcal{H}}^0$.

(iii) ⇒ (i): We prove that the dual spaces, $H = \mathcal{C}^{-1/2}(L_2)$, $\widetilde{H} = \widetilde{\mathcal{C}}^{-1/2}(L_2)$, are norm equivalent, which implies the result for H^* and $\widetilde{H}^*$. Norm equivalence of $\mathcal{H}^0$ and $\widetilde{\mathcal{H}}^0$ implies continuity of the inclusion maps $\mathcal{I}_{\mathcal{H}^0\to\widetilde{\mathcal{H}}^0}$, $\mathcal{I}_{\widetilde{\mathcal{H}}^0\to\mathcal{H}^0}$, which similarly as above, yields continuity of $\widetilde{\mathcal{J}}^{-1}\mathcal{I}_{\mathcal{H}^0\to\widetilde{\mathcal{H}}^0}\mathcal{J}\colon H \to \widetilde{H}$ and of $\mathcal{J}^{-1}\mathcal{I}_{\widetilde{\mathcal{H}}^0\to\mathcal{H}^0}\widetilde{\mathcal{J}}\colon \widetilde{H} \to H$. Thus, it follows that $\|v\|_{\widetilde{H}} \le c_0\|v\|_H$ and $\|\widetilde{v}\|_H \le c_1\|\widetilde{v}\|_{\widetilde{H}}$ hold for all $v \in H$, $\widetilde{v} \in \widetilde{H}$ with some constants $c_0, c_1 \in (0,\infty)$, since $\|\widetilde{\mathcal{J}}^{-1}\mathcal{I}_{\mathcal{H}^0\to\widetilde{\mathcal{H}}^0}\mathcal{J}v\|_{\widetilde{H}} = \|v\|_{\widetilde{H}}$ and $\|\mathcal{J}^{-1}\mathcal{I}_{\widetilde{\mathcal{H}}^0\to\mathcal{H}^0}\widetilde{\mathcal{J}}\widetilde{v}\|_H = \|\widetilde{v}\|_H$.

(iii) ⇒ (iv): Suppose that $k_0\,\mathsf{Var}[h] \le \widetilde{\mathsf{Var}}[h] \le k_1\,\mathsf{Var}[h]$ holds for every $h \in \mathcal{H} = \widetilde{\mathcal{H}}$ and let $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}^{\mu}_{\mathrm{adm}}$. Then, for every $n \in \mathbb{N}$ and all $h \in \mathcal{H}_{-n}$, $k_0 \le \frac{\widetilde{\mathsf{Var}}[h_n-h]}{\mathsf{Var}[h_n-h]} \le k_1$ as well as $k_1^{-1} \le \frac{\mathsf{Var}[\widetilde{h}_n-h]}{\widetilde{\mathsf{Var}}[\widetilde{h}_n-h]} \le k_0^{-1}$ readily follow. Subsequently, we find that

$$1 \le \frac{\mathsf{Var}[\widetilde{h}_n - h]}{\mathsf{Var}[h_n - h]} = \frac{\mathsf{Var}[\widetilde{h}_n - h]}{\widetilde{\mathsf{Var}}[\widetilde{h}_n - h]}\,\frac{\widetilde{\mathsf{Var}}[\widetilde{h}_n - h]}{\widetilde{\mathsf{Var}}[h_n - h]}\,\frac{\widetilde{\mathsf{Var}}[h_n - h]}{\mathsf{Var}[h_n - h]} \le k_0^{-1}k_1, \tag{4.3}$$

and a similar trick shows that also $\frac{\widetilde{\mathsf{Var}}[h_n-h]}{\widetilde{\mathsf{Var}}[\widetilde{h}_n-h]} \in [1, k_1k_0^{-1}]$ for all n and h.

(iv) ⇒ (iii): We first show necessity of (iii) for uniform boundedness (from above and below) of the first two fractions in (3.6). For this, let $h \in \mathcal{Z}^0 \setminus \{0\}$. By positive definiteness of ϱ, there exists $\phi \in \mathcal{Z}^0$ so that $\{h, \phi\}$ are linearly independent. Define $\psi_1' := \phi - \frac{(\phi,h)_{\mathcal{H}^0}}{(h,h)_{\mathcal{H}^0}}h$, $\psi_1 := \frac{1}{\|\psi_1'\|_{\mathcal{H}^0}}\psi_1'$ and $\widetilde{\psi}_1' := \phi - \frac{(\phi,h)_{\widetilde{\mathcal{H}}^0}}{(h,h)_{\widetilde{\mathcal{H}}^0}}h$, $\widetilde{\psi}_1 := \frac{1}{\|\widetilde{\psi}_1'\|_{\mathcal{H}^0}}\widetilde{\psi}_1'$, and note that $h \in \mathcal{Z}^0$ is orthogonal to $\psi_1 \in \mathcal{Z}^0$ in $\mathcal{H}^0$ and to $\widetilde{\psi}_1 \in \mathcal{Z}^0$ in $\widetilde{\mathcal{H}}^0$. By separability of $\mathcal{H}^0$, there exist sequences $\{\psi_j\}_{j\ge2}$ and $\{\widetilde{\psi}_j\}_{j\ge2}$ such that $\{\psi_j\}_{j\in\mathbb{N}}$ and $\{\widetilde{\psi}_j\}_{j\in\mathbb{N}}$ are orthonormal bases for $\mathcal{H}^0$. For $n \in \mathbb{N}$, define the spaces $\mathcal{H}_n^\star := \mathbb{R} \oplus \operatorname{span}\{\psi_1,\dots,\psi_n\}$, $\mathcal{H}_n^{\circledast} := \mathbb{R} \oplus \operatorname{span}\{\widetilde{\psi}_1,\dots,\widetilde{\psi}_n\}$. Then $\{\mathcal{H}_n^\star\}_{n\in\mathbb{N}}, \{\mathcal{H}_n^{\circledast}\}_{n\in\mathbb{N}} \in \mathcal{S}^{\mu}_{\mathrm{adm}}$ and, if h_1 denotes the best linear predictor of h based on $\mathcal{H}_1^\star$ and μ, and $\widetilde{h}_1$ denotes the best linear predictor of h based on $\mathcal{H}_1^{\circledast}$ and $\widetilde{\mu}$, then $h_1 = \widetilde{h}_1 = 0$ follows. By boundedness of the first or second fraction in (3.6) (with $n = 1$), we have that $\frac{\widetilde{\mathsf{Var}}[h]}{\mathsf{Var}[h]} = \frac{\widetilde{\mathsf{Var}}[h_1-h]}{\mathsf{Var}[h_1-h]} \in [k, K]$ or $\frac{\mathsf{Var}[h]}{\widetilde{\mathsf{Var}}[h]} = \frac{\mathsf{Var}[\widetilde{h}_1-h]}{\widetilde{\mathsf{Var}}[\widetilde{h}_1-h]} \in [k, K]$. Thus, in both cases $k_0\,\mathsf{Var}[h] \le \widetilde{\mathsf{Var}}[h] \le k_1\,\mathsf{Var}[h]$ holds, where $k_0 := \min\{k, K^{-1}\}$ and $k_1 := \max\{K, k^{-1}\}$. Since $h \in \mathcal{Z}^0$ was arbitrary and since the constants $k, K \in (0,\infty)$ do not depend on h (as the fractions in (3.6) are bounded *uniformly* in $\{\mathcal{H}_n\}_{n\in\mathbb{N}}$, n and h), assertion (iii) follows by density of $\mathcal{Z}^0$ in $\mathcal{H}^0$ and in $\widetilde{\mathcal{H}}^0$.

Assume next that the third fraction in (3.6) is bounded, uniformly with respect to $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}^{\mu}_{\mathrm{adm}}$, $n \in \mathbb{N}$ and $h \in \mathcal{H}_{-n}$. In either of the two cases, $\alpha_0 = 0$ or $\alpha_1 = \infty$, where $\alpha_0 := \inf_{h\in\mathcal{H}^0\cap\widetilde{\mathcal{H}}^0}\frac{\widetilde{\mathsf{Var}}[h]}{\mathsf{Var}[h]}$, $\alpha_1 := \sup_{h\in\mathcal{H}^0\cap\widetilde{\mathcal{H}}^0}\frac{\widetilde{\mathsf{Var}}[h]}{\mathsf{Var}[h]}$, it follows as in [5], Proof of Theorem 5, that there exist sequences $\{h^{(\ell)}\}_{\ell\in\mathbb{N}}, \{\psi_1^{(\ell)}\}_{\ell\in\mathbb{N}} \subset \mathcal{H}^0 \cap \widetilde{\mathcal{H}}^0$, normalized in $\mathcal{H}^0$, with

$\frac{\mathsf{Var}[\widetilde{h}_1^{(\ell)}-h^{(\ell)}]}{\mathsf{Var}[h_1^{(\ell)}-h^{(\ell)}]} \geq \ell$ for all $\ell \in \mathbb{N}$, where $h_1^{(\ell)}$, $\widetilde{h}_1^{(\ell)}$ are the best linear predictors of $h^{(\ell)}$ based on μ, respectively, $\widetilde{\mu}$ and $\mathcal{H}_1^{(\ell)} := \mathbb{R} \oplus \operatorname{span}\{\psi_1^{(\ell)}\}$. By separability of $\mathcal{H}^0$, for each $\ell \in \mathbb{N}$, we may complement $\psi_1^{(\ell)}$ to an orthonormal basis $\{\psi_j^{(\ell)}\}_{j\in\mathbb{N}}$ for $\mathcal{H}^0$. Thus, for all $\ell \in \mathbb{N}$, $\{\mathcal{H}_n^{(\ell)}\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^{\mu}$ holds, where $\mathcal{H}_n^{(\ell)} := \mathbb{R} \oplus \operatorname{span}\{\psi_1^{(\ell)}, \ldots, \psi_n^{(\ell)}\}$ and

$$\sup_{\ell\in\mathbb{N}} \sup_{n\in\mathbb{N}} \sup_{h\in\mathcal{H}_{-n}^{(\ell)}} \frac{\mathsf{Var}[\widetilde{h}_n^{(\ell)} - h]}{\mathsf{Var}[h_n^{(\ell)} - h]} \geq \sup_{\ell\in\mathbb{N}} \frac{\mathsf{Var}[\widetilde{h}_1^{(\ell)} - h^{(\ell)}]}{\mathsf{Var}[h_1^{(\ell)} - h^{(\ell)}]} = \infty,$$

contradicting uniform boundedness of the third fraction in (3.6). We therefore conclude that $\alpha_0, \alpha_1 \in (0, \infty)$, $\mathcal{H}^0 \cap \widetilde{\mathcal{H}}^0 = \mathcal{H}^0 = \widetilde{\mathcal{H}}^0$ and (iii) follows.

Finally, assuming uniform boundedness of the last fraction in (3.6), analogous arguments show that $\widetilde{\alpha}_0 := \inf_{h\in\mathcal{H}^0\cap\widetilde{\mathcal{H}}^0} \frac{\widetilde{\mathsf{Var}}[h]}{\mathsf{Var}[h]}$, $\widetilde{\alpha}_1 := \sup_{h\in\mathcal{H}^0\cap\widetilde{\mathcal{H}}^0} \frac{\widetilde{\mathsf{Var}}[h]}{\mathsf{Var}[h]}$ satisfy $\widetilde{\alpha}_0, \widetilde{\alpha}_1 \in (0, \infty)$, again yielding (iii). □

REMARK 4.2. The arguments in the proof of Proposition 3.5 imply, in particular, that under Assumption 3.3.I we have, for all $v, v' \in H = \widetilde{H}$, that

$$(v, v')_{\widetilde{H}} = (\widetilde{\mathcal{C}}^{1/2} v, \widetilde{\mathcal{C}}^{1/2} v')_{L_2} = \widetilde{\mathsf{Cov}}[\mathcal{J} v, \mathcal{J} v']. \tag{4.4}$$

LEMMA 4.3. *Suppose Assumption* 3.3.I *is satisfied and let* $\{\mathcal{H}_n\}_{n\in\mathbb{N}}$ *be a sequence of subspaces of* $\mathcal{H}$ *such that, for all* $n \in \mathbb{N}$, $\mathcal{H}_n$ *is of the form* (2.9). *Then, for every* $h \in \mathcal{H}$, *the kriging predictors* $\{h_n\}_{n\in\mathbb{N}}$ *based on* $\{\mathcal{H}_n\}_{n\in\mathbb{N}}$ *and the measure* μ *are* μ*-consistent if and only if the kriging predictors* $\{\widetilde{h}_n\}_{n\in\mathbb{N}}$ *based on* $\{\mathcal{H}_n\}_{n\in\mathbb{N}}$ *and* $\widetilde{\mu}$ *are* $\widetilde{\mu}$*-consistent. In particular, the sets* $\mathcal{S}_{\mathrm{adm}}^{\mu}$ *and* $\mathcal{S}_{\mathrm{adm}}^{\widetilde{\mu}}$ *are equal,* $\mathcal{S}_{\mathrm{adm}}^{\mu} = \mathcal{S}_{\mathrm{adm}}^{\widetilde{\mu}}$.

PROOF. Let $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^{\mu}$. By Proposition 3.5(i) ⇔ (iii), $\mathcal{H}^0$ and $\widetilde{\mathcal{H}}^0$ are norm equivalent. Thus, $\widetilde{\mathsf{E}}[(\widetilde{h}_n - h)^2] = \widetilde{\mathsf{Var}}[\widetilde{h}_n - h] \leq \widetilde{\mathsf{Var}}[h_n - h] \leq k_1 \mathsf{Var}[h_n - h] = k_1 \mathsf{E}[(h_n - h)^2]$ holds, for any $h \in \mathcal{H} = \widetilde{\mathcal{H}}$, where $k_1 \in (0, \infty)$ is independent of n and h. This shows that $\mathcal{S}_{\mathrm{adm}}^{\mu} \subseteq \mathcal{S}_{\mathrm{adm}}^{\widetilde{\mu}}$. Analogously, $\mathsf{E}[(h_n - h)^2] \leq k_0^{-1} \widetilde{\mathsf{E}}[(\widetilde{h}_n - h)^2]$ follows for $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^{\widetilde{\mu}}$, with $k_0 \in (0, \infty)$ independent of n and h, showing the reverse inclusion $\mathcal{S}_{\mathrm{adm}}^{\widetilde{\mu}} \subseteq \mathcal{S}_{\mathrm{adm}}^{\mu}$. □

PROPOSITION 4.4. *Let* $\mu = \mathsf{N}(m, \mathcal{C})$ *and* $\widetilde{\mu} = \mathsf{N}(\widetilde{m}, \widetilde{\mathcal{C}})$. *Suppose that Assumptions* 3.3.I *and* 3.3.III *are satisfied and let* $a \in (0, \infty)$ *be the constant in* (3.5) *of Assumption* 3.3.III. *In addition, let* h_n, $\widetilde{h}_n$ *denote the best linear predictors of* h *based on* $\mathcal{H}_n$ *and the measures* μ *and* $\widetilde{\mu}$, *respectively. Then* (3.7)–(3.10) *hold for all* $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^{\mu}$. *If, in addition, Assumption* 3.3.II *is fulfilled, then* (3.11)–(3.14) *hold for all* $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^{\mu}$.

PROOF. Let $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^{\mu}$. As shown in Proposition C.2 (see Appendix C in the Supplementary Material [9]), we can without loss of generality assume that μ has zero mean and that $\widetilde{\mu}$ has mean $\widetilde{m} - m$. We first show that Assumptions 3.3.I and 3.3.III imply that (3.9) and (3.10) hold. To this end, let $n \in \mathbb{N}$ and recall that h_n is the kriging predictor of h based on $\mathcal{H}_n = \mathbb{R} \oplus \mathcal{H}_n^0$ and μ. We let $\{\psi_1^{(n)}, \ldots, \psi_n^{(n)}\}$ be an $\mathcal{H}^0$-orthonormal basis for $\mathcal{H}_n^0$, that is, $\mathsf{E}[\psi_k^{(n)} \psi_\ell^{(n)}] = \delta_{k\ell}$. Since $\mathcal{H}^0$ is a separable Hilbert space there exists a countable orthonormal basis of the orthogonal complement of $\mathcal{H}_n^0$ in $\mathcal{H}^0$, which will be denoted by $\{\psi_k^{(n)}\}_{k>n}$. Then, by construction $\{\psi_k^{(n)}\}_{k\in\mathbb{N}}$ is an orthonormal basis for $\mathcal{H}^0$. We identify $\psi_k^{(n)} \in \mathcal{H}^0$ with $v_k^{(n)} := \mathcal{J}^{-1} \psi_k^{(n)} \in H$, where $\mathcal{J}\colon H \to \mathcal{H}^0$ is the isometric isomorphism in (4.1) from

Lemma 4.1(iii). Due to (4.2), $\{v_k^{(n)}\}_{k\in\mathbb{N}}$ is then an orthonormal basis for $H = \mathcal{C}^{-1/2}(L_2)$. Furthermore, we note that, for every $h \in \mathcal{H}_{-n}$, the vector $h_n - h \in \mathcal{H}^0$ can be written as a linear combination of $\{\psi_k^{(n)}\}_{k>n}$, that is, $h_n - h = \sum_{k=n+1}^{\infty} c_k^{(n)} \psi_k^{(n)}$ with $\sum_{k=n+1}^{\infty} |c_k^{(n)}|^2 < \infty$.

We recall the identities in (4.2) and (4.4) from Lemma 4.1(iii) and Remark 4.2 and rewrite the term (A) $:= |\widetilde{\mathsf{Var}}[h_n - h] - a\,\mathsf{Var}[h_n - h]|$ as follows:

$$
\begin{aligned}
(\mathrm{A}) &= \left| \sum_{k,\ell=n+1}^{\infty} c_k^{(n)} c_\ell^{(n)} \big(\widetilde{\mathsf{Cov}}[\psi_k^{(n)}, \psi_\ell^{(n)}] - a\,\mathsf{Cov}[\psi_k^{(n)}, \psi_\ell^{(n)}]\big) \right| \\
&= \left| \sum_{k,\ell=n+1}^{\infty} c_k^{(n)} c_\ell^{(n)} \big((\widetilde{\mathcal{C}}^{1/2} v_k^{(n)}, \widetilde{\mathcal{C}}^{1/2} v_\ell^{(n)})_{L_2} - a (\mathcal{C}^{1/2} v_k^{(n)}, \mathcal{C}^{1/2} v_\ell^{(n)})_{L_2} \big) \right|.
\end{aligned}
$$

Since $\{v_k^{(n)}\}_{k\in\mathbb{N}}$ is an orthonormal basis for $H = \mathcal{C}^{-1/2}(L_2)$, so is $\{w_k^{(n)}\}_{k\in\mathbb{N}}$ for L_2, where $w_k^{(n)} := \mathcal{C}^{1/2} v_k^{(n)}$. We set $w_n^h := \sum_{k=n+1}^{\infty} c_k^{(n)} w_k^{(n)}$ and obtain

$$
(\mathrm{A}) = \big| \big((\mathcal{C}^{-1/2} \widetilde{\mathcal{C}} \mathcal{C}^{-1/2} - a\mathcal{I}) Q_n^{\perp} w_n^h, Q_n^{\perp} w_n^h \big)_{L_2} \big|,
$$

where $Q_n^{\perp} := \mathcal{I} - Q_n$ and $Q_n \colon L_2 \to W_n$ denotes the L_2-orthogonal projection onto the subspace $W_n := \operatorname{span}\{w_1^{(n)}, \dots, w_n^{(n)}\}$. By Assumption 3.3.III, $T_a = \mathcal{C}^{-1/2} \widetilde{\mathcal{C}} \mathcal{C}^{-1/2} - a\mathcal{I}$ is compact on L_2. For this reason, there exists an orthonormal basis $\{b_j\}_{j\in\mathbb{N}}$ for L_2 consisting of eigenvectors of T_a with corresponding eigenvalues $\{\tau_j\}_{j\in\mathbb{N}} \subset \mathbb{R}$ accumulating only at zero. For $J \in \mathbb{N}$, we define $V_J := \operatorname{span}\{b_1, \dots, b_J\}$. We write $P_J \colon L_2 \to V_J$ for the corresponding L_2-orthogonal projection and set $P_J^{\perp} := \mathcal{I} - P_J$. Then, by invoking the chain of identities $\|w_n^h\|_{L_2}^2 = \sum_{k=n+1}^{\infty} |c_k^{(n)}|^2 = \mathsf{E}[(h_n - h)^2] = \mathsf{Var}[h_n - h]$, we estimate

$$
(\mathrm{A}) \le \mathsf{Var}[h_n - h] \sup_{\|w\|_{L_2}=1} \big| (T_a Q_n^{\perp} w, Q_n^{\perp} w)_{L_2} \big|.
$$

Clearly, if $T_a = 0$, we obtain that (A) $= 0$. Thus, from now on we assume that $\|T_a\|_{\mathcal{L}(L_2)} > 0$. Since $\mathcal{I} = P_J + P_J^{\perp}$ and $P_J^{\perp} T_a P_J = P_J T_a P_J^{\perp} = 0$, we find

$$
\begin{aligned}
\frac{(\mathrm{A})}{\mathsf{Var}[h_n - h]} &\le \sup_{\|w\|_{L_2}=1} \big| (P_J^{\perp} T_a P_J^{\perp} Q_n^{\perp} w, Q_n^{\perp} w)_{L_2} + (P_J T_a P_J Q_n^{\perp} w, Q_n^{\perp} w)_{L_2} \big| \\
&\le \sup_{\|w\|_{L_2}=1} \big| (T_a P_J^{\perp} w, P_J^{\perp} w)_{L_2} \big| + \sup_{\|w\|_{L_2}=1} \| Q_n^{\perp} P_J T_a P_J Q_n^{\perp} w \|_{L_2}.
\end{aligned} \tag{4.5}
$$

Here, we have used self-adjointness of T_a, P_J, $P_J^{\perp}$ and $Q_n^{\perp}$ on L_2 in the last step. Now fix $\varepsilon \in (0, \infty)$. Since $\lim_{j\to\infty} \tau_j = 0$, there exists $J_\varepsilon \in \mathbb{N}$ with

$$
\sup_{\|w\|_{L_2}=1} \big| (T_a P_{J_\varepsilon}^{\perp} w, P_{J_\varepsilon}^{\perp} w)_{L_2} \big| = \sup_{j > J_\varepsilon} |\tau_j| < \frac{\varepsilon}{2}. \tag{4.6}
$$

In addition, for $w := \sum_{k\in\mathbb{N}} \alpha_k^{(n)} w_k^{(n)} \in L_2$ and $h^w := \sum_{k\in\mathbb{N}} \alpha_k^{(n)} \psi_k^{(n)} \in \mathcal{H}^0$, with some square-summable coefficients $\{\alpha_k^{(n)}\}_{k\in\mathbb{N}}$, we find that

$$
\|Q_n^{\perp} w\|_{L_2}^2 = \sum_{k=n+1}^{\infty} |\alpha_k^{(n)}|^2 = \left\| \sum_{k=n+1}^{\infty} \alpha_k^{(n)} \psi_k^{(n)} \right\|_{\mathcal{H}^0}^2 = \|h_n^w - h^w\|_{\mathcal{H}}^2. \tag{4.7}
$$

Because of this relation and thanks to the assumption that $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^{\mu}$, there exists $n_\varepsilon \in \mathbb{N}$ such that $\max_{1\le j\le J_\varepsilon} \|Q_n^{\perp} b_j\|_{L_2} < \frac{\varepsilon}{2\|T_a\|_{\mathcal{L}(L_2)} \sqrt{J_\varepsilon}}$ holds for every $n \ge n_\varepsilon$; cf. (2.12).

Therefore, for all $n \geq n_\varepsilon$, we obtain that

$$\|Q_n^\perp P_{J_\varepsilon} w\|_{L_2} < \frac{\varepsilon}{2\|T_a\|_{\mathcal{L}(L_2)}\sqrt{J_\varepsilon}} \sum_{j=1}^{J_\varepsilon} |(w, b_j)_{L_2}| \leq \frac{\varepsilon}{2\|T_a\|_{\mathcal{L}(L_2)}} \|P_{J_\varepsilon} w\|_{L_2} \quad \forall w \in L_2.$$

The norm identities $\|P_{J_\varepsilon}\|_{\mathcal{L}(L_2)} = \|Q_n^\perp\|_{\mathcal{L}(L_2)} = 1$ thus imply that, for every $n \geq n_\varepsilon$, and for all $w \in L_2$,

$$\|Q_n^\perp P_{J_\varepsilon} T_a P_{J_\varepsilon} Q_n^\perp w\|_{L_2} < \frac{\varepsilon}{2\|T_a\|_{\mathcal{L}(L_2)}} \|P_{J_\varepsilon} T_a P_{J_\varepsilon} Q_n^\perp w\|_{L_2} \leq \frac{\varepsilon}{2}\|w\|_{L_2} \tag{4.8}$$

holds. Combining (4.5), (4.6) and (4.8) shows that $\sup_{h \in \mathcal{H}_{-n}} \frac{(\mathrm{A})}{\mathsf{Var}[h_n - h]} < \varepsilon$ for every $n \geq n_\varepsilon$ and, since $\varepsilon \in (0, \infty)$ was arbitrary,

$$\lim_{n\to\infty} \sup_{h \in \mathcal{H}_{-n}} \frac{(\mathrm{A})}{\mathsf{Var}[h_n - h]} = \lim_{n\to\infty} \sup_{h \in \mathcal{H}_{-n}} \left| \frac{\widetilde{\mathsf{Var}}[h_n - h]}{\mathsf{Var}[h_n - h]} - a \right| = 0,$$

and (3.9) follows. Furthermore, $\widetilde{\mathcal{C}}^{-1/2}\mathcal{C}\widetilde{\mathcal{C}}^{-1/2} - a^{-1}\mathcal{I}$ is compact on L_2 by Lemma B.1 (see Appendix B in the Supplementary Material [9]) and $\mathcal{S}_{\mathrm{adm}}^{\mu} = \mathcal{S}_{\mathrm{adm}}^{\widetilde{\mu}}$ by Lemma 4.3 so that, after changing the roles of the measures μ and $\widetilde{\mu}$, (3.9) implies (3.10).

Next, we show validity of (3.13) under Assumptions 3.3.I–III. To this end, we first split $|\widetilde{\mathsf{E}}[(h_n - h)^2] - a\mathsf{E}[(h_n - h)^2]| \leq (\mathrm{A}) + (\mathrm{B})$ in term (A), which is defined as above, and term $(\mathrm{B}) := |\widetilde{\mathsf{E}}[h_n - h]|^2$. By the Cauchy–Schwarz inequality,

$$(\mathrm{B}) = \left| \sum_{k=n+1}^{\infty} c_k^{(n)} \widetilde{\mathsf{E}}[\psi_k^{(n)}] \right|^2 \leq \mathsf{E}[(h_n - h)^2] \sum_{k=n+1}^{\infty} |\widetilde{\mathsf{E}}[\psi_k^{(n)}]|^2.$$

For each $k \geq n + 1$, we let $\{\psi_{kj}^{(n)}\}_{j \in \mathbb{N}}$ be the coefficients of $\psi_k^{(n)}$ when represented with respect to the orthonormal basis $\{z_j\}_{j \in \mathbb{N}}$ from Lemma 4.1(ii). We then find (recall that we have centered μ so that $\widetilde{\mu}$ has mean $\widetilde{m} - m$):

$$\begin{aligned}
\sum_{k=n+1}^{\infty} |\widetilde{\mathsf{E}}[\psi_k^{(n)}]|^2 &= \sum_{k=n+1}^{\infty} \left| \sum_{j \in \mathbb{N}} \psi_{kj}^{(n)} \widetilde{\mathsf{E}}[z_j] \right|^2 \\
&= \sum_{k=n+1}^{\infty} \left| \sum_{j \in \mathbb{N}} \psi_{kj}^{(n)} \widetilde{\mathsf{E}}[(Z^0, v_j)_{L_2}] \right|^2 \\
&= \sum_{k=n+1}^{\infty} \left| \sum_{j \in \mathbb{N}} \psi_{kj}^{(n)} (\widetilde{m} - m, \mathcal{C}^{-1/2} e_j)_{L_2} \right|^2 \\
&= \sum_{k=n+1}^{\infty} \left(\mathcal{C}^{-1/2}(\widetilde{m} - m), w_k^{(n)}\right)_{L_2}^2,
\end{aligned}$$

since $w_k^{(n)} = \mathcal{C}^{1/2} v_k^{(n)} = \mathcal{C}^{1/2}\mathcal{J}^{-1}\psi_k^{(n)} = \sum_{j \in \mathbb{N}} \psi_{kj}^{(n)} e_j$ and this series converges in L_2. Therefore, $\sum_{k=n+1}^{\infty} |\widetilde{\mathsf{E}}[\psi_k^{(n)}]|^2 = \|Q_n^\perp \mathcal{C}^{-1/2}(m - \widetilde{m})\|_{L_2}^2$ follows. By Assumption 3.3.II, the difference of the means $m - \widetilde{m}$ is an element of the Cameron–Martin space $H^* = \mathcal{C}^{1/2}(L_2)$. Consequently, $\mathcal{C}^{-1/2}(m - \widetilde{m}) \in L_2$ and the norm on the right-hand side converges to zero as $n \to \infty$ by (4.7) and (2.11). This shows that also

$$\lim_{n\to\infty} \sup_{h \in \mathcal{H}_{-n}} \frac{(\mathrm{B})}{\mathsf{E}[(h_n - h)^2]} = \lim_{n\to\infty} \sup_{h \in \mathcal{H}_{-n}} \frac{|\widetilde{\mathsf{E}}[h_n - h]|^2}{\mathsf{E}[(h_n - h)^2]} = 0. \tag{4.9}$$

We thus conclude with (3.9) and (4.9) that, uniformly in h,

$$\left|\frac{\widetilde{\mathsf{E}}[(h_n-h)^2]}{\mathsf{E}[(h_n-h)^2]}-a\right| \leq \left|\frac{\widetilde{\mathsf{Var}}[h_n-h]}{\mathsf{Var}[h_n-h]}-a\right| + \frac{|\widetilde{\mathsf{E}}[h_n-h]|^2}{\mathsf{E}[(h_n-h)^2]} \to 0 \quad \text{as } n\to\infty,$$

and (3.13) follows. Again, by virtue of Lemma 4.3 and Lemma B.1 (see Appendix B in the Supplementary Material [9]) we may change the roles of μ and $\widetilde{\mu}$ which gives (3.14).

To derive (3.11), note that $\mathsf{E}[(h_n-h)^2] \leq \mathsf{E}[(\widetilde{h}_n-h)^2]$ as h_n is the μ-best linear predictor. For the same reason, we obtain $\widetilde{\mathsf{E}}[(\widetilde{h}_n-h)^2] \leq \widetilde{\mathsf{E}}[(h_n-h)^2]$, and the estimates $1 \leq \frac{\mathsf{E}[(\widetilde{h}_n-h)^2]}{\mathsf{E}[(h_n-h)^2]} \leq \frac{\mathsf{E}[(\widetilde{h}_n-h)^2]}{\widetilde{\mathsf{E}}[(\widetilde{h}_n-h)^2]}\frac{\widetilde{\mathsf{E}}[(h_n-h)^2]}{\mathsf{E}[(h_n-h)^2]}$ follow similarly as in (4.3). By (3.14) and (3.13), the last two fractions converge to a^{-1} and to a, uniformly in h, as $n\to\infty$ and (3.11) follows. Changing the roles of μ and $\widetilde{\mu}$ implies (3.12).

Finally, note that if $\mu = \mathsf{N}(m,\mathcal{C})$ and $\widetilde{\mu} = \mathsf{N}(\widetilde{m},\widetilde{\mathcal{C}})$ are such that Assumptions 3.3.I and 3.3.III are satisfied, then the centered measures $\mu_{\mathrm{c}} = \mathsf{N}(0,\mathcal{C})$ and $\widetilde{\mu}_{\mathrm{c}} = \mathsf{N}(0,\widetilde{\mathcal{C}})$ satisfy Assumptions 3.3.I–III so that (3.11), (3.12) hold for the pair μ_{c}, $\widetilde{\mu}_{\mathrm{c}}$ and (3.7), (3.8) follow from the identities in (C.2) and (C.3); see Appendix C in the Supplementary Material [9]. □

PROOF OF THEOREM 3.1 AND COROLLARY 3.2. If the measures μ and $\widetilde{\mu}$ are equivalent, then by the Feldman–Hájek theorem (see Theorem A.1 in Appendix A of the Supplementary Material [9]) Assumptions 3.3.I–II hold and $T_1 = \mathcal{C}^{-1/2}\widetilde{\mathcal{C}}\mathcal{C}^{-1/2} - \mathcal{I} \in \mathcal{L}_2(L_2)$. Since every Hilbert–Schmidt operator is compact, this implies that also Assumption 3.3.III is fulfilled for $a=1$. Therefore, for every $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}^{\mu}_{\mathrm{adm}}$, all assertions in (3.1), (3.2), (3.3), (3.4) follow from Proposition 4.4. □

LEMMA 4.5. *Let $\mu = \mathsf{N}(m,\mathcal{C})$, $\widetilde{\mu} = \mathsf{N}(\widetilde{m},\widetilde{\mathcal{C}})$. In (3.7)–(3.10), let h_n, $\widetilde{h}_n$ denote the best linear predictors of h based on $\mathcal{H}_n$ and the measures μ, $\widetilde{\mu}$. Then validity of any of the statements (3.7), (3.8), (3.9) or (3.10) for all $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}^{\mu}_{\mathrm{adm}}$ implies that the Assumptions 3.3.I and 3.3.III are satisfied, and the constant $a\in(0,\infty)$ in (3.9), (3.10) is the same as in (3.5).*

PROOF. By (C.2)–(C.5) (see Appendix C in the Supplementary Material [9]), we can without loss of generality assume that $m=\widetilde{m}=0$. Then $\mathsf{Var}[\widetilde{h}_n-h] = \mathsf{E}[(\widetilde{h}_n-h)^2]$ and $\widetilde{\mathsf{Var}}[h_n-h] = \widetilde{\mathsf{E}}[(h_n-h)^2]$ follow. Furthermore, $\mathsf{Var}[h_n-h] = \mathsf{E}[(h_n-h)^2]$ and $\widetilde{\mathsf{Var}}[\widetilde{h}_n-h] = \widetilde{\mathsf{E}}[(\widetilde{h}_n-h)^2]$ always hold by unbiasedness of the kriging predictor. Recall from Lemma 4.1 the orthonormal bases $\{e_j\}_{j\in\mathbb{N}}$ for L_2, $\{v_j\}_{j\in\mathbb{N}}$ for $H = \mathcal{C}^{-1/2}(L_2)$, and $\{z_j\}_{j\in\mathbb{N}}$ for $\mathcal{H}^0$ as well as the isometry $\mathcal{J}\colon H\to\mathcal{H}^0$, which identifies v_j with z_j.

If any of the statements (3.7), (3.8), (3.9) or (3.10) holds for every $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}^{\mu}_{\mathrm{adm}}$, then by Lemma B.4 (see Appendix B in the Supplementary Material [9]), all four assertions of Proposition 3.5 and, in particular, part (i) hold, that is, Assumption 3.3.I is satisfied.

Next, we prove that validity of (3.8) for all $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}^{\mu}_{\mathrm{adm}}$ implies Assumption 3.3.III. For $n\in\mathbb{N}$, define $E_n := \operatorname{span}\{e_1,\ldots,e_n\} \subset L_2$ and $H_n := \operatorname{span}\{v_1,\ldots,v_n\} \subset H$, and let $E_n^{\perp} = \operatorname{span}\{e_j\}_{j>n}$ as well as $H_n^{\perp} = \operatorname{span}\{v_j\}_{j>n}$ be their orthogonal complements in L_2 and H, respectively. Note that $E_n = H_n$ and $E_n^{\perp} \subset H_n^{\perp}$. Now suppose that, for all $a\in(0,\infty)$, the linear operator $T_a = \mathcal{C}^{-1/2}\widetilde{\mathcal{C}}\mathcal{C}^{-1/2} - a\mathcal{I}$ is not compact on L_2, and define $\underline{\alpha} := \|\mathcal{C}^{1/2}\widetilde{\mathcal{C}}^{-1/2}\|^{-2}_{\mathcal{L}(L_2)}$, $\overline{\alpha} := \|\widetilde{\mathcal{C}}^{1/2}\mathcal{C}^{-1/2}\|^{2}_{\mathcal{L}(L_2)}$. Then, by Lemma B.2 (see Appendix B in the Supplementary Material [9]) there exist $\delta\in(0,\infty)$ and, for every $n\in\mathbb{N}$, $\underline{a}_n, \overline{a}_n \in [\underline{\alpha},\overline{\alpha}]$ and $\underline{w}_n, \overline{w}_n \in E_n^{\perp}\setminus\{0\}$ such that, for all $n\in\mathbb{N}$, we have $\overline{a}_n - \underline{a}_n \geq \delta$ and

$$\left|\frac{(\mathcal{C}^{-1/2}\widetilde{\mathcal{C}}\mathcal{C}^{-1/2}\underline{w}_n,\underline{w}_n)_{L_2}}{(\underline{w}_n,\underline{w}_n)_{L_2}} - \underline{a}_n\right| < \frac{\delta\underline{\alpha}^2}{3\overline{\alpha}^2}, \qquad \left|\frac{(\mathcal{C}^{-1/2}\widetilde{\mathcal{C}}\mathcal{C}^{-1/2}\overline{w}_n,\overline{w}_n)_{L_2}}{(\overline{w}_n,\overline{w}_n)_{L_2}} - \overline{a}_n\right| < \frac{\delta\underline{\alpha}^2}{3\overline{\alpha}^2}.$$

We set $\underline{c}_n := \overline{a}_n^{-1}$, $\overline{c}_n := \underline{a}_n^{-1}$, and $\underline{v}_n := \mathcal{C}^{-1/2}\overline{w}_n$, $\overline{v}_n := \mathcal{C}^{-1/2}\underline{w}_n$. Then we obtain that, for all $n\in\mathbb{N}$, $\underline{c}_n, \overline{c}_n \in [\overline{\alpha}^{-1},\underline{\alpha}^{-1}]$, and $\overline{c}_n - \underline{c}_n \geq \delta' := \delta\overline{\alpha}^{-2}$. The vectors $\underline{v}_n, \overline{v}_n \in H_n^{\perp}$ satisfy

$\left|\frac{\|\mathcal{C}^{1/2}\underline{v}_n\|_{L_2}^2}{\|\widetilde{\mathcal{C}}^{1/2}\underline{v}_n\|_{L_2}^2} - \underline{c}_n\right| < \frac{\delta'}{3}$ and $\left|\frac{\|\mathcal{C}^{1/2}\overline{v}_n\|_{L_2}^2}{\|\widetilde{\mathcal{C}}^{1/2}\overline{v}_n\|_{L_2}^2} - \overline{c}_n\right| < \frac{\delta'}{3}$. We then define $\underline{\phi}_n := \mathcal{J}\underline{v}_n \in \mathcal{H}^0$ as well as $\bar{\phi}_n := \mathcal{J}\overline{v}_n \in \mathcal{H}^0$ and find that

$$\frac{\mathsf{E}[\underline{\phi}_n^2]}{\widetilde{\mathsf{E}}[\underline{\phi}_n^2]} = \frac{\|\mathcal{C}^{1/2}\underline{v}_n\|_{L_2}^2}{\|\widetilde{\mathcal{C}}^{1/2}\underline{v}_n\|_{L_2}^2} \in \left(\underline{c}_n - \frac{\delta'}{3}, \underline{c}_n + \frac{\delta'}{3}\right),$$
$$\frac{\mathsf{E}[\bar{\phi}_n^2]}{\widetilde{\mathsf{E}}[\bar{\phi}_n^2]} = \frac{\|\mathcal{C}^{1/2}\overline{v}_n\|_{L_2}^2}{\|\widetilde{\mathcal{C}}^{1/2}\overline{v}_n\|_{L_2}^2} \in \left(\overline{c}_n - \frac{\delta'}{3}, \overline{c}_n + \frac{\delta'}{3}\right).$$

As in [5], Proof of Theorem 5, it follows that there exist $h^{(n)}, \psi_n \in \operatorname{span}\{\underline{\phi}_n, \bar{\phi}_n\}$ such that

$$\frac{\widetilde{\mathsf{E}}[(h_1^{(n)} - h^{(n)})^2]}{\widetilde{\mathsf{E}}[(\widetilde{h}_1^{(n)} - h^{(n)})^2]} = \frac{(\widetilde{\theta}_n + \widetilde{\Theta}_n)^2}{4\widetilde{\theta}_n\widetilde{\Theta}_n} \tag{4.10}$$

holds, where $h_1^{(n)}$ and $\widetilde{h}_1^{(n)}$ are the best linear predictors of $h^{(n)}$ based on the subspace $\mathcal{V}_n := \mathbb{R} \oplus \operatorname{span}\{\psi_n\} \subset \mathcal{H}$ and the measures μ and $\widetilde{\mu}$, respectively. Moreover,

$$\widetilde{\theta}_n := \min\{\mathsf{E}[h^2]/\widetilde{\mathsf{E}}[h^2] : h \in \operatorname{span}\{\underline{\phi}_n, \bar{\phi}_n\}, h \neq 0\},$$

and $\widetilde{\Theta}_n \in (0, \infty)$ is defined as $\widetilde{\theta}_n$ with min replaced by max. Clearly, these definitions yield that $\widetilde{\theta}_n \le \mathsf{E}[\underline{\phi}_n^2]/\widetilde{\mathsf{E}}[\underline{\phi}_n^2] < \underline{c}_n + \frac{\delta'}{3}$ and $\widetilde{\Theta}_n \ge \mathsf{E}[\bar{\phi}_n^2]/\widetilde{\mathsf{E}}[\bar{\phi}_n^2] > \overline{c}_n - \frac{\delta'}{3}$, which implies that $\widetilde{\Theta}_n - \widetilde{\theta}_n > \overline{c}_n - \underline{c}_n - \frac{2\delta'}{3} = \frac{\delta'}{3}$. As we have already derived fulfillment of Assumption 3.3.I, Proposition 3.5(i) $\Leftrightarrow$ (ii) shows that $\widetilde{\Theta}_n \le \sup_{h\in\mathcal{H}^0\setminus\{0\}} \frac{\mathsf{E}[h^2]}{\widetilde{\mathsf{E}}[h^2]} \le \|\mathcal{C}^{1/2}\widetilde{\mathcal{C}}^{-1/2}\|_{\mathcal{L}(L_2)}^2 < \infty$. Define $\mathcal{H}_1^\star := \mathcal{V}_1$ and, for $n \ge 2$, set $\mathcal{H}_n^\star := \mathbb{R} \oplus \operatorname{span}\{z_1, \ldots, z_{n-1}, \psi_n\}$. By the basis property of $\{z_j\}_{j\in\mathbb{N}}$ in $\mathcal{H}^0$ (see Lemma 4.1(ii)), the so constructed subspaces are admissible, that is, $\{\mathcal{H}_n^\star\}_{n\in\mathbb{N}} \in \mathcal{S}_{\text{adm}}^\mu$. Since $h^{(n)}, \psi_n \in \operatorname{span}\{\underline{\phi}_n, \bar{\phi}_n\}$ and since $\underline{\phi}_n, \bar{\phi}_n$ are $\mathcal{H}^0$-orthogonal to $z_1, \ldots, z_{n-1}$, we obtain $h_n^{(n)} = h_1^{(n)}$, where $h_n^{(n)}$ is the best linear predictor of $h^{(n)}$ based on $\mathcal{H}_n^\star$ and μ. Thus, by using (4.10) we obtain, for all $n \in \mathbb{N}$,

$$\frac{\widetilde{\mathsf{E}}[(h_n^{(n)} - h^{(n)})^2]}{\widetilde{\mathsf{E}}[(\widetilde{h}_n^{(n)} - h^{(n)})^2]} - 1 \ge \frac{(\widetilde{\Theta}_n - \widetilde{\theta}_n)^2}{4\widetilde{\theta}_n\widetilde{\Theta}_n} > \frac{\delta'^2}{36\widetilde{\Theta}_n^2} \ge \frac{\delta'^2}{36\|\mathcal{C}^{1/2}\widetilde{\mathcal{C}}^{-1/2}\|_{\mathcal{L}(L_2)}^4},$$

a contradiction to (3.8) for the sequence $\{\mathcal{H}_n^\star\}_{n\in\mathbb{N}} \in \mathcal{S}_{\text{adm}}^\mu$, which proves that (3.8) holding for all $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\text{adm}}^\mu$ implies Assumption 3.3.III.

Next, we show that validity of (3.9) for all $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\text{adm}}^\mu$ also implies that Assumption 3.3.III is satisfied. To this end, suppose that this assumption does not hold. It then again follows from Lemma B.2 (see Appendix B in the Supplementary Material [9]) that there are $\delta \in (0, \infty)$ and, for all $n \in \mathbb{N}$, $\underline{a}_n, \overline{a}_n \in [\underline{\alpha}, \overline{\alpha}]$ and $\underline{v}_n, \overline{v}_n \in H_n^\perp$, linearly independent, such that $\left|\frac{\|\widetilde{\mathcal{C}}^{1/2}\underline{v}_n\|_{L_2}^2}{\|\mathcal{C}^{1/2}\underline{v}_n\|_{L_2}^2} - \underline{a}_n\right| < \frac{\delta}{3}$, $\left|\frac{\|\widetilde{\mathcal{C}}^{1/2}\overline{v}_n\|_{L_2}^2}{\|\mathcal{C}^{1/2}\overline{v}_n\|_{L_2}^2} - \overline{a}_n\right| < \frac{\delta}{3}$, and $\overline{a}_n - \underline{a}_n \ge \delta$ for all $n \in \mathbb{N}$. Define $\bar{h}^{(n)} := \mathcal{J}\overline{v}_n \in \mathcal{H}^0$ and $\underline{h}^{(n)} := \mathcal{J}\underline{v}_n \in \mathcal{H}^0$. Then, for $\mathcal{H}_n^\star := \mathbb{R} \oplus \operatorname{span}\{z_1, \ldots, z_n\} \subset \mathcal{H}$,

$$\forall n \in \mathbb{N}: \quad \frac{\widetilde{\mathsf{Var}}[\bar{h}_n^{(n)} - \bar{h}^{(n)}]}{\mathsf{Var}[\bar{h}_n^{(n)} - \bar{h}^{(n)}]} - \frac{\widetilde{\mathsf{Var}}[\underline{h}_n^{(n)} - \underline{h}^{(n)}]}{\mathsf{Var}[\underline{h}_n^{(n)} - \underline{h}^{(n)}]} = \frac{\widetilde{\mathsf{Var}}[\bar{h}^{(n)}]}{\mathsf{Var}[\bar{h}^{(n)}]} - \frac{\widetilde{\mathsf{Var}}[\underline{h}^{(n)}]}{\mathsf{Var}[\underline{h}^{(n)}]} \ge \frac{\delta}{3}. \tag{4.11}$$

Note that, if (3.9) holds for all sequences $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\text{adm}}^\mu$, then, in particular,

$$\forall \{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\text{adm}}^\mu: \quad \lim_{n\to\infty}\left(\sup_{g\in\mathcal{H}_{-n}} \frac{\widetilde{\mathsf{Var}}[g_n - g]}{\mathsf{Var}[g_n - g]} - \inf_{h\in\mathcal{H}_{-n}} \frac{\widetilde{\mathsf{Var}}[h_n - h]}{\mathsf{Var}[h_n - h]}\right) = 0$$

follows. Therefore, (4.11) contradicts (3.9) for the sequence $\{\mathcal{H}_n^\star\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^\mu$, and thus, Assumption 3.3.III is satisfied if (3.9) holds for all $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^\mu$.

Finally, necessity of Assumption 3.3.III for validity of (3.7) (or (3.10)) holding for all $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^\mu$ follows from changing the roles of μ and $\widetilde{\mu}$: If μ and $\widetilde{\mu}$ are such that (3.7) (or (3.10)) holds for every $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^\mu$, then (3.8) (or (3.9)) is true for the pair $\widetilde{\mu}$, μ and every $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^\mu$. Since necessity of Assumption 3.3.I has already been derived, by Lemma 4.3 we have $\mathcal{S}_{\mathrm{adm}}^\mu = \mathcal{S}_{\mathrm{adm}}^{\widetilde{\mu}}$ so that the above arguments combined with Lemma B.1 in Appendix B of the Supplementary Material [9] show that Assumption 3.3.III also holds. □

PROOF OF THEOREM 3.6. Sufficiency and necessity of Assumptions 3.3.I and 3.3.III for (3.7)–(3.10) to hold for all $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^\mu$ have been proven in Proposition 4.4 and Lemma 4.5, respectively. □

LEMMA 4.6. *Define the Gaussian measures* $\mu_{\mathrm{c}} = \mathsf{N}(0, \mathcal{C})$ *and* $\mu_{\mathrm{s}} = \mathsf{N}(\widetilde{m} - m, \mathcal{C})$, *with corresponding expectation operators* E_{c} *and* E_{s}. *Let* h_n^{c} *and* h_n^{s} *denote the best linear predictors of* h *based on* $\mathcal{H}_n \in \{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^\mu$ *and the measures* μ_{c} *and* μ_{s}, *respectively. For* $n \in \mathbb{N}$ *and* $h \in \mathcal{H}_{-n}$, *consider the errors of the predictors,* $e_{\mathrm{c}} = e_{\mathrm{c}}(h, n) := h_n^{\mathrm{c}} - h$ *and* $e_{\mathrm{s}} = e_{\mathrm{s}}(h, n) := h_n^{\mathrm{s}} - h$. *Then*

$$\frac{\mathsf{E}_{\mathrm{c}}[e_{\mathrm{s}}^2]}{\mathsf{E}_{\mathrm{c}}[e_{\mathrm{c}}^2]} - 1 = \frac{\mathsf{E}_{\mathrm{s}}[e_{\mathrm{c}}^2]}{\mathsf{E}_{\mathrm{s}}[e_{\mathrm{s}}^2]} - 1 = \left|\frac{\mathsf{E}_{\mathrm{s}}[e_{\mathrm{c}}^2]}{\mathsf{E}_{\mathrm{c}}[e_{\mathrm{c}}^2]} - 1\right| = \left|\frac{\mathsf{E}_{\mathrm{c}}[e_{\mathrm{s}}^2]}{\mathsf{E}_{\mathrm{s}}[e_{\mathrm{s}}^2]} - 1\right| = \frac{|\mathsf{E}_{\mathrm{s}}[e_{\mathrm{c}}]|^2}{\mathsf{E}_{\mathrm{c}}[e_{\mathrm{c}}^2]}. \tag{4.12}$$

Furthermore, for all $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^\mu$, *this term is bounded, uniformly with respect to* $n \in \mathbb{N}$ *and* $h \in \mathcal{H}_{-n}$, *if and only if Assumption* 3.3.II *is satisfied. Under Assumption* 3.3.II, $\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \frac{|\mathsf{E}_{\mathrm{s}}[e_{\mathrm{c}}(h,n)]|^2}{\mathsf{E}_{\mathrm{c}}[e_{\mathrm{c}}(h,n)^2]} = 0$ *holds for all* $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^\mu$.

PROOF. Let $n \in \mathbb{N}$ and $h \in \mathcal{H}_{-n}$. By $\mathcal{H}$-orthogonality of $e_{\mathrm{c}} = h_n^{\mathrm{c}} - h$ to $\mathcal{H}_n$, we obtain

$$\mathsf{E}_{\mathrm{c}}\big[e_{\mathrm{s}}^2\big] - \mathsf{E}_{\mathrm{c}}\big[e_{\mathrm{c}}^2\big] = \mathsf{E}_{\mathrm{c}}[e_{\mathrm{s}}e_{\mathrm{c}}] + \mathsf{E}_{\mathrm{c}}\big[e_{\mathrm{s}}\big(h_n^{\mathrm{s}} - h_n^{\mathrm{c}}\big)\big] - \mathsf{E}_{\mathrm{c}}[e_{\mathrm{c}}e_{\mathrm{s}}] = \mathsf{E}_{\mathrm{c}}\big[\big(h_n^{\mathrm{s}} - h_n^{\mathrm{c}}\big)^2\big].$$

Since μ_{c} and μ_{s} have the same covariance operator, we can combine the above equality with (C.1) from Lemma C.1 (see Appendix C in the Supplementary Material [9]), which gives $\frac{\mathsf{E}_{\mathrm{c}}[e_{\mathrm{s}}^2]}{\mathsf{E}_{\mathrm{c}}[e_{\mathrm{c}}^2]} - 1 = \frac{\mathsf{E}_{\mathrm{c}}[(h_n^{\mathrm{s}} - h_n^{\mathrm{c}})^2]}{\mathsf{E}_{\mathrm{c}}[e_{\mathrm{c}}^2]} = \frac{|\mathsf{E}_{\mathrm{s}}[e_{\mathrm{c}}]|^2}{\mathsf{E}_{\mathrm{c}}[e_{\mathrm{c}}^2]}$. Noting that $|\mathsf{E}_{\mathrm{c}}[e_{\mathrm{s}}]| = |\mathsf{E}_{\mathrm{s}}[e_{\mathrm{c}}]|$ and $\mathsf{E}_{\mathrm{s}}[e_{\mathrm{s}}^2] = \mathsf{E}_{\mathrm{c}}[e_{\mathrm{c}}^2]$ due to the identical covariance operators of μ_{c}, μ_{s} yields the relation $\frac{\mathsf{E}_{\mathrm{s}}[e_{\mathrm{c}}^2]}{\mathsf{E}_{\mathrm{s}}[e_{\mathrm{s}}^2]} - 1 = \frac{|\mathsf{E}_{\mathrm{s}}[e_{\mathrm{c}}]|^2}{\mathsf{E}_{\mathrm{c}}[e_{\mathrm{c}}^2]}$. Next, again by equality of the covariance operators, we find that $\mathsf{E}_{\mathrm{s}}[e_{\mathrm{c}}^2] - \mathsf{E}_{\mathrm{c}}[e_{\mathrm{c}}^2] = |\mathsf{E}_{\mathrm{s}}[e_{\mathrm{c}}]|^2$ and $\mathsf{E}_{\mathrm{c}}[e_{\mathrm{s}}^2] - \mathsf{E}_{\mathrm{s}}[e_{\mathrm{s}}^2] = |\mathsf{E}_{\mathrm{c}}[e_{\mathrm{s}}]|^2 = |\mathsf{E}_{\mathrm{s}}[e_{\mathrm{c}}]|^2$, which completes the proof of (4.12).

Now suppose that Assumption 3.3.II is satisfied and let $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^\mu$. Then we obtain $\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \frac{|\mathsf{E}_{\mathrm{s}}[e_{\mathrm{c}}(h,n)]|^2}{\mathsf{E}_{\mathrm{c}}[e_{\mathrm{c}}(h,n)^2]} = 0$ as in (4.9) with $\mathcal{C} = \widetilde{\mathcal{C}}$. In particular, there exists a constant $K \in (0, \infty)$ such that $\sup_{n\in\mathbb{N}} \sup_{h\in\mathcal{H}_{-n}} \frac{|\mathsf{E}_{\mathrm{s}}[e_{\mathrm{c}}(h,n)]|^2}{\mathsf{E}_{\mathrm{c}}[e_{\mathrm{c}}(h,n)^2]} \le K$.

Finally, assume that $m - \widetilde{m} \notin H^* = \mathcal{C}^{1/2}(L_2)$, that is, Assumption 3.3.II is not satisfied. For $n \in \mathbb{N}$, define $H_n = \operatorname{span}\{v_1, \dots, v_n\}$, where $\{v_j\}_{j\in\mathbb{N}}$ is the orthonormal basis of H from Lemma 4.1(i), and let $H_n^\perp$ be the H-orthogonal complement of H_n. Since $m - \widetilde{m} \notin H^*$ and L_2 is dense in H, we can find $\{\overline{v}_n\}_{n\in\mathbb{N}} \subset L_2 \setminus \{0\}$ such that $(m - \widetilde{m}, \overline{v}_n)_{L_2} \ge n\|\overline{v}_n\|_H$. Furthermore, we may pick $\overline{v}_n$ in $H_n^\perp \subset H$, since $\dim(H_n) < \infty$. In summary,

$$\forall n \in \mathbb{N} \quad \exists \overline{v}_n \in L_2 \cap H_n^\perp, \overline{v}_n \ne 0: \quad (m - \widetilde{m}, \overline{v}_n)_{L_2} \ge n\|\mathcal{C}^{1/2}\overline{v}_n\|_{L_2}.$$

By (4.2), $h^{(n)} := \mathcal{J}\overline{v}_n \in \mathcal{H}^0$ is $\mathcal{H}^0$-orthogonal to $\mathcal{H}_n^0 := \operatorname{span}\{z_1, \dots, z_n\}$ if $\{z_j\}_{j\in\mathbb{N}}$ is the orthonormal basis for $\mathcal{H}^0$ from Lemma 4.1(ii). Therefore, the kriging predictor of $h^{(n)}$ based on $\mathcal{H}_n^\star := \mathbb{R} \oplus \mathcal{H}_n^0$ and $\mu_{\mathrm{c}} = \mathsf{N}(0, \mathcal{C})$ vanishes, $h_n^{(n),c} = 0$. Thus, there exist square-summable coefficients $\{c_j^{(n)}\}_{j>n}$ such that $h^{(n)} = \sum_{j>n} c_j^{(n)} z_j$ and $\overline{v}_n = \sum_{j>n} c_j^{(n)} v_j$, and we find that

$$\frac{|\mathsf{E}_{\mathrm{s}}[h_n^{(n),c} - h^{(n)}]|^2}{\mathsf{E}_{\mathrm{c}}[(h_n^{(n),c} - h^{(n)})^2]} = \frac{|\mathsf{E}_{\mathrm{s}}[h^{(n)}]|^2}{\mathsf{E}_{\mathrm{c}}[|h^{(n)}|^2]} = \frac{|\mathsf{E}_{\mathrm{s}}[\sum_{j>n} c_j^{(n)}(Z^0, v_j)_{L_2}]|^2}{\|\mathcal{C}^{1/2}\overline{v}_n\|_{L_2}^2} = \frac{(\widetilde{m} - m, \overline{v}_n)_{L_2}^2}{\|\mathcal{C}^{1/2}\overline{v}_n\|_{L_2}^2} \geq n^2.$$

Furthermore, $\{\mathcal{H}_n^\star\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^\mu$ so that this yields a contradiction. □

PROOF OF THEOREM 3.8. In this proof, all references starting with "C" are referring to Appendix C in the Supplementary Material [9]. As shown in (C.2)–(C.5) (see Proposition C.2), we can equivalently prove the claim for the pair of measures $\mu_{\mathrm{c}} = \mathsf{N}(0, \mathcal{C})$ and $\widetilde{\mu}_{\mathrm{s}} = \mathsf{N}(\widetilde{m} - m, \widetilde{\mathcal{C}})$ in place of $\mu = \mathsf{N}(m, \mathcal{C})$ and $\widetilde{\mu} = \mathsf{N}(\widetilde{m}, \widetilde{\mathcal{C}})$. Sufficiency of Assumptions 3.3.I–III for each of the assertions (3.11)–(3.14) to hold for all $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^\mu$ is shown in Proposition 4.4.

Conversely, if (3.11) (or (3.12)) holds for μ_{c}, $\widetilde{\mu}_{\mathrm{s}}$ and all $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^\mu$, then by (C.7) the relation (3.7) (or (3.8)) holds for the pair μ, $\widetilde{\mu}$ and all $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^\mu$. By Lemma 4.5, Assumptions 3.3.I and 3.3.III have to be satisfied. Subsequently, necessity of Assumption 3.3.II for (3.12) follows from (C.8) combined with Lemma 4.6. Since we have already derived Assumption 3.3.I, we have $\mathcal{S}_{\mathrm{adm}}^\mu = \mathcal{S}_{\mathrm{adm}}^{\widetilde{\mu}}$ by Lemma 4.3, and we may also combine (C.8) with Lemma 4.6 applied for $\widetilde{\mu}_{\mathrm{c}}$ and $\widetilde{\mu}_{\mathrm{s}}$, showing necessity of Assumption 3.3.II for (3.11).

If (3.13) holds for μ_{c}, $\widetilde{\mu}_{\mathrm{s}}$ and all $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^\mu$, then the first identity in (C.6) combined with Lemma 4.6 (see also Remark C.3) show that Assumption 3.3.II has to be satisfied. Subsequently, again the first identity in (C.6) implies that (3.9) holds for the pair μ, $\widetilde{\mu}$ and all $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}_{\mathrm{adm}}^\mu$ and Assumptions 3.3.I and 3.3.III follow from Lemma 4.5. For (3.14) we may proceed analogously, since uniform boundedness of the terms in the second identity of (C.6) in particular implies that $\mathcal{S}_{\mathrm{adm}}^{\widetilde{\mu}} \subseteq \mathcal{S}_{\mathrm{adm}}^\mu$, so that Lemma 4.6 is applicable for the pair of measures $\widetilde{\mu}_c$ and $\widetilde{\mu}_{\mathrm{s}}$. □

5. Simplified necessary and sufficient conditions. In order to exploit Theorem 3.8 to check if two models provide uniformly asymptotically equivalent linear predictions, one has to verify Assumptions 3.3.I–III. Depending on the form of the covariance operators, this may be difficult. In this section, we provide equivalent formulations of Assumptions 3.3.I and III for two important cases: 1. the two covariance operators diagonalize with respect to the same eigenbasis, and 2. $\varrho, \widetilde{\varrho}\colon \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ are covariance functions of weakly stationary random fields on $\mathcal{X} \subset \mathbb{R}^d$, a priori defined on all of $\mathbb{R}^d$, which have well-defined spectral densities $f, \widetilde{f}\colon \mathbb{R}^d \to [0, \infty)$.

5.1. *Common eigenbasis*. In the case that the two covariance operators diagonalize with respect to the same eigenbasis, conditions I and III of Assumption 3.3 can be formulated as conditions on the ratios of the eigenvalues. We consider this scenario in the next corollary.

COROLLARY 5.1. *Suppose that $\mathcal{C}$, $\widetilde{\mathcal{C}}$ are self-adjoint, positive definite, compact operators on $L_2(\mathcal{X}, \nu_{\mathcal{X}})$, which diagonalize with respect to the same orthonormal basis $\{e_j\}_{j\in\mathbb{N}}$ for $L_2(\mathcal{X}, \nu_{\mathcal{X}})$, that is, there exist corresponding eigenvalues $\gamma_j, \widetilde{\gamma}_j \in (0, \infty)$, $j \in \mathbb{N}$, accumulating only at zero such that $\mathcal{C}e_j = \gamma_j e_j$ and $\widetilde{\mathcal{C}}e_j = \widetilde{\gamma}_j e_j$ for all $j \in \mathbb{N}$. Then Assumptions* 3.3.I *and* 3.3.III *are satisfied if and only if there exists $a \in (0, \infty)$ such that $\lim_{j\to\infty} \widetilde{\gamma}_j/\gamma_j = a$.*

PROOF. We start by showing that $\lim_{j\to\infty} \widetilde{\gamma}_j/\gamma_j = a \in (0, \infty)$ is sufficient for Assumptions 3.3.I and 3.3.III. By Proposition 3.5, Assumption 3.3.I is equivalent to requiring that $\widetilde{\mathcal{C}}^{1/2}\mathcal{C}^{-1/2}$ is an isomorphism on L_2. If $\mathcal{C}$ and $\widetilde{\mathcal{C}}$ admit the same eigenbasis $\{e_j\}_{j\in\mathbb{N}}$, then these are also eigenvectors of the self-adjoint, positive definite linear operator $\mathcal{C}^{-1/2}\widetilde{\mathcal{C}}\mathcal{C}^{-1/2}$

with corresponding eigenvalues $\{\widetilde{\gamma}_j/\gamma_j\}_{j\in\mathbb{N}}$. By assumption, this sequence converges. Hence, $\|\widetilde{\mathcal{C}}^{1/2}\mathcal{C}^{-1/2}\|^2_{\mathcal{L}(L_2)} = \sup_{\|v\|_{L_2}=1}(\mathcal{C}^{-1/2}\widetilde{\mathcal{C}}\mathcal{C}^{-1/2}v, v)_{L_2} = \sup_{j\in\mathbb{N}} \widetilde{\gamma}_j/\gamma_j \in (0,\infty)$ follows, and $\lim_{j\to\infty} \gamma_j/\widetilde{\gamma}_j = 1/a$ implies that $\|\mathcal{C}^{1/2}\widetilde{\mathcal{C}}^{-1/2}\|^2_{\mathcal{L}(L_2)} = \sup_{j\in\mathbb{N}} \gamma_j/\widetilde{\gamma}_j \in (0,\infty)$ by the same argument. Thus, Assumption 3.3.I is satisfied. Furthermore, also Assumption 3.3.III follows, since $T_a = \mathcal{C}^{-1/2}\widetilde{\mathcal{C}}\mathcal{C}^{-1/2} - a\mathcal{I}$ diagonalizes with respect to $\{e_j\}_{j\in\mathbb{N}}$ with corresponding eigenvalues $\{\widetilde{\gamma}_j/\gamma_j - a\}_{j\in\mathbb{N}}$ which by assumption accumulate only at zero, and hence, T_a is compact on $L_2(\mathcal{X}, \nu_{\mathcal{X}})$. Conversely, if Assumptions 3.3.I and 3.3.III are satisfied, then by the latter there exists $a \in (0,\infty)$ such that T_a is compact and $\{\widetilde{\gamma}_j/\gamma_j - a\}_{j\in\mathbb{N}}$ is a null sequence, that is, $\{\widetilde{\gamma}_j/\gamma_j\}_{j\in\mathbb{N}}$ converges to $a \in (0,\infty)$. □

5.2. *Weakly stationary random fields.* We consider a connected, compact subset $\mathcal{X}$ of $\mathbb{R}^d$ equipped with the Euclidean metric and the Lebesgue measure λ_d. For brevity, we omit λ_d in the notation $L_2(\mathcal{X})$, $L_2(\mathbb{R}^d)$. We assume that the operators $\mathcal{C}, \widetilde{\mathcal{C}} \colon L_2(\mathcal{X}) \to L_2(\mathcal{X})$ are induced by continuous, (strictly) positive definite kernels $\varrho|_{\mathcal{X}\times\mathcal{X}}$ and $\widetilde{\varrho}|_{\mathcal{X}\times\mathcal{X}}$, which are restrictions of translation invariant covariance functions $\varrho, \widetilde{\varrho} \colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$. Translation invariance of ϱ implies that there exists an even function $\varrho_0 \colon \mathbb{R}^d \to \mathbb{R}$ such that $\varrho(x,x') = \varrho_0(x-x')$ for all $x, x' \in \mathbb{R}^d$, and similarly $\widetilde{\varrho}_0$ is defined for $\widetilde{\varrho}$. We assume $\varrho_0, \widetilde{\varrho}_0 \in L_1(\mathbb{R}^d)$, so that the corresponding spectral densities $f, \widetilde{f} \colon \mathbb{R}^d \to [0,\infty)$ exist. Recall that the spectral density f and ϱ_0 relate via the inversion formula (see, e.g., [25], page 25): For all $\omega \in \mathbb{R}^d$, we have

$$f(\omega) = \frac{1}{(2\pi)^d}(\mathcal{F}\varrho_0)(\omega), \qquad (\mathcal{F}\varrho_0)(\omega) := \int_{\mathbb{R}^d} \exp(-i\omega\cdot x)\varrho_0(x)\,\mathrm{d}x. \tag{5.1}$$

Using this convention for the Fourier transform $\mathcal{F}$, its inverse becomes

$$(\mathcal{F}^{-1}\hat{v})(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \exp(i\omega\cdot x)\hat{v}(\omega)\,\mathrm{d}\omega, \quad x \in \mathbb{R}^d.$$

Let the linear operator $\mathcal{F}_{\mathcal{X}} \colon L_2(\mathcal{X}) \to L_2(\mathbb{R}^d;\mathbb{C})$ be the composition $\mathcal{F}_{\mathcal{X}} := \mathcal{F} \circ E^0_{\mathcal{X}}$, where $E^0_{\mathcal{X}}$ is the zero extension $L_2(\mathcal{X}) \ni w \mapsto E^0_{\mathcal{X}} w \in L_1(\mathbb{R}^d) \cap L_2(\mathbb{R}^d)$ that sets $(E^0_{\mathcal{X}}w)(x) = 0$ for all $x \in \mathbb{R}^d \setminus \mathcal{X}$. We then consider the following subset of the space of complex-valued square-integrable functions $L_2(\mathbb{R}^d;\mathbb{C})$, which itself is a vector space over $\mathbb{R}$,

$$\mathcal{F}_{\mathcal{X}}(L_2(\mathcal{X})) = \{\hat{w} \colon \mathbb{R}^d \to \mathbb{C} \mid \exists w \in L_2(\mathcal{X}) : \hat{w} = \mathcal{F}_{\mathcal{X}} w\} \subset L_2(\mathbb{R}^d;\mathbb{C}),$$

and define the Hilbert space H_f (over $\mathbb{R}$) as the closure of $\mathcal{F}_{\mathcal{X}}(L_2(\mathcal{X}))$ with respect to norm induced by the weighted $L_2(\mathbb{R}^d;\mathbb{C})$-inner product with weight f,

$$(\hat{v}_1, \hat{v}_2)_{H_f} := \int_{\mathbb{R}^d} f(\omega)\hat{v}_1(\omega)\overline{\hat{v}_2(\omega)}\,\mathrm{d}\omega, \qquad H_f := \overline{\mathcal{F}_{\mathcal{X}}(L_2(\mathcal{X}))}^{\|\cdot\|_{H_f}}. \tag{5.2}$$

We recall the Hilbert space $H = \mathcal{C}^{-1/2}(L_2(\mathcal{X}))$ with $(\cdot,\cdot)_H = (\mathcal{C}^{1/2}\cdot, \mathcal{C}^{1/2}\cdot)_{L_2(\mathcal{X})}$ from Lemma 4.1(i) and find by invoking (5.1) that, for all $v_1, v_2 \in L_2(\mathcal{X})$,

$$\begin{aligned}(\mathcal{F}_{\mathcal{X}}v_1, \mathcal{F}_{\mathcal{X}}v_2)_{H_f} &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}(\mathcal{F}\varrho_0)(\omega)(\mathcal{F}_{\mathcal{X}}v_1)(\omega)\overline{(\mathcal{F}_{\mathcal{X}}v_2)(\omega)}\,\mathrm{d}\omega \\ &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\mathcal{F}(\varrho_0 * (E^0_{\mathcal{X}}v_1))(\omega)\overline{\mathcal{F}(E^0_{\mathcal{X}}v_2)(\omega)}\,\mathrm{d}\omega \\ &= (\varrho_0 * (E^0_{\mathcal{X}}v_1), E^0_{\mathcal{X}}v_2)_{L_2(\mathbb{R}^d)} = (\mathcal{C}v_1, v_2)_{L_2(\mathcal{X})} = (v_1, v_2)_H.\end{aligned} \tag{5.3}$$

By density of $L_2(\mathcal{X})$ in H and of $\mathcal{F}_{\mathcal{X}}(L_2(\mathcal{X}))$ in H_f, $\mathcal{F}_{\mathcal{X}}$ thus admits a unique continuous linear extension to an inner product preserving isometric isomorphism between H and H_f. Its inverse $\mathcal{F}^{-1}_{\mathcal{X}} \colon H_f \to H$ is the unique continuous linear extension of $R_{\mathcal{X}} \circ \mathcal{F}^{-1} \colon \mathcal{F}_{\mathcal{X}}(L_2(\mathcal{X})) \to L_2(\mathcal{X}) \subset H$, where $R_{\mathcal{X}} \colon L_2(\mathbb{R}^d) \to L_2(\mathcal{X})$ denotes the restriction to $\mathcal{X}$.

PROPOSITION 5.2. *Suppose that the self-adjoint, positive definite, compact operators $\mathcal{C}, \widetilde{\mathcal{C}} \colon L_2(\mathcal{X}) \to L_2(\mathcal{X})$ are induced by restrictions (to $\mathcal{X} \times \mathcal{X}$) of translation invariant covariance functions $\varrho, \widetilde{\varrho} \colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, which have spectral densities $f, \widetilde{f} \colon \mathbb{R}^d \to [0, \infty)$ defined via (5.1). Then, Assumptions 3.3.I and 3.3.III are satisfied if and only if:*

(I') *The spaces H_f and $H_{\widetilde{f}}$ are isomorphic with equivalent norms, that is, there exist constants $0 < k \le K < \infty$ such that*

$$k \|\hat{v}\|_{H_f}^2 \le \int_{\mathbb{R}^d} \widetilde{f}(\omega) |\hat{v}(\omega)|^2 \,\mathrm{d}\omega \le K \|\hat{v}\|_{H_f}^2 \quad \forall \hat{v} \in \mathcal{F}_{\mathcal{X}}(L_2(\mathcal{X})). \tag{5.4}$$

(III') *There exists a constant $a \in (0, \infty)$ such that the linear operator $\widehat{T}_a := S - a\mathcal{I}_{H_f}$ is compact on H_f, where $\mathcal{I}_{H_f}$ denotes the identity on H_f and $S \colon H_f \to H_f$ is defined by*

$$(S\hat{v}_1, \hat{v}_2)_{H_f} = \int_{\mathbb{R}^d} \widetilde{f}(\omega) \hat{v}_1(\omega) \overline{\hat{v}_2(\omega)} \,\mathrm{d}\omega \quad \forall \hat{v}_1, \hat{v}_2 \in H_f. \tag{5.5}$$

PROOF. (I') Let $\hat{v}_1, \hat{v}_2 \in \mathcal{F}_{\mathcal{X}}(L_2(\mathcal{X}))$ and $v_1, v_2 \in L_2(\mathcal{X})$ be such that $\hat{v}_1 = \mathcal{F}_{\mathcal{X}} v_1$ and $\hat{v}_2 = \mathcal{F}_{\mathcal{X}} v_2$. Applying the inversion formula (5.1) for $\widetilde{f}$ gives, similarly as in (5.3),

$$(\widetilde{\mathcal{C}} v_1, v_2)_{L_2(\mathcal{X})} = \int_{\mathbb{R}^d} \frac{(\mathcal{F}\widetilde{\varrho}_0)(\omega)}{(2\pi)^d} \hat{v}_1(\omega) \overline{\hat{v}_2(\omega)} \,\mathrm{d}\omega = \int_{\mathbb{R}^d} \widetilde{f}(\omega) \hat{v}_1(\omega) \overline{\hat{v}_2(\omega)} \,\mathrm{d}\omega. \tag{5.6}$$

Therefore, (5.4) is equivalent to $k(\mathcal{C}v, v)_{L_2(\mathcal{X})} \le (\widetilde{\mathcal{C}}v, v)_{L_2(\mathcal{X})} \le K(\mathcal{C}v, v)_{L_2(\mathcal{X})}$ holding for all $v \in L_2(\mathcal{X})$, which by density of $L_2(\mathcal{X})$ in H can be reformulated as the relation $\|\widetilde{\mathcal{C}}^{1/2} \mathcal{C}^{-1/2} w\|_{L_2(\mathcal{X})}^2 \in [k, K]$ for all $w \in L_2(\mathcal{X})$ with $\|w\|_{L_2(\mathcal{X})} = 1$, that is, $\widetilde{\mathcal{C}}^{1/2} \mathcal{C}^{-1/2}$ is an isomorphism on $L_2(\mathcal{X})$. By Proposition 3.5(i) $\Leftrightarrow$ (ii) this is equivalent to Assumption 3.3.I.

(III') We proceed as illustrated below: We prove that T_a can be expressed as the composition $T_a = \mathcal{C}^{1/2} \mathcal{F}_{\mathcal{X}}^{-1} \widehat{T}_a \mathcal{F}_{\mathcal{X}} \mathcal{C}^{-1/2}$. Since $\mathcal{C}^{-1/2} \colon L_2(\mathcal{X}) \to H$ and $\mathcal{F}_{\mathcal{X}} \colon H \to H_f$ are inner product preserving isometric isomorphisms, this shows that $T_a \in \mathcal{K}(L_2(\mathcal{X}))$ is equivalent to compactness of $\widehat{T}_a$ on H_f.

$$\begin{array}{ccc}
(L_2(\mathcal{X}), (\cdot,\cdot)_{L_2(\mathcal{X})}) & \xrightarrow{\;T_a = \mathcal{C}^{-1/2}\widetilde{\mathcal{C}}\mathcal{C}^{-1/2} - a\mathcal{I}\;} & (L_2(\mathcal{X}), (\cdot,\cdot)_{L_2(\mathcal{X})}) \\
\downarrow \mathcal{C}^{-1/2} & & \uparrow \mathcal{C}^{1/2} \\
(H, (\mathcal{C}^{1/2}\cdot, \mathcal{C}^{1/2}\cdot)_{L_2(\mathcal{X})}) & \xrightarrow{\;\mathcal{C}^{-1}\widetilde{\mathcal{C}} - a\mathcal{I}_H\;} & (H, (\mathcal{C}^{1/2}\cdot, \mathcal{C}^{1/2}\cdot)_{L_2(\mathcal{X})}) \\
\downarrow \mathcal{F} \circ E_{\mathcal{X}}^0 & & \uparrow R_{\mathcal{X}} \circ \mathcal{F}^{-1} \\
(H_f, (\cdot,\cdot)_{H_f}) & \xrightarrow{\;\widehat{T}_a = S - a\mathcal{I}_{H_f}\;} & (H_f, (\cdot,\cdot)_{H_f})
\end{array}$$

Part (I') implies that $\mathfrak{b}(\hat{v}_1, \hat{v}_2) := \int_{\mathbb{R}^d} \widetilde{f}(\omega) \hat{v}_1(\omega) \overline{\hat{v}_2(\omega)} \,\mathrm{d}\omega$ defines a continuous, coercive bilinear form on the real Hilbert space H_f. Thus, for every $\hat{v}_1 \in H_f$, existence and uniqueness of $S\hat{v}_1$ satisfying (5.5) follows from the Riesz representation theorem, and $S \colon H_f \to H_f$ is well-defined, linear and bounded. For $v_1, v_2 \in L_2(\mathcal{X})$ and $\hat{v}_1 := \mathcal{F}_{\mathcal{X}} v_1$, $\hat{v}_2 := \mathcal{F}_{\mathcal{X}} v_2$, we have

$$\begin{aligned}
\big((\mathcal{C}^{-1}\widetilde{\mathcal{C}} - a\mathcal{I}_H) v_1, v_2\big)_H &= \big((\widetilde{\mathcal{C}} - a\mathcal{C}) v_1, v_2\big)_{L_2(\mathcal{X})} \\
&= \int_{\mathbb{R}^d} \big(\widetilde{f}(\omega) - a f(\omega)\big) \hat{v}_1(\omega) \overline{\hat{v}_2(\omega)} \,\mathrm{d}\omega \\
&= \big((S - a\mathcal{I}_{H_f}) \hat{v}_1, \hat{v}_2\big)_{H_f} \\
&= (\widehat{T}_a \hat{v}_1, \hat{v}_2)_{H_f} = \big(\mathcal{F}_{\mathcal{X}}^{-1} \widehat{T}_a \mathcal{F}_{\mathcal{X}} v_1, v_2\big)_H,
\end{aligned}$$

where we used (5.2), (5.3) and (5.6). By density of $L_2(\mathcal{X})$ in H and continuity of $\mathcal{F}_{\mathcal{X}}^{-1} \widehat{T}_a \mathcal{F}_{\mathcal{X}}$ on H, this equality holds also for all $v_1, v_2 \in H$. Consequently, we obtain the chain of identities $T_a = \mathcal{C}^{-1/2} \widetilde{\mathcal{C}} \mathcal{C}^{-1/2} - a\mathcal{I} = \mathcal{C}^{1/2} (\mathcal{C}^{-1} \widetilde{\mathcal{C}} - a\mathcal{I}_H) \mathcal{C}^{-1/2} = \mathcal{C}^{1/2} \mathcal{F}_{\mathcal{X}}^{-1} \widehat{T}_a \mathcal{F}_{\mathcal{X}} \mathcal{C}^{-1/2}$. $\square$

REMARK 5.3. We emphasize that the Hilbert spaces $\mathcal{H}^0$ in (2.4), H from Lemma 4.1(i) and H_f in (5.2) are mutually isomorphic, with inner product preserving isomorphisms $\mathcal{J}\colon H \to \mathcal{H}^0$ and $\mathcal{F}_{\mathcal{X}}\colon H \to H_f$.

For two continuous functions $g, \widetilde{g}\colon \mathbb{R}^d \to [0,\infty)$, the notation $g \asymp \widetilde{g}$ indicates that there exist $k, K \in (0,\infty)$ such that the relations $kg(\omega) \le \widetilde{g}(\omega) \le Kg(\omega)$ hold for all $\omega \in \mathbb{R}^d$.

COROLLARY 5.4. *Suppose the setting of Proposition 5.2.*

(i) *Assumption 3.3.I is satisfied whenever $f \asymp \widetilde{f}$.*

(ii) *Suppose that $\varrho_0\colon \mathbb{R}^d \to \mathbb{R}$ related to $f\colon \mathbb{R}^d \to [0,\infty)$ via (5.1) is not infinitely differentiable in at least one Cartesian coordinate direction. Then, in either of the cases $\frac{\widetilde{f}(\omega)}{f(\omega)} \to 0$ or $\frac{\widetilde{f}(\omega)}{f(\omega)} \to \infty$ as $\|\omega\|_{\mathbb{R}^d} \to \infty$, Assumption 3.3.I is not satisfied.*

PROOF. (i) Clearly, the relation $f \asymp \widetilde{f}$ in (i) implies that (5.4) holds. Therefore, by Proposition 5.2(I') Assumption 3.3.I is satisfied.

(ii) Without loss of generality, we may assume that $\mathcal{X}$ contains an open subset containing the origin and pick $L \in (0,\infty)$ with $[-L^{-1}, L^{-1}]^d \subseteq \mathcal{X}$. By assumption on the differentiability of ϱ_0, there exist $j \in \{1,\dots,d\}$ and $p \in \mathbb{N}$ such that $\int_{\mathbb{R}} f_j(\omega_j)\omega_j^{2p}\,\mathrm{d}\omega_j = \infty$, where $f_j\colon \mathbb{R} \to [0,\infty)$ is defined by

$$f_j(\omega_j) := \int_{\mathbb{R}^{d-1}} f(\omega_1,\dots,\omega_j,\dots,\omega_d)\,\mathrm{d}\lambda_{d-1}(\omega_1,\dots,\omega_{j-1},\omega_{j+1},\dots,\omega_d).$$

Let $r(\,\cdot\,)$ be the rectangular function, defined as $r(x) = 1$ for $|x| \le \frac{1}{2}$ and $r(x) = 0$ for $|x| > \frac{1}{2}$. For $n \in \mathbb{N}$ such that $n > pL$, consider the forward difference operator $\Delta_n\colon L_2(\mathbb{R}) \to L_2(\mathbb{R})$, $[\Delta_n g](x) := n(g(x+1/n) - g(x))$, and set

$$v_n(x) := n^d[\Delta_n^p r](nx_j - \tfrac{1}{2}) \prod_{k\neq j} r(nx_k - \tfrac{1}{2}), \quad n \in \mathbb{N}, \quad n > pL.$$

Each function v_n has compact support in $[-L^{-1}, L^{-1}]^d \subseteq \mathcal{X}$ and $v_n \in L_2(\mathcal{X})$. Furthermore, its Fourier transform is $\hat{v}_n(\omega) := (\mathcal{F}_{\mathcal{X}} v_n)(\omega) = n^p(e^{i\omega_j/n} - 1)^p \prod_{k=1}^d e^{-i\omega_k/2}\operatorname{sinc}(\frac{\omega_k}{2\pi n})$, where $\operatorname{sinc}(x) := \frac{\sin(\pi x)}{\pi x}$ for all $x \in \mathbb{R}$, and by basic trigonometric identities

$$|\hat{v}_n(\omega)|^2 = \left[2n\sin\left(\frac{\omega_j}{2n}\right)\right]^{2p} \prod_{k=1}^d \operatorname{sinc}^2\left(\frac{\omega_k}{2\pi n}\right) \tag{5.7}$$

follows. Fix $\ell \in \mathbb{N}$. By assumption, there exists a constant $M_\ell \in (0,\infty)$ such that one of the following holds for all ω with $\|\omega\|_{\mathbb{R}^d} > M_\ell$: **(a)** $\widetilde{f}(\omega) < \frac{1}{2\ell} f(\omega)$ or **(b)** $\widetilde{f}(\omega) > 2\ell f(\omega)$. Next, define the infinite strip $A_\ell := \{\omega \in \mathbb{R}^d : |\omega_j| \le M_\ell\}$. Then, for every $n > pL$, we find by (5.7)

$$\int_{A_\ell} f(\omega)|\hat{v}_n(\omega)|^2\,\mathrm{d}\omega \le \int_{A_\ell} f(\omega)\omega_j^{2p}\,\mathrm{d}\omega \le M_\ell^{2p}\int_{\mathbb{R}^d} f(\omega)\,\mathrm{d}\omega = M_\ell^{2p}\varrho_0(0),$$

since $2n\sin(\frac{\omega_j}{2n}) = \operatorname{sinc}(\frac{\omega_j}{2\pi n})\omega_j$ and $|\operatorname{sinc}(x)| \le 1$ for all $x \in \mathbb{R}$. By the same arguments, $\int_{A_\ell} \widetilde{f}(\omega)|\hat{v}_n(\omega)|^2\,\mathrm{d}\omega \le M_\ell^{2p}\widetilde{\varrho}_0(0)$ holds for all $n > pL$. In addition, for every $n \in \mathbb{N}$ with $n > \max\{pL, M_\ell/\pi\}$, define the set

$$B_\ell^n := \{\omega \in \mathbb{R}^d : M_\ell < |\omega_j| < n\pi,\ |\omega_k| < n\pi \ \ \forall k \neq j\} \subset A_\ell^c := \mathbb{R}^d \setminus A_\ell.$$

Since $\operatorname{sinc}^2(\theta/(2\pi n)) > (2/\pi)^2$ for $\theta \in (-\pi n, \pi n)$, we obtain again by (5.7) that

$$\int_{B_\ell^n} f(\omega)|\hat{v}_n(\omega)|^2\,\mathrm{d}\omega > \frac{4^d}{\pi^{2d}}\int_{B_\ell^n} f(\omega)\left[2n\sin\left(\frac{\omega_j}{2n}\right)\right]^{2p}\mathrm{d}\omega > \frac{4^{d+p}}{\pi^{2(d+p)}}\int_{B_\ell^n} f(\omega)\omega_j^{2p}\,\mathrm{d}\omega.$$

Furthermore, note that $\lim_{n\to\infty}\int_{B_\ell^n} f(\omega)\omega_j^{2p}\,\mathrm{d}\omega = \int_{A_\ell^c} f(\omega)\omega_j^{2p}\,\mathrm{d}\omega = \infty$, since the integral $\int_{\mathbb{R}^d} f(\omega)\omega_j^{2p}\,\mathrm{d}\omega = \int_{\mathbb{R}} f_j(\omega_j)\omega_j^{2p}\,\mathrm{d}\omega_j = \infty$ diverges and $\int_{A_\ell} f(\omega)\omega_j^{2p}\,\mathrm{d}\omega \le M_\ell^{2p}\varrho_0(0)$ is finite. For this reason, there exists an integer $n_0 = n_0(\ell) > \max\{pL, M_\ell/\pi\}$ such that

$$\int_{A_\ell^c} f(\omega)\big|\hat{v}_n(\omega)\big|^2\,\mathrm{d}\omega \ge \int_{B_\ell^n} f(\omega)\big|\hat{v}_n(\omega)\big|^2\,\mathrm{d}\omega > 2\ell M_\ell^{2p}\max\{\varrho_0(0), \widetilde{\varrho}_0(0)\},$$

for all $n \ge n_0$. We then obtain, for every $n \ge n_0$, in case **(a)** the estimate

$$\frac{\int_{\mathbb{R}^d}\widetilde{f}(\omega)|\hat{v}_n(\omega)|^2\,\mathrm{d}\omega}{\|\hat{v}_n\|_{H_f}^2} \le \frac{\int_{A_\ell}\widetilde{f}(\omega)|\hat{v}_n(\omega)|^2\,\mathrm{d}\omega}{\int_{A_\ell^c} f(\omega)|\hat{v}_n(\omega)|^2\,\mathrm{d}\omega} + \frac{\int_{A_\ell^c}\widetilde{f}(\omega)|\hat{v}_n(\omega)|^2\,\mathrm{d}\omega}{\int_{A_\ell^c} f(\omega)|\hat{v}_n(\omega)|^2\,\mathrm{d}\omega} < \ell^{-1},$$

and in case **(b)** we have, for all $n \ge n_0$,

$$\begin{aligned}\frac{\int_{\mathbb{R}^d}\widetilde{f}(\omega)|\hat{v}_n(\omega)|^2\,\mathrm{d}\omega}{\|\hat{v}_n\|_{H_f}^2} &\ge \frac{\int_{A_\ell^c}\widetilde{f}(\omega)|\hat{v}_n(\omega)|^2\,\mathrm{d}\omega}{\int_{A_\ell} f(\omega)|\hat{v}_n(\omega)|^2\,\mathrm{d}\omega + \int_{A_\ell^c} f(\omega)|\hat{v}_n(\omega)|^2\,\mathrm{d}\omega} \\ &= \frac{\int_{A_\ell^c}\widetilde{f}(\omega)|\hat{v}_n(\omega)|^2\,\mathrm{d}\omega}{\int_{A_\ell^c} f(\omega)|\hat{v}_n(\omega)|^2\,\mathrm{d}\omega}\left(\frac{\int_{A_\ell} f(\omega)|\hat{v}_n(\omega)|^2\,\mathrm{d}\omega}{\int_{A_\ell^c} f(\omega)|\hat{v}_n(\omega)|^2\,\mathrm{d}\omega} + 1\right)^{-1} \\ &\ge 2\ell\left(\frac{1}{2\ell} + 1\right)^{-1} = \frac{4\ell^2}{1+2\ell} \ge \ell.\end{aligned}$$

Since $\ell \in \mathbb{N}$ was arbitrary, in case **(a)** there is no constant $k \in (0,\infty)$ such that the lower bound in (5.4) holds and in case **(b)** we cannot find $K \in (0,\infty)$ for the upper bound in (5.4). Thus, the result follows by Proposition 5.2(I'). $\square$

In what follows, we let $\mathcal{W}_{\Pi_{\mathbf{R}}}$ denote the class of Fourier transforms $\mathcal{F}v$ of square-integrable functions $v \in L_2(\mathbb{R}^d;\mathbb{C})$ with support $\operatorname{supp}(v) \subseteq \Pi_{\mathbf{R}}$ inside the bounded parallelepiped $\Pi_{\mathbf{R}} = \{x \in \mathbb{R}^d : -R_j \le x_j \le R_j,\ j = 1,\dots,d\}$.

COROLLARY 5.5. *Suppose the setting of Proposition* 5.2, $f \asymp \widetilde{f}$, *and furthermore that there exists* $\varphi_0 \in \mathcal{W}_{\Pi_{\mathbf{R}}}$ *such that* $f \asymp |\varphi_0|^2$. *Then Assumptions* 3.3.I *and* 3.3.III *are satisfied whenever there exists a constant* $a \in (0,\infty)$ *such that* $\frac{\widetilde{f}(\omega)}{f(\omega)} \to a$ *as* $\|\omega\|_{\mathbb{R}^d} \to \infty$.

PROOF. By Corollary 5.4(i), $f \asymp \widetilde{f}$ implies that Assumption 3.3.I holds.

Next, recall the bounded linear operator $S\colon H_f \to H_f$ from (5.5) and, for $\ell \in \mathbb{N}$, define the self-adjoint linear operator $\widehat{T}_a^\ell\colon H_f \to H_f$ similarly via

$$\big(\widehat{T}_a^\ell \hat{v}_1, \hat{v}_2\big)_{H_f} = \int_{B_\ell}\big(\widetilde{f}(\omega) - af(\omega)\big)\hat{v}_1(\omega)\overline{\hat{v}_2(\omega)}\,\mathrm{d}\omega \quad \forall \hat{v}_1, \hat{v}_2 \in H_f,$$

where $B_\ell := \{w \in \mathbb{R}^d : \|\omega\|_{\mathbb{R}^d} < \ell\}$ is the ball around the origin with radius ℓ. By Proposition 5.2(I'), $\widehat{T}_a^\ell$ is bounded with $\|\widehat{T}_a^\ell\|_{\mathcal{L}(H_f)} \le a + \|S\|_{\mathcal{L}(H_f)}$ for all $\ell \in \mathbb{N}$. We now proceed in two steps: We first show that, for every $\ell \in \mathbb{N}$, $\widehat{T}_a^\ell$ is compact on H_f. Second, we prove convergence $\lim_{\ell\to\infty}\|\widehat{T}_a^\ell - \widehat{T}_a\|_{\mathcal{L}(H_f)} = 0$, which implies that $\widehat{T}_a = S - a\mathcal{I}_{H_f}$ is compact on H_f, since $\mathcal{K}(H_f)$ is closed in $\mathcal{L}(H_f)$. Then Assumption 3.3.III holds by Proposition 5.2(III').

By Proposition 5.2(I'), H_f and $H_{|\varphi_0|^2}$ are isomorphic and $c\|\hat{v}\|_{H_f}^2 \le \|\hat{v}\|_{H_{|\varphi_0|^2}}^2 \le C\|\hat{v}\|_{H_f}^2$ for some constants $c, C \in (0,\infty)$ independent of $\hat{v} \in H_f$. For this reason, the operator $\widehat{T}_a^\ell$ is compact on H_f if and only if it is compact on $H_{|\varphi_0|^2}$. To see that, for every $\ell \in \mathbb{N}$, the operator $\widehat{T}_a^\ell$ is compact on $H_{|\varphi_0|^2}$, we prove the stronger result that $\widehat{T}_a^\ell$ is Hilbert–Schmidt on $H_{|\varphi_0|^2}$.

Let $\{\hat{v}_j\}_{j\in\mathbb{N}}$ be an orthonormal basis for $H_{|\varphi_0|^2}$. Then we note that the relations $f \asymp \widetilde{f}$ and $f \asymp |\varphi_0|^2$ imply equality of the supports, $\operatorname{supp}(\widetilde{f}) = \operatorname{supp}(f) = \operatorname{supp}(\varphi_0)$, and estimate

$$
\begin{aligned}
C^{-1} &\sum_{j\in\mathbb{N}} \|\widehat{T}_a^\ell \hat{v}_j\|_{H_{|\varphi_0|^2}}^2 \\
&\leq \sum_{j\in\mathbb{N}} \|\widehat{T}_a^\ell \hat{v}_j\|_{H_f}^2 = \sum_{j\in\mathbb{N}} \sup_{\|\hat{w}\|_{H_f}=1} |(\widehat{T}_a^\ell \hat{v}_j, \hat{w})_{H_f}|^2 \\
&= \sum_{j\in\mathbb{N}} \sup_{\|\hat{w}\|_{H_f}=1} \left| \int_{\operatorname{supp}(\varphi_0)\cap B_\ell} \frac{\widetilde{f}(\omega) - af(\omega)}{|\varphi_0(\omega)|^2} |\varphi_0(\omega)|^2 \hat{v}_j(\omega)\overline{\hat{w}(\omega)}\, \mathrm{d}\omega \right|^2 \\
&\leq \sum_{j\in\mathbb{N}} \sup_{\|\hat{w}\|_{H_f}=1} \|\hat{w}\|_{H_{|\varphi_0|^2}}^2 \int_{\operatorname{supp}(\varphi_0)\cap B_\ell} \frac{|\widetilde{f}(\omega) - af(\omega)|^2}{|\varphi_0(\omega)|^4} |\varphi_0(\omega)|^2 |\hat{v}_j(\omega)|^2\, \mathrm{d}\omega \\
&\leq C \left| \sup_{\omega\in\operatorname{supp}(\varphi_0)} \frac{\widetilde{f}(\omega) + af(\omega)}{|\varphi_0(\omega)|^2} \right|^2 \int_{\operatorname{supp}(\varphi_0)\cap B_\ell} |\varphi_0(\omega)|^2 \sum_{j\in\mathbb{N}} |\hat{v}_j(\omega)|^2\, \mathrm{d}\omega.
\end{aligned}
$$

Since $f \asymp |\varphi_0|^2$ and $f \asymp \widetilde{f}$, the supremum in this bound is finite. Furthermore, as $\varphi_0 \in \mathcal{W}_{\Pi_\mathbf{R}}$ by [19], Lemma on page 34, we obtain the bound

$$
|\varphi_0(\omega)|^2 \sum_{j\in\mathbb{N}} |\hat{v}_j(\omega)|^2 \leq C_{\mathbf{R},\mathbf{K}}, \quad \text{where } C_{\mathbf{R},\mathbf{K}} := \prod_{j=1}^d \frac{(R_j + K_j)}{\pi},
$$

and $\Pi_\mathbf{K}$ is a parallelepiped enclosing the compact set $\mathcal{X} \subseteq \Pi_\mathbf{K}$. Consequently,

$$
\sum_{j\in\mathbb{N}} \|\widehat{T}_a^\ell \hat{v}_j\|_{H_{|\varphi_0|^2}}^2 \leq C^2 (2\ell)^d C_{\mathbf{R},\mathbf{K}} \left| \sup_{\omega\in\operatorname{supp}(\varphi_0)} \frac{\widetilde{f}(\omega) + af(\omega)}{|\varphi_0(\omega)|^2} \right|^2 < \infty,
$$

that is, for every $\ell \in \mathbb{N}$, $\widehat{T}_a^\ell$ is Hilbert–Schmidt on $H_{|\varphi_0|^2}$, and thus, compact on H_f.

Finally, let $\varepsilon \in (0, \infty)$ and $\ell_\varepsilon \in \mathbb{N}$ be such that $\sup_{\omega\in B_\ell^c} |\frac{\widetilde{f}(\omega)}{f(\omega)} - a| < \varepsilon$ for all $\ell \geq \ell_\varepsilon$, where $B_\ell^c := \mathbb{R}^d \setminus B_\ell$. Then, for every $\ell \geq \ell_\varepsilon$ and all $\hat{v}_1, \hat{v}_2 \in H_f$,

$$
((\widehat{T}_a - \widehat{T}_a^\ell)\hat{v}_1, \hat{v}_2)_{H_f} = \int_{B_\ell^c} (\widetilde{f}(\omega) - af(\omega))\hat{v}_1(\omega)\overline{\hat{v}_2(\omega)}\, \mathrm{d}\omega \leq \varepsilon \|\hat{v}_1\|_{H_f} \|\hat{v}_2\|_{H_f}.
$$

Thus, $\lim_{\ell\to\infty} \|\widehat{T}_a^\ell - \widehat{T}_a\|_{\mathcal{L}(H_f)} = 0$ and $\widehat{T}_a$ is compact on H_f. □

6. Applications. In the following, we exemplify the results of Section 3 and Section 5 by three specific applications. Corollary 5.4 and Corollary 5.5 can be used to check for uniformly asymptotically optimal linear prediction in the case of weakly stationary processes on compact subsets of $\mathbb{R}^d$, using their spectral densities. As an explicit example, we consider the Matérn covariance family in Section 6.1. Corollary 5.1 is applicable, for example, to periodic random fields on $\mathcal{X} = [0, 1]^d$ as considered by Stein [23]; see Section 6.2. Moreover, as Theorem 3.8 it also holds for random fields on more general domains. As a further illustration, we consider an application on the sphere $\mathcal{X} = \mathbb{S}^2$ in Section 6.3.

6.1. *The Matérn covariance family.* The Matérn covariance function $\varrho|_{\mathcal{X}\times\mathcal{X}}$ on $\mathcal{X} \subset \mathbb{R}^d$ with parameters $\sigma, \nu, \kappa \in (0, \infty)$ (see Example 2.1(a)) has the spectral density

$$
f_\mathsf{M}(\omega) = \frac{1}{(2\pi)^d} (\mathcal{F}\varrho_\mathsf{M})(\omega) = \frac{\Gamma(\nu + d/2)}{\Gamma(\nu)\pi^{d/2}} \frac{\sigma^2\kappa^{2\nu}}{(\kappa^2 + \|\omega\|_{\mathbb{R}^d}^2)^{\nu+d/2}}, \quad \omega \in \mathbb{R}^d; \tag{6.1}
$$

cf. [25], equation (32) on page 49. Assume that $\widetilde{\varrho}$ is a further Matérn covariance function with parameters $\widetilde{\sigma}, \widetilde{\nu}, \widetilde{\kappa} \in (0, \infty)$ and corresponding spectral density $\widetilde{f}_{\mathsf{M}}$. Since

$$\frac{\widetilde{f}_{\mathsf{M}}(\omega)}{f_{\mathsf{M}}(\omega)} = \frac{\Gamma(\widetilde{\nu}+d/2)}{\Gamma(\widetilde{\nu})} \frac{\Gamma(\nu)}{\Gamma(\nu+d/2)} \frac{\widetilde{\sigma}^2\widetilde{\kappa}^{2\widetilde{\nu}}}{\sigma^2\kappa^{2\nu}} \frac{(\kappa^2+\|\omega\|^2_{\mathbb{R}^d})^{\nu+d/2}}{(\widetilde{\kappa}^2+\|\omega\|^2_{\mathbb{R}^d})^{\widetilde{\nu}+d/2}}, \quad \omega \in \mathbb{R}^d,$$

we conclude with Corollary 5.4(ii) that Assumption 3.3.I can only be satisfied if $\widetilde{\nu} = \nu$. In this case, $f_{\mathsf{M}} \asymp \widetilde{f}_{\mathsf{M}}$ and, since by [10], Remark 4.1, also $f_{\mathsf{M}} \asymp |\varphi_0|^2$ holds for some $\varphi_0 \in \mathcal{W}_{\Pi_{\mathbf{R}}}$ and some parallelepiped $\mathbf{R}$, Corollary 5.5 is applicable and shows that Assumptions 3.3.I and 3.3.III hold, with $a = \frac{\widetilde{\sigma}^2\widetilde{\kappa}^{2\nu}}{\sigma^2\kappa^{2\nu}}$ in (3.5). Thus, misspecifying the second-order structure $(0, \varrho)$ by $(0, \widetilde{\varrho})$ yields uniformly asymptotically optimal linear prediction if and only if $\nu = \widetilde{\nu}$. For equivalence of the corresponding Gaussian measures, $a = 1$ is necessary, that is, the microergodic parameter $\sigma^2\kappa^{2\nu}$ has to coincide for the two models. This is in accordance with the identifiability of this parameter under infill asymptotics; see [28].

6.2. *Periodic random fields.* A stochastic process $\{Z(x)\}_{x\in\mathcal{X}}$ indexed by $\mathcal{X} := [0,1]^d$ is said to be weakly periodic if its mean value function $\mathsf{E}[Z] \equiv m$ is constant on $[0,1]^d$ and, in addition, its covariance function $\varrho(x, x')$ only depends on the difference $x - x'$, where the difference is taken modulo 1 in each coordinate (see [23]). Let $\mathbb{Z}^d_+$ denote all elements $\mathbf{k} = (k_1, \ldots, k_d)^\top \in \mathbb{Z}^d$ such that at least one element in the vector is nonzero, and the first nonzero component is positive. A weakly periodic process admits the series expansion

$$Z(x) = X_0 + \sum_{\mathbf{k}\in\mathbb{Z}^d_+} \left[X^{\mathrm{c}}_{\mathbf{k}} \cos(2\pi\mathbf{k}\cdot x) + X^{\mathrm{s}}_{\mathbf{k}} \sin(2\pi\mathbf{k}\cdot x)\right],$$

where X_0, $X^{\mathrm{c}}_{\mathbf{k}}$, $X^{\mathrm{s}}_{\mathbf{k}}$ are pairwise uncorrelated random variables such that $\mathsf{E}[X_0] = \mathsf{E}[Z] = m$ and, for all $\mathbf{k} \in \mathbb{Z}^d_+$, one has $\mathsf{E}[X^{\mathrm{c}}_{\mathbf{k}}] = \mathsf{E}[X^{\mathrm{s}}_{\mathbf{k}}] = 0$ as well as $\mathsf{E}[|X^{\mathrm{c}}_{\mathbf{k}}|^2] = \mathsf{E}[|X^{\mathrm{s}}_{\mathbf{k}}|^2]$. Define $f\colon \mathbb{Z}^d \to [0,\infty)$ by $f(\mathbf{0}) := \mathsf{Var}[X_0]$, $f(\mathbf{k}) := \frac{1}{2}\mathsf{Var}[X^{\mathrm{c}}_{\mathbf{k}}]$ for $\mathbf{k} \in \mathbb{Z}^d_+$, and $f(-\mathbf{k}) = f(\mathbf{k})$. Then we can represent the covariance function of Z as

$$\begin{aligned}
\varrho(x, x') &= \sum_{\mathbf{k}\in\mathbb{Z}^d} f(\mathbf{k})\left[\cos(2\pi\mathbf{k}\cdot x)\cos(2\pi\mathbf{k}\cdot x') + \sin(2\pi\mathbf{k}\cdot x)\sin(2\pi\mathbf{k}\cdot x')\right] \\
&= \sum_{\mathbf{k}\in\mathbb{Z}^d} f(\mathbf{k})\cos\big(2\pi\mathbf{k}\cdot(x - x')\big) =: \varrho_0(x - x').
\end{aligned}$$

For this reason, f can be viewed as the spectral density with respect to the counting measure on $\mathbb{Z}^d$. It is not difficult to show that the set

$$\{1, e^{\mathrm{c}}_{\mathbf{k}}, e^{\mathrm{s}}_{\mathbf{k}} : \mathbf{k} \in \mathbb{Z}^d_+\}, \qquad e^{\mathrm{c}}_{\mathbf{k}}(x) := \sqrt{2}\cos(2\pi\mathbf{k}\cdot x), \qquad e^{\mathrm{s}}_{\mathbf{k}}(x) := \sqrt{2}\sin(2\pi\mathbf{k}\cdot x),$$

forms an orthonormal basis for $L_2([0,1]^d)$. Moreover, it is an eigenbasis of the covariance operator with kernel ϱ. Indeed, $\int_{\mathcal{X}} \varrho(x, x')\,\mathrm{d}x' = f(\mathbf{0})$ and

$$\forall \mathbf{k} \in \mathbb{Z}^d_+: \quad \int_{\mathcal{X}} \varrho(x, x') e^{\iota}_{\mathbf{k}}(x')\,\mathrm{d}x' = f(\mathbf{k}) e^{\iota}_{\mathbf{k}}(x), \quad \iota \in \{\mathrm{c}, \mathrm{s}\}.$$

Since $\varrho_0(0) = \sum_{\mathbf{k}\in\mathbb{Z}^d} f(\mathbf{k}) < \infty$ and $f(\mathbf{k}) \geq 0$, it is clear that $f(\mathbf{k})$ accumulates only at zero. Thus, for any two weakly periodic random fields on $[0,1]^d$ with corresponding spectral densities $f, \widetilde{f}\colon \mathbb{Z}^d \to [0,\infty)$ defined as above, we are in the setting of Corollary 5.1: Assumptions 3.3.I and 3.3.III are satisfied if and only if $\widetilde{f}(\mathbf{k})/f(\mathbf{k}) \to a$ for some $a \in (0,\infty)$ as $|\mathbf{k}| \to \infty$. This result holds without any further assumptions on the spectral densities, and can be viewed as a version of [25], Chapter 3, Theorem 10, for periodic random fields.

6.3. *Random fields on the sphere.* Due to the popularity of the Matérn covariance family on $\mathbb{R}^d$ (see Example 2.1(a) and Section 6.1) it is highly desirable to have a corresponding covariance model also on the sphere $\mathbb{S}^2$. A simple remedy for this is to define the covariance function as in Example 2.1(c), that is, via the *chordal distance* $d_{\mathbb{R}^3}(x, x') = \|x - x'\|_{\mathbb{R}^3}$. One reason for why this is a common choice is that the (more suitable) great circle distance $d_{\mathbb{S}^2}(x, x') = \arccos((x, x')_{\mathbb{R}^3})$ results in a kernel ϱ which is (strictly) positive definite only for $\nu \leq 1/2$; see Example 2.1(b) and [6]. As this severely limits the flexibility of the model, several authors have suggested alternative "Matérn-like" covariances on $\mathbb{S}^2$.

Guinness and Fuentes [7] proposed the *Legendre–Matérn* covariance,

$$\varrho_1(x, x') := \sum_{\ell=0}^{\infty} \frac{\sigma_1^2}{(\kappa_1^2 + \ell^2)^{\nu_1 + 1/2}} P_\ell\big(\cos d_{\mathbb{S}^2}(x, x')\big), \quad x, x' \in \mathbb{S}^2, \tag{6.2}$$

where $\sigma_1, \nu_1, \kappa_1 \in (0, \infty)$ are model parameters and $P_\ell \colon [-1, 1] \to \mathbb{R}$ is the ℓth Legendre polynomial, that is,

$$P_\ell(y) = 2^{-\ell} \frac{1}{\ell!} \frac{\mathrm{d}^\ell}{\mathrm{d}y^\ell} (y^2 - 1)^\ell, \quad y \in [-1, 1], \quad \ell \in \mathbb{N}_0 := \{0, 1, 2, \ldots\}.$$

This choice is motivated first by the Legendre polynomial representation of positive definite functions on $\mathbb{S}^2$ (see [17]), and second by the fact that the spectral density $f_{\mathsf{M}}(\omega)$ for the Matérn covariance on $\mathbb{R}^d$ is proportional to $\sigma^2(\kappa^2 + \|\omega\|_{\mathbb{R}^d}^2)^{-(\nu + d/2)}$; see (6.1). However, note that the parameter σ_1^2 in (6.2) is not the variance since

$$\varrho_1(x, x) = \sum_{\ell=0}^{\infty} \frac{\sigma_1^2}{(\kappa_1^2 + \ell^2)^{\nu_1 + 1/2}}.$$

Another plausible way of defining a Matérn model on $\mathbb{S}^2$ is to use the stochastic partial differential equation (SPDE) representation of Gaussian Matérn fields derived by Whittle [27], according to which a centered Gaussian Matérn field $\{Z^0(x) : x \in \mathbb{R}^d\}$ can be viewed as a solution to the SPDE

$$\big(\kappa^2 - \Delta\big)^{(\nu + d/2)/2} \big(\tau Z^0\big) = \mathcal{W} \quad \text{on } \mathbb{R}^d, \tag{6.3}$$

where the parameter $\tau \in (0, \infty)$ controls the variance of Z^0, $\mathcal{W}$ is Gaussian white noise and Δ is the Laplacian. Lindgren, Rue and Lindström [12] proposed Gaussian Matérn fields on the sphere as solutions to (6.3) formulated on $\mathbb{S}^2$ instead of $\mathbb{R}^d$. In this case, Δ is the Laplace–Beltrami operator.

In order to state the corresponding covariance function $\varrho_2 \colon \mathbb{S}^2 \times \mathbb{S}^2 \to \mathbb{R}$, we introduce the spherical coordinates $(\vartheta, \varphi) \in [0, \pi] \times [0, 2\pi)$ of a point $(x_1, x_2, x_3)^\top \in \mathbb{R}^3$ on $\mathbb{S}^2$ by $\vartheta = \arccos(x_3)$ and $\varphi = \arccos(x_1(x_1^2 + x_2^2)^{-1/2})$. For all $\ell \in \mathbb{N}_0$ and $m \in \{-\ell, \ldots, \ell\}$, we then define the (complex-valued) spherical harmonic $Y_{\ell,m} : \mathbb{S}^2 \to \mathbb{C}$ as (see [13], page 64)

$$Y_{\ell,m}(\vartheta, \varphi) = C_{\ell,m} P_{\ell,m}(\cos \vartheta) e^{im\varphi}, \qquad m \geq 0,$$

$$Y_{\ell,m}(\vartheta, \varphi) = (-1)^m \overline{Y}_{\ell,-m}(\vartheta, \varphi), \qquad m < 0,$$

where, for $\ell \in \mathbb{N}_0$ and $m \in \{0, \ldots, \ell\}$, we set $C_{\ell,m} := \sqrt{\frac{2\ell+1}{4\pi} \frac{(\ell-m)!}{(\ell+m)!}}$ and $P_{\ell,m} \colon [-1, 1] \to \mathbb{R}$ denotes the associated Legendre polynomial, given by

$$P_{\ell,m}(y) = (-1)^m \big(1 - y^2\big)^{m/2} \frac{\mathrm{d}^m}{\mathrm{d}y^m} P_\ell(y), \quad y \in [-1, 1].$$

The spherical harmonics $\{Y_{\ell,m} : \ell \in \mathbb{N}_0, m = -\ell, \ldots, \ell\}$ are eigenfunctions of the Laplace–Beltrami operator, with corresponding eigenvalues given by $\lambda_{\ell,m} = -\ell(\ell + 1)$. In addition, they form an orthonormal basis of the complex-valued Lebesgue space $L_2(\mathbb{S}^2, \nu_{\mathbb{S}^2}; \mathbb{C})$; see [13], Proposition 3.29. Here, $\nu_{\mathbb{S}^2}$ denotes the Lebesgue measure on the sphere which, in spherical coordinates, can be expressed as $d\nu_{\mathbb{S}^2}(x) = \sin\vartheta \, d\vartheta \, d\varphi$.

The covariance function of the solution Z^0 to the SPDE (6.3) on $\mathbb{S}^2$ can thus be represented using the spherical harmonics via the series expansion (cf. [13], Theorem 5.13 and page 125)

$$\varrho_2(x, x') = \sum_{\ell=0}^{\infty} \frac{\tau^{-2}}{(\kappa^2 + \ell(\ell+1))^{\nu+1}} \sum_{m=-\ell}^{\ell} Y_{\ell,m}(\vartheta, \varphi) \overline{Y}_{\ell,m}(\vartheta', \varphi'),$$

where (ϑ, φ), (ϑ', φ') are the spherical coordinates of x and x', respectively. Then, by expressing also the Legendre–Matérn covariance function in (6.2) in spherical coordinates and by using the addition formula for the spherical harmonics ([13], equation (3.42)) we find that

$$\begin{aligned} \varrho_1(x, x') &= \sum_{\ell=0}^{\infty} \frac{\sigma_1^2}{(\kappa_1^2 + \ell^2)^{\nu_1+1/2}} P_\ell((x, x')_{\mathbb{R}^3}) \\ &= \sum_{\ell=0}^{\infty} \frac{\sigma_1^2}{(\kappa_1^2 + \ell^2)^{\nu_1+1/2}} \frac{4\pi}{2\ell+1} \sum_{m=-\ell}^{\ell} Y_{\ell,m}(\vartheta, \varphi) \overline{Y}_{\ell,m}(\vartheta', \varphi'). \end{aligned}$$

Thus, the covariance functions ϱ_1, ϱ_2 are similar, but not identical. Due to the SPDE representation of ϱ_2, we believe that this is the preferable model. However, an immediate question is now if the two models provide similar kriging predictions. The answer to this is given by Corollary 5.1: Since $\sum_{m=-\ell}^{\ell} Y_{\ell,m}(\vartheta, \varphi) \overline{Y}_{\ell,m}(\vartheta', \varphi') = \sum_{m=-\ell}^{\ell} v_{\ell,m}(\vartheta, \varphi) v_{\ell,m}(\vartheta', \varphi')$, where

$$v_{\ell,m}(\vartheta, \varphi) := \begin{cases} \sqrt{2} C_{\ell,-m} P_{\ell,-m}(\cos\vartheta) \cos(m\varphi) & \text{if } m < 0, \\ (1/\sqrt{4\pi}) P_\ell(\cos\vartheta) & \text{if } m = 0, \\ \sqrt{2} C_{\ell,m} P_{\ell,m}(\cos\vartheta) \sin(m\varphi) & \text{if } m > 0, \end{cases}$$

the two covariance operators have the same (orthonormal, real-valued) eigenfunctions in $L_2(\mathbb{S}^2, \nu_{\mathbb{S}^2}; \mathbb{R})$. Thus, we are in the setting of Corollary 5.1 and consider the limit of the ratio of the corresponding eigenvalues:

$$\lim_{\ell\to\infty} \frac{(\kappa_1^2 + \ell^2)^{\nu_1+1/2}(2\ell+1)}{(\kappa^2 + \ell(\ell+1))^{\nu+1}} \frac{1}{\tau^2\sigma_1^2 4\pi} = \begin{cases} 0 & \text{if } \nu_1 < \nu, \\ \infty & \text{if } \nu_1 > \nu, \\ \dfrac{1}{\tau^2\sigma_1^2 2\pi} & \text{if } \nu_1 = \nu. \end{cases}$$

We conclude that the models will provide asymptotically equivalent kriging prediction as long as they have the same smoothness parameter ν (and positive, finite variance parameters). By the same reasoning, it is easy to see that one may misspecify both τ and κ as well as σ_1 and κ_1 for the two covariance models and still obtain asymptotically optimal linear prediction.

7. Discussion. For statistical applications, it is crucial to understand the effect that misspecifying the mean or the covariance function has on linear prediction. We have addressed this by providing three necessary and sufficient conditions, Assumptions 3.3.I–III, for uniformly asymptotically optimal linear prediction of random fields on compact metric spaces.

There are several directions in which this work can be continued in the future. An interesting question is whether Assumptions 3.3.I–III can be relaxed if the uniformity requirement on the optimality is dropped. Furthermore, the results of Section 5.2 can likely be refined to obtain *necessary and sufficient* conditions on the spectral densities f and $\widetilde{f}$. This should be possible at least in the case that $f \asymp |\varphi_0|^2$ holds for some $\varphi_0 \in \mathcal{W}_{\Pi_{\mathbf{R}}}$.

A more challenging problem would be to generalize our results to the setting of *locally compact* spaces. This extension is conceivable, but it would require substantial changes to both the problem formulation and the methods of proving. For the current setting of compact metric spaces, there are several additional applications that can be considered. For example, the application to Gaussian Matérn fields on the sphere in Section 6.3 can easily be extended to SPDE-based Gaussian Matérn fields on more general domains, since our arguments depend only on the asymptotic behavior of the eigenvalues of the Laplace–Beltrami operator, which is also known, for instance, on compact Riemannian manifolds; see, for example, [18], Theorem 15.2.

Acknowledgments. The authors thank S.G. Cox and J.M.A.M. van Neerven for fruitful discussions on spectral theory, which considerably contributed to the proof of Lemma B.2; see Appendix B in the Supplementary Material [9]. In addition, we thank the Editor and an anonymous reviewer for their valuable comments.

SUPPLEMENTARY MATERIAL

Supplement to "Necessary and sufficient conditions for asymptotically optimal linear prediction of random fields on compact metric spaces" (DOI: 10.1214/21-AOS2138SUPP; .pdf). Three appendices of the manuscript.

REFERENCES

[1] ANDERES, E., MØLLER, J. and RASMUSSEN, J. G. (2020). Isotropic covariance functions on graphs and their edges. *Ann. Statist.* **48** 2478–2503. MR4134803 https://doi.org/10.1214/19-AOS1896

[2] ARAFAT, A., PORCU, E., BEVILACQUA, M. and MATEU, J. (2018). Equivalence and orthogonality of Gaussian measures on spheres. *J. Multivariate Anal.* **167** 306–318. MR3830648 https://doi.org/10.1016/j.jmva.2018.05.005

[3] BERLINET, A. and THOMAS-AGNAN, C. (2004). *Reproducing Kernel Hilbert Spaces in Probability and Statistics*. Kluwer Academic, Boston, MA. MR2239907 https://doi.org/10.1007/978-1-4419-9096-9

[4] BEVILACQUA, M., FAOUZI, T., FURRER, R. and PORCU, E. (2019). Estimation and prediction using generalized Wendland covariance functions under fixed domain asymptotics. *Ann. Statist.* **47** 828–856. MR3909952 https://doi.org/10.1214/17-AOS1652

[5] CLEVELAND, W. S. (1971). Projection with the wrong inner product and its application to regression with correlated errors and linear filtering of time series. *Ann. Math. Stat.* **42** 616–624. MR0286236 https://doi.org/10.1214/aoms/1177693411

[6] GNEITING, T. (2013). Strictly and non-strictly positive definite functions on spheres. *Bernoulli* **19** 1327–1349. MR3102554 https://doi.org/10.3150/12-BEJSP06

[7] GUINNESS, J. and FUENTES, M. (2016). Isotropic covariance functions on spheres: Some properties and modeling considerations. *J. Multivariate Anal.* **143** 143–152. MR3431424 https://doi.org/10.1016/j.jmva.2015.08.018

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

[9] KIRCHNER, K. and BOLIN, D. (2022). Supplement to "Necessary and sufficient conditions for asymptotically optimal linear prediction of random fields on compact metric spaces." https://doi.org/10.1214/21-AOS2138SUPP

[10] KRASNITSKII, S. M. (1990). Conditions for the equivalence of distributions of homogeneous Gaussian fields. *Theory Probab. Appl.* **34** 720–725.

[11] KUKUSH, A. (2019). *Gaussian Measures in Hilbert Space: Construction and Properties. Mathematics and Statistics*. ISTE Ltd, London.

[12] LINDGREN, F., RUE, H. and LINDSTRÖM, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73** 423–498. MR2853727 https://doi.org/10.1111/j.1467-9868.2011.00777.x

[13] MARINUCCI, D. and PECCATI, G. (2011). *Random Fields on the Sphere. London Mathematical Society Lecture Note Series* **389**. Cambridge Univ. Press, Cambridge. MR2840154 https://doi.org/10.1017/CBO9780511751677

[14] MATÉRN, B. (1960). *Spatial Variation*: *Stochastic Models and Their Application to Some Problems in Forest Surveys and Other Sampling Investigations*. Meddelanden Från Statens Skogsforskningsinstitut, Band 49, Nr. 5, Stockholm. MR0169346

[15] MERCER, J. (1909). Functions of positive and negative type and their connection with the theory of integral equations. *Philos*. *Trans*. *R*. *Soc*. *A* **209** 415–446.

[16] PARZEN, E. (1959). Statistical inference on time series by Hilbert space methods, I. Technical Report No. 23, Department of Statistics, Stanford Univ.

[17] SCHOENBERG, I. J. (1942). Positive definite functions on spheres. *Duke Math*. *J*. **9** 96–108. MR0005922

[18] SHUBIN, M. A. (2001). *Pseudodifferential Operators and Spectral Theory*, 2nd ed. Springer, Berlin. MR1852334 https://doi.org/10.1007/978-3-642-56579-3

[19] SKOROKHOD, A. V. and YADRENKO, M. I. (1973). On absolute continuity of measures corresponding to homogeneous Gaussian fields. *Theory Probab*. *Appl*. **18** 27–40.

[20] STEIN, M. L. (1988). Asymptotically efficient prediction of a random field with a misspecified covariance function. *Ann*. *Statist*. **16** 55–63. MR0924856 https://doi.org/10.1214/aos/1176350690

[21] STEIN, M. L. (1990). Uniform asymptotic optimality of linear predictions of a random field using an incorrect second-order structure. *Ann*. *Statist*. **18** 850–872. MR1056340 https://doi.org/10.1214/aos/1176347629

[22] STEIN, M. L. (1993). A simple condition for asymptotic optimality of linear predictions of random fields. *Statist*. *Probab*. *Lett*. **17** 399–404. MR1237787 https://doi.org/10.1016/0167-7152(93)90261-G

[23] STEIN, M. L. (1997). Efficiency of linear predictors for periodic processes using an incorrect covariance function. *J*. *Statist*. *Plann*. *Inference* **58** 321–331. MR1450019 https://doi.org/10.1016/S0378-3758(96)00088-2

[24] STEIN, M. L. (1999). Predicting random fields with increasing dense observations. *Ann*. *Appl*. *Probab*. **9** 242–273. MR1682572 https://doi.org/10.1214/aoap/1029962604

[25] STEIN, M. L. (1999). *Interpolation of Spatial Data*: *Some Theory for Kriging*. *Springer Series in Statistics*. Springer, New York. MR1697409 https://doi.org/10.1007/978-1-4612-1494-6

[26] STEINWART, I. and SCOVEL, C. (2012). Mercer's theorem on general domains: On the interaction between measures, kernels, and RKHSs. *Constr*. *Approx*. **35** 363–417. MR2914365 https://doi.org/10.1007/s00365-012-9153-3

[27] WHITTLE, P. (1963). Stochastic processes in several dimensions. *Bull*. *Int*. *Stat*. *Inst*. **40** 974–994. MR0173287

[28] ZHANG, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *J*. *Amer*. *Statist*. *Assoc*. **99** 250–261. MR2054303 https://doi.org/10.1198/016214504000000241