2021

Change detection, a functional analysis perspective

CHANGE DETECTION: A FUNCTIONAL ANALYSIS PERSPECTIVE

JULIO E. CASTRILLÓN-CANDÁS[‡], MARK KON[‡]

Abstract. We develop a new approach for detecting changes in the behavior of stochastic processes and random fields based on tensor product representations such as the Karhunen-Loève expansion. From the associated eigenspaces of the covariance operator a series of nested function spaces are constructed, allowing detection of signals lying in orthogonal subspaces. In particular this can succeed even if the stochastic behavior of the signal changes either in a global or local sense. A mathematical approach is developed to locate and measure sizes of extraneous components based on construction of multilevel nested subspaces. We show examples in $\mathbb{R}$ and on a spherical domain $\mathbb{S}^2$. However, the method is flexible, allowing the detection of orthogonal signals on general topologies, including spatio-temporal domains.

Keywords: Hilbert spaces, Karhunen-Loève Expansions, Stochastic Processes, Random Fields, Multilevel spaces, Optimization

1. Introduction

Change detection is an important topic in statistics and has received much attention, particularly in the context of time series and break detection (see the literature review in [3, 17]). There are many approaches to this problem, including a posteriori change point analysis [10, 2, 18]. Other directions concentrate on parameter changes [27, 14, 25, 20]. More recently, an avenue based on tracking changes in a linear model was proposed in [8] and extended in [15, 12, 1, 34, 19, 28]. This direction has been recently expanded to problems in $\mathbb{R}^d$ [11] and combined with ideas involving self-normalization [31, 30, 35].

In this paper an orthogonal and new direction to change detection is developed, framed in the context of functional analysis and tensor product representations such as the Karhunen-Loève [23] expansion. This method is very different from the previous approaches – Karhunen-Loève (KL) expansions are an important method for representating stochastic processes and random fields, forming optimal tensor product representations. Due to the generality of this approach, a large class of processes and fields can be represented with high accuracy. Detection is achieved by constructing nested subspaces adapted to eigenspaces of truncated KL expansions.

In Section 2 the mathematical background is discussed. In particular, the KL expansion of a stochastic process is defined. In Section 3 the theory of change detection via application of nested function spaces is developed and applied to stochastic process examples. In Section 4 an algorithm for the construction of these spaces is shown in detail. This method is very general, allowing construction of multilevel bases on very general simplicial complex domains. An example application of this method to Spherical Fractional Brownian Motion (SFBM) is shown in Section 5.

2. Mathematical Background

The Karhunen Loève expansion is an important methodology that represents random fields in terms of spatial-stochastic tensor expansions. It has been shown to be optimal in several ways, making it attractive for analysis of random fields.

‡ Department of Mathematics and Statistics, Boston University, Boston, MA
E-mail address: jcandas@bu.edu, mkon@bu.edu.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, with Ω a set of outcomes, and $\mathcal{F}$ a σ-algebra of events equipped with the probability measure $\mathbb{P}$. Let U be a domain of $\mathbb{R}^d$ and $L^2(U)$ be the Hilbert space of all square integrable functions $v : U \to \mathbb{R}$ equipped with the standard inner product

$$\langle u, v \rangle = \int_U uv \, \mathrm{d}\mathbf{x},$$

for all $u(\mathbf{x}), v(\mathbf{x}) \in L^2(U)$. In addition, let $L^2_{\mathbb{P}}(\Omega; L^2(U))$ be the space of all functions $v : \Omega \to L^2(U)$ equipped with the inner product

$$\langle u, v \rangle_{L^2_{\mathbb{P}}(\Omega; L^2(U))} = \int_\Omega \langle u, v \rangle \, \mathrm{d}\mathbb{P},$$

for all $u, v \in L^2_{\mathbb{P}}(\Omega; L^2(U))$.

Definition 1.

- *Suppose* $v \in L^2_{\mathbb{P}}(\Omega; L^2(U))$. *Then* $E_v := \mathbb{E}[v]$ *is denoted as the mean of* v.
- *For* $v \in L^2_{\mathbb{P}}(\Omega; L^2(U))$ *define the covariance function as*

$$\mathrm{Cov}(v(\mathbf{x}, \omega), v(\mathbf{y}, \omega)) := \mathbb{E}\left[(v(\mathbf{x}, \omega) - \mathbb{E}[v(\mathbf{x}, \omega)])(v(\mathbf{y}, \omega) - \mathbb{E}[v(\mathbf{y}, \omega)])\right].$$

From the properties of Bochner integrals (see [22]) we have that $E_v \in L^2(U)$ and that the covariance function $\mathrm{Cov}(\mathbf{x}, \mathbf{y}) \in L^2(U \times U)$. Define the operator $T : L^2(U) \to L^2(U)$

$$T(u)(\mathbf{x}) := \int_U \mathrm{Cov}(\mathbf{x}, \mathbf{y}) u(\mathbf{y}) \, \mathrm{d}\mathbf{y}$$

for all $u \in L^2(U)$. From Lemma 2 and Theorem 1 in [13] there exists an orthonormal set of eigenfunctions $\{\phi_k\}_{k \in \mathbb{N}}$, where $\phi_k \in L^2(U)$ and a sequence of non-negative eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots$ such that $T\phi_k = \lambda_k \phi_k$ for all $k \in \mathbb{N}$.

If $v \in L^2(\Omega; L^2(U))$ then from Proposition 2.8 in [29] the random field v can be represented in terms of the *Karhunen-Loève* (KL) tensor product expansion

$$v(\mathbf{x}, \omega) = E_v + \sum_{k \in \mathbb{N}} \lambda_k^{\frac{1}{2}} \phi_k(\mathbf{x}) Y_k(\omega), \tag{1}$$

where $\mathbb{E}[Y_k Y_l] = \delta_{kl}$ and $\mathbb{E}[Y_k] = 0$ for all $k, l \in \mathbb{N}$.

We now focus our attention on the truncated KL expansion, as this will be important in the construction of change detection filters. For any $M \in \mathbb{N}$ it is not hard to show that

$$\|v(\mathbf{x}, \omega) - E_v - \sum_{k=1}^{M} \lambda_k^{\frac{1}{2}} \phi_k(\mathbf{x}) Y_k(\omega)\|_{L^2_{\mathbb{P}}(\Omega; L^2(U))} = \left(\sum_{k \geq M+1} \lambda_k \right)^{\frac{1}{2}}. \tag{2}$$

Thus the decay of the eigenvalues controls the error of the representation. Additionally, the truncated KL expansion has the property of being optimal i.e. no other expansion of the same form is better in a sense to be specified.

To examine this further we note that from the definitions, $L^2_{\mathbb{P}}(\Omega; L^2(U))$ is isomorphic to the tensor product space $L^2_{\mathbb{P}}(\Omega) \otimes L^2(U)$. Suppose that $H_M \subset L^2(U)$ is a finite dimensional subspace of $L^2(U)$ such that $\dim H_M = M$ and $P_{H_M \otimes L^2_{\mathbb{P}}(\Omega)} : L^2(U) \otimes L^2_{\mathbb{P}}(\Omega) \to H_M \otimes L^2_{\mathbb{P}}(\Omega)$ is an orthogonal projection operator. Suppose $f \in L^2(U) \otimes L^2_{\mathbb{P}}(\Omega)$, with $\mathbb{E}[f] = 0$, then from Theorem 2.7 in [29]

$$\inf_{\substack{H_M \subset L^2(U) \\ \dim S = M}} \|f - P_{H_M \otimes L^2_{\mathbb{P}}(\Omega)} f\|_{L^2_{\mathbb{P}}(\Omega) \otimes L^2(U)} = \left(\sum_{k \geq M+1} \lambda_k \right)^{\frac{1}{2}}$$

where the infimum is achieved when $H_M = \mathrm{span}\{\phi_1, \dots, \phi_M\}$.

The optimal expansion of any random field $v \in L^2_{\mathbb{P}}(\Omega; L^2(U))$ will depend on the smoothness of v, which will have a direct impact on the decay of the eigenvalues $\{\lambda_k\}_{k=1}^{\infty}$. Consider the Sobolev space $H^p(U)$ with $p > 0$, and its dual space $\tilde{H}^{-p}(U)$. For functions $v \in L^2_{\mathbb{P}}(\Omega; H^p(U))$ (almost surely p-Sobolev smooth) we have as a consequence of Theorem 2 in [13]):

Theorem 2.1. *If $v \in L^2_{\mathbb{P}}(\Omega; H^p(U))$, then the eigenvalues $\{\lambda_k\}_{k=1}^{\infty}$ of the covariance operator $T : \tilde{H}^{-p}(U) \to H^p(U)$ satisfy $\lambda_k \leq Ck^{-2p}$ for some constant $C > 0$ independent of k and p.*

Example 1. (*Brownian Motion*) Suppose that $U = [0, 1]$ and W_t is the Wiener process with covariance function $\mathrm{Cov}(t, s) = \min\{s, t\}$. The KL expansion of W_t requires solving for eigenspaces defined by

$$\int_U \min\{s, t\}\phi_k(s)\, ds = \lambda_k \phi_k(t).$$

For this type of stochastic process it is possible to analytically solve for the eigenpair (λ_k, ϕ_k) for all $k \in \mathbb{N}$. In [33] it is shown that $\lambda_k = \frac{4}{(2k-1)^2\pi^2}$, $\phi_k(t) = \sqrt{2}\sin(t/\sqrt{\lambda_k})$ and $Y_k(\omega) \sim \mathcal{N}(0, 1)$ i.i.d. Thus we have

$$W_t = \sqrt{2}\sum_{k\geq 1} \frac{2}{(2k-1)\pi} \sin((k - 1/2)\pi t)Y_k(\omega).$$

Remark 1. In general the KL expansion can be difficult to obtain. In particular, for non-Gaussian processes the random coefficients $\{Y_k(\omega)\}_{k=1}^{M}$ are not generally independent. However, as will be shown in Section 3, to build a change detection filter it is not necessary to explicitly obtain the random coefficients $\{Y_k(\omega)\}_{k=1}^{M}$. It is sufficient only to characterize the eigenspaces from the decomposition $\{(\lambda_k, \phi_k)\}_{k=1}^{M}$. In practice, from a set of realizations of $v(\mathbf{x}, \omega) \in L^2_{\mathbb{P}}(\Omega; L^2(U))$ the pair (λ_k, ϕ_k), for $k = 1, \ldots, M$, can be estimated empirically using the method of snapshots [5].

3. Multilevel orthogonal eigenspaces

Our main goal is to detect signals defined on the domain U that do not belong to the family of finite dimensional truncated KL expansions

$$v_M(\mathbf{x}, \omega) - E_v = \sum_{k=1}^{M} \lambda_k^{\frac{1}{2}} \phi_k(\mathbf{x})Y_k(\omega). \tag{3}$$

To be more precise, we seek to detect signals that are orthogonal to the eigenspace spanned by $\{\phi_1, \ldots, \phi_M\}$ in a local and/or global sense.

Assumption 1. *Without loss of generality assume that $E_v = 0$, and consider a sequence of nested subspaces $V_0 \subset V_1 \cdots \subset L^2(U)$ such that*

$$\overline{\bigcup_{k\in\mathbb{N}_0} V_k} = L^2(U)$$

and $V_0 := \mathrm{span}\{\phi_1, \phi_2, \ldots, \phi_M\}$. Furthermore, let the subspaces $W_k \subset L^2(U)$, for $k = 0, 1, 2, \ldots$, be defined by $V_{k+1} = V_k \oplus W_k$ (all direct sums are orthogonal), so that

$$\overline{V_0 \bigoplus_{k\in\mathbb{N}_0} W_k} = L^2(U).$$

Proposition 1. *For all $k \in \mathbb{N}_0$ and any function $\psi \in W_k$ we have*

$$\int_U \phi_l \psi \, \mathrm{d}\mathbf{x} = 0$$

for $l = 1, \ldots, M$.

Proof. Since $V_k = V_{k-1} \oplus W_{k-1}$, it follows $W_k \perp V_0$. $\square$

Suppose $v \in L^2_{\mathbb{P}}(\Omega; L^2(U))$. Then the projection of v onto the multilevel spaces $\{W_k\}_{k\in\mathbb{N}}$ characterizes the signal in terms of components orthogonal to the eigenspace V_0. Given that computational power is limited, we seek to construct the space V_0 from the eigenfunctions $\phi_1, \phi_2, \ldots, \phi_M$ so that

$$\|v(\mathbf{x},\omega) - v_M(\mathbf{x},\omega)\|_{L^2_{\mathbb{P}}(\Omega;L^2(U))} = \left(\sum_{k \geq M+1} \lambda_k \right)^{\frac{1}{2}} \leq \mathtt{tol}$$

for a desired tolerance $\mathtt{tol} > 0$. The choice of $M \in \mathbb{N}$ has a direct impact on the magnitude of the projection of v onto the sum of the remainder spaces $\{W_k\}_{k\in\mathbb{N}_0}$. We shall now study the effect of truncating the KL expansion of $v \in L^2_{\mathbb{P}}(\Omega; L^2(U))$ onto the above projections.

Remark 2. The following discussion is also applicable to other non-KL expansions of random fields of the same form. The coefficients λ_k and functions ϕ_k do not necessarily need to be restricted to the eigenvalue decomposition of the covariance operator $T : L^2(U) \to L^2(U)$. However, for simplicity of the exposition, for non-KL expansions we use the same notation and it is still assumed that $\lambda_1 \geq \lambda_2 \geq \cdots > 0$ and $\phi_1, \phi_2, \ldots$ form an orthogonal set. In the rest of the paper we assume KL expansion unless otherwise noted.

Assumption 2. *For all $l \in \mathbb{N}_0$ let $\{\{\psi^l_k\}_{k=1}^{M_l}\}_{l\in\mathbb{N}_0}$ be a collection of orthonormal functions with $W_l = \text{span}\{\psi^l_1, \ldots, \psi^l_{M_l}\}$ and $M_l := \dim W_l$.*

Since the $L^2(U)$ basis $\{\{\psi^l_k\}_{k=1}^{M_l}\}_{l\in\mathbb{N}}$ is orthonormal, for any function $g \in L^2(U)$ the orthogonal projection coefficient onto the function $\psi^l_k \in W_l$ is

$$d^l_k := \int_U g\psi^l_k \, d\mathbf{x}. \tag{4}$$

We will now study the effect of the truncation parameter M on the projection coefficients on the spaces W_l for $l \in \mathbb{N}_0$.

Theorem 3.1. *Suppose that $v \in L^2_{\mathbb{P}}(\Omega; L^2(U))$ with KL expansion*

$$v(\mathbf{x},\omega) = \sum_{p\in\mathbb{N}} \lambda_p^{\frac{1}{2}} \phi_p(\mathbf{x}) Y_p(\omega).$$

Then for all $l \in \mathbb{N}_0$, $k = 1, \ldots, M_l$ and projection coefficients

$$d^l_k(\omega) = \int_U v(\mathbf{x},\omega)\psi^l_k \, d\mathbf{x}$$

a.s. then

$$\mathbb{E}\left[d^l_k\right] = 0 \quad \text{and} \quad \mathbb{E}\left[(d^l_k)^2\right] \leq \sum_{i\geq M+1} \lambda_i.$$

Proof. $\mathbb{E}\left[d^l_k\right] = 0$ follows trivially from $\mathbb{E}\left[Y_i\right] = 0$ for all $i \in \mathbb{N}_0$. From Proposition 1 and $W_k \perp V_0$ we have that

$$\mathbb{E}\left[(d^l_k)^2\right] = \mathbb{E}\left[\left(\int_U v\psi^l_k \, d\mathbf{x}\right)^2\right] = \mathbb{E}\left[\left(\int_U \sum_{i\geq M+1} \lambda_i^{\frac{1}{2}} \phi_i(\mathbf{x})\psi^l_k(\mathbf{x}) Y_i(\omega) \, d\mathbf{x}\right)^2\right]$$

$$= \mathbb{E}\left[\sum_{i\geq M+1}\sum_{j\geq M+1} Y_i(\omega)Y_j(\omega)\lambda_i^{\frac{1}{2}}\lambda_j^{\frac{1}{2}} \int_U \phi_i(\mathbf{x})\psi^l_k(\mathbf{x}) \, d\mathbf{x} \int_U \phi_j(\mathbf{x})\psi^l_k(\mathbf{x}) \, d\mathbf{x}\right]$$

From the property that $\mathbb{E}[Y_i Y_j] = \delta_{ij}$ $(i, j \in \mathbb{N}_0)$, we have that

$$\mathbb{E}\left[(d_k^l)^2\right] = \sum_{i \geq M+1} \lambda_i \int_U \phi_i(\mathbf{x}) \psi_k^l(\mathbf{x}) \, \mathrm{d}\mathbf{x} \int_U \phi_i(\mathbf{x}) \psi_k^l(\mathbf{x}) \, \mathrm{d}\mathbf{x}.$$

The conclusion follows from Cauchy-Schwartz and the orthonormality of the basis $\{\{\psi_k^l\}_{k=1}^{M_l}\}_{l=0}^{\infty}$. □

Remark 3. As M increases not only the approximation error of the KL expansion is reduced and dominated by the sum of eigenvalues, but the variance of the coefficients d_k^l is also controlled by the same quantity. Thus by using the Chebyshev inequality the projection coefficients on W_l converge in probability to 0 if $\left(\sum_{i \geq M+1} \lambda_i\right) \to 0$. The significance of this theorem is t hat each of the projection coefficients becomes more deterministic as the truncation parameter M is increased.

3.1. **Global detector.** Suppose that d_k^l are the projection coefficients of a novel signal $u(\mathbf{x}, \omega) \in L^2_{\mathbb{P}}(\Omega; L^2(U))$. The coefficients d_k^l provide a mechanism to detect the magnitude of the novel part of the signal that is orthogonal to the eigenspace V_0. In more colloquial terms, we desire to detect the components of $u(\mathbf{x}, \omega)$ with stochastic properties that are different from the eigenspace.

Suppose that $u(\mathbf{x}, \omega) = v_M(\mathbf{x}, \omega) + w(\mathbf{x}, \omega)$ i.e. the signal $u(\mathbf{x}, \omega)$ is formed from the components $v_M(\mathbf{x}, \omega) \in V_0$ and $w(\mathbf{x}, \omega) \in V_0^{\perp}$ (See Figure 1). The goal then is to detect the orthogonal component $w(\mathbf{x}, \omega)$ that does not belong in the eigenspace V_0.

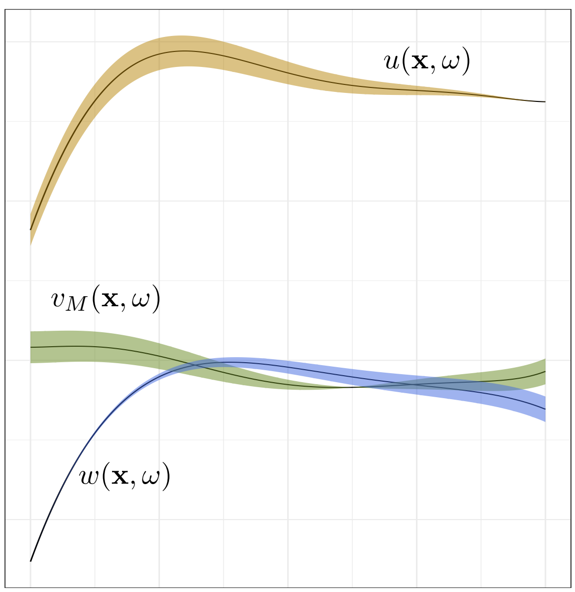


FIGURE 1. Decomposition of the signal $u(\mathbf{x}, \omega)$ into the two orthogonal components $v_M(\mathbf{x}, \omega)$ and $w(\mathbf{x}, \omega)$. Given the known random field $v_M(\mathbf{x}, \omega) \in V_0$ a.s. in Ω the objective is to detect the orthogonal component $w(\mathbf{x}, \omega) \in V_0^{\perp}$ a.s..

In the following theorems it is assumed that d_k^l are defined as in equation (4).

Theorem 3.2. *Suppose that $u(\mathbf{x}, \omega) = v_M(\mathbf{x}, \omega) + w(\mathbf{x}, \omega)$ for some $w(\mathbf{x}, \omega) \in L^2_{\mathbb{P}}(\Omega; L^2(U))$ and $w(\mathbf{x}, \omega) \perp V_0$ almost surely. Then, almost surely,*

$$\sum_{l \in \mathbb{N}_0} \sum_{k=1}^{M_l} (d_k^l)^2 = \|w(\mathbf{x}, \omega)\|^2_{L^2(U)}$$

and

$$\sum_{l \in \mathbb{N}_0} \sum_{k=1}^{M_l} \mathbb{E}\left[(d^l_k)^2\right] = \|w(\mathbf{x},\omega)\|^2_{L^2_{\mathbb{P}}(\Omega;L^2(U))}.$$

Proof. This is immediate from the fact that $\{\{\psi^l_k\}_{k=1}^{M_l}\}_{l=0}^{\infty}$ are orthonormal. □

The implication of the previous theorem is that the orthogonal signal $w(\mathbf{x},\omega)$ can be determined exactly from the projection coefficients in the space $V_0^{\perp}$.

We now study the effect of replacing the finite dimensional signal $v_M(\mathbf{x},\omega)$ with the full KL expansion of the signal. In many cases the infinite dimensional signal will be provide a more informative and useful model, including for KL expansions of Gaussian random fields. Nevertheless, in practical situations computational limitations among others will limit the dimensionality of the eigenspace V_0 to a finite level M. The detector coefficients d^l_k will be affected by the magnitude of the truncation of the KL expansion. However, as M increases the error due to the truncation will rapidly decay as the sum of the remaining eigenvalues.

Theorem 3.3. *Let* $t_M := \sum_{j \geq M+1} \lambda_j$ *and suppose that* $u(\mathbf{x},\omega) = v(\mathbf{x},\omega) + w(\mathbf{x},\omega)$ *for some* $w(\mathbf{x},\omega) \in L^2_{\mathbb{P}}(\Omega; L^2(U))$*, with* $w(\mathbf{x},\omega) \perp V_0$ *almost surely. Then*

$$\|w(\mathbf{x},\omega)\|^2_{L^2_{\mathbb{P}}(\Omega;L^2(U))}(1-2t_M) + t_M \leq \sum_{l \in \mathbb{N}_0} \sum_{k=1}^{M_l} \mathbb{E}\left[(d^l_k)^2\right] \leq \|w(\mathbf{x},\omega)\|^2_{L^2_{\mathbb{P}}(\Omega;L^2(U))}(1+2t_M) + t_M.$$

Proof. Let $P : L^2(U) \to V_0^{\perp}$ be the orthogonal projection. Since

$$u(\mathbf{x},\omega) = v_M(\mathbf{x},\omega) + \sum_{p \geq M+1} \lambda_p^{\frac{1}{2}} \phi_p(\mathbf{x}) Y_p(\omega) + w(\mathbf{x},\omega),$$

it follows

$$Pu(\mathbf{x},\omega) = \sum_{p \geq M+1} \lambda_p^{\frac{1}{2}} \phi_p(\mathbf{x}) Y_p(\omega) + w(\mathbf{x},\omega).$$

Given that $\{\{\psi^l_k\}_{k=1}^{M_l}\}_{l=0}^{\infty}$ forms an orthonormal basis for $V_0^{\perp}$, we have

$$\begin{aligned}
\sum_{l \in \mathbb{N}_0} \sum_{k=1}^{M_l} \mathbb{E}\left[d^l_k\right]^2 &= \|Pu(\mathbf{x},\omega)\|^2_{L^2_{\mathbb{P}}(\Omega;L^2(U))} = \|\sum_{p \geq M+1} \lambda_p^{\frac{1}{2}} \phi_p(\mathbf{x}) Y_p(\omega) + w(\mathbf{x},\omega)\|^2_{L^2_{\mathbb{P}}(\Omega;L^2(U))} \\
&= \|\sum_{p \geq M+1} \lambda_p^{\frac{1}{2}} \phi_p(\mathbf{x}) Y_p(\omega)\|^2_{L^2_{\mathbb{P}}(\Omega;L^2(U))} + \|w(\mathbf{x},\omega)\|^2_{L^2_{\mathbb{P}}(\Omega;L^2(U))} \\
&+ 2\mathbb{E}\left[\int_U \Big(\sum_{p \geq M+1} \lambda_p^{\frac{1}{2}} \phi_p(\mathbf{x}) Y_p(\omega)\Big) w(\mathbf{x},\omega)\, \mathrm{d}\mathbf{x}\right],
\end{aligned} \tag{5}$$

$$\|\sum_{p \geq M+1} \lambda_p^{\frac{1}{2}} \phi_p(\mathbf{x}) Y_p(\omega)\|^2_{L^2_{\mathbb{P}}(\Omega;L^2(U))} = t_M. \tag{6}$$

From Cauchy-Schwartz (both with respect to the probability and the Lesbegue measure) we have

$$
\begin{aligned}
\left|\mathbb{E}\left[\int_U \phi_p(\mathbf{x})Y_p(\omega)w(\mathbf{x},\omega)\,\mathrm{d}\mathbf{x}\right]\right| &\leq \int_U \|\phi_p(\mathbf{x})Y_p(\omega)\|_{L^2_{\mathbb{P}}(\Omega)}\|w(\mathbf{x},\omega)\|_{L^2_{\mathbb{P}}(\Omega)}\,\mathrm{d}\mathbf{x} \\
&= \int_U |\phi_p(\mathbf{x})|\|w(\mathbf{x},\omega)\|_{L^2_{\mathbb{P}}(\Omega)}\mathrm{d}\mathbf{x} \\
&\leq \|\phi_p(\mathbf{x})\|_{L^2(U)}\|w(\mathbf{x},\omega)\|_{L^2_{\mathbb{P}}(\Omega;L^2(U))} \\
&= \|w(\mathbf{x},\omega)\|_{L^2_{\mathbb{P}}(\Omega;L^2(U))}.
\end{aligned} \tag{7}
$$

Inserting equations (7) and (6) in (5) we reach the conclusion. □

Remark 4. The previous theorem provided a mechanism to determine the intensity of the orthogonal signal $w(\mathbf{x},\omega)$ given the size of the truncation parameter M. Thus the larger M is, depending on the decay of λ_p, $p = M+1,\dots$, the size of the coefficients d^l_k can be used to determine more precisely the size of the perturbation $w(\mathbf{x},\omega)$ both in the local and global sense.

Theorem 3.4. *Suppose that a signal* $u(\mathbf{x},\omega) = v_M(\mathbf{x},\omega) + w(\mathbf{x},\omega)$ *for* $v_M \in V_0$*, and for some* $w(\mathbf{x},\omega) \in L^2_{\mathbb{P}}(\Omega; L^2(U))$ *then*

$$
Pu(\mathbf{x},\omega) = Pw(\mathbf{x},\omega) = \sum_{l\in\mathbb{N}_0}\sum_{k=1}^{M_l} d^l_k(\omega)\psi^l_k(\mathbf{x}) \text{ and } \sum_{l\in\mathbb{N}_0}\sum_{k=1}^{M_l}\mathbb{E}\left[(d^l_k)^2\right] = \|Pw(\mathbf{x},\omega)\|^2_{L^2_{\mathbb{P}}(\Omega;L^2(U))},
$$

where $P : L^2(U) \to V_0^{\perp}$ *is the orthogonal projection.*

Proof. Immediate □

Remark 5. The sharpness of the bound in Theorem 3.3 depends on the decay of the coefficients $\lambda_{M+1}, \lambda_{M+2}, \dots$. From Theorem 2.1 the decay of the eigenvalues is related to the smoothness of the realization $v(\cdot,\omega) \in H^p(U)$, where $p \in \mathbb{N}_0$, from which we have $t_M \leq C\sum_{k=M+1}^{\infty} k^{-2p}$, with C independent of k and p. For $p > 0$ we have that $t_M \approx (M+1)^{-2p}$.

3.2. **Global - local detector.** Suppose that $\{\chi^l_k\}_{k\in\mathcal{K}(l)}$ is a finite disjoint partition at level l of the domain U such that

$$
U = \bigcup_{k\in\mathcal{K}(l)} \chi^l_k,
$$

where $\mathcal{K}(l)$ is an index set corresponding to each element in the the partition. For each $l \in \mathbb{N}_0$ let $U^l := \{\chi^l_k\}_{k\in\mathcal{K}(l)}$ and assume that U^{l+1} is refinement of U^l. We make the assumption that the support of any basis function $\psi^l_k \in W_l$ is given by a union of sets in U^l. In Section 4 the construction of the finite dimensional spaces $V_0,\dots V_n$ and $W_0,\dots W_{n-1}$ with compactly supported basis functions will be described in detail.

Suppose that $\tilde{U} \subset U$ and let

$$
\mathcal{C}(\tilde{U}) := \{(k,l);\ l \in \mathbb{N}_0, k \in \mathcal{K}(l), \operatorname{supp}\tilde{U} \cap \operatorname{supp}\psi^l_k \neq \emptyset\},
$$

where by $\operatorname{supp}\tilde{U}$ we mean the union of all sets in $\tilde{U}$. It is clear that any projection coefficient $d^l_k = 0$ (see eq. (4)) if $(l,k) \notin \mathcal{C}(\tilde{U})$. Thus the coefficients that correspond to the intersection of the domain $\tilde{U}$ and the support of the basis functions $\{\{\psi^l_k\}_{k=1}^{M_l}\}_{l\in\mathbb{N}_0}$ are sufficient to detect any changes on $\tilde{U}$. The following example explores this property in more detail.

Example 2. Let $v_M(x,\omega) = 1 + Y_1(\omega)\left(\frac{\sqrt{\pi}L}{2}\right)^{\frac{1}{2}} + \sum_{k=2}^{M} \lambda_k^{\frac{1}{2}} \phi_k(x) Y_k(\omega)$ be a stochastic process defined on the domain $[0,\tau]$ where

$$\phi_k(x) := \begin{cases} \sin\left(\frac{\lfloor \frac{k}{2} \rfloor \pi x}{L_p}\right) & \text{if } k \text{ is even} \\ \cos\left(\frac{\lfloor \frac{k}{2} \rfloor \pi x}{L_p}\right) & \text{if } k \text{ is odd} \end{cases}$$

are orthogonal and

$$\sqrt{\lambda_k} := (\sqrt{\pi}L)^{\frac{1}{2}} \exp\left(-\frac{(\lfloor \frac{k}{2} \rfloor \pi L)^2}{8}\right).$$

The random variables $Y_1, \dots, Y_M$ are assumed to be uniform in $[-\sqrt{3}, \sqrt{3}]$ and independently identically distributed. In [26] the authors show that $v_M(x,\omega)$ is the truncation of the infinite dimensional random field v with the covariance:

$$\text{Cov}(v(x,\omega), v(y,\omega)) = \exp\left(-\frac{(x-y)^2}{L_c^2}\right), \tag{8}$$

where $L_p = \max\{\tau, 2L_c\}$ is the length correlation and $L = L_c/L_p$. Suppose that $\tau = 1$ and $L_c = 0.01$, so that $L_p = 1$ and $L = 0.01$. Furthermore, let $u(x,\omega) = v_M(x,\omega) + w(x)$ where $\tilde{U} = [0.3, 0.7]$, $w(x) = \mathbb{1}_{\tilde{U}} 0.05 \exp\left(-\frac{(x-x_s)^2}{\sigma^2}\right)$, $x_s = 0.5$ and $\sigma = 10^{-3/2}$. The compactly supported *multilevel basis* (MB) for $V_0 \oplus W_0 \oplus W_1 \oplus \dots$ is constructed such that V_0 is the span of $\{\phi_1, \dots, \phi_M\}$. The objective now is to detect the presence of the smooth Gaussian function $w(x)$ given the signal $u(x,\omega)$.

In Figure 2 (a) the signals $u(x,\omega)$, $v_M(x,\omega)$ (solid and dashed lines with left vertical axis) and $w(x)$ (solid orange line with right vertical axis) are plotted. The deformation $w(x)$ is added to the baseline stochastic process $v_M(x,\omega)$ to obtain $u(x,\omega)$. In Figure 2 (b), (c) and (d) the projection coefficients of the signal $u(x,\omega)$ onto the multilevel spaces W_{n-1}, W_{n-2} and W_{n-3}, for $n = 6$, are plotted. Notice that around the center of $w(x)$ at $x_s = 0.5$ the projection coefficients are clearly non-zero. Thus the local presence of the non-zero coefficients indicated that around $x_s = 0.5$ the usual behavior of the signal $v_M(x,\omega)$ changes. Furthermore, from Theorem 3.3 we can conclude that

$$Pw(\mathbf{x}) = \sum_{(k,l) \in \mathcal{C}(\tilde{U})} d_k^l \psi_k^l(x) \text{ and } \sum_{(k,l) \in \mathcal{C}(\tilde{U})} (d_k^l)^2 = \|Pw(x)\|_{L^2(U)}^2,$$

where P is as in Theorem 3.4.

Remark 6. In practice we assume that $u(\mathbf{x},\omega)$, $v_M(\mathbf{x},\omega)$ and $w(\mathbf{x},\omega)$ belong in the finite dimensional space $V_n = V_0 \oplus W_0 \oplus \dots \oplus W_{n-1}$ for some finite fixed $n \in \mathbb{N}_0$. For example, V_n can be the span of disjoint characteristic functions (Haar basis) over the domain U. This is a reasonable assumption since data are collected as samples. In Example 2 the signal is formed from 500 equally spaced samples of $v_M(x,\omega)$ from $[0,1]$, and we have chosen $n = 6$ in V_n above.

Example 3. The multilevel approach can also be applied to non-KL expansions, such as the following. Let $v_M(x,\omega) = 1 + Y_1(\omega)\left(\frac{\sqrt{\pi}L}{2}\right)^{1/2} + \sum_{k=2}^{M} \sqrt{\mu_k} \phi_k(x) Y_k(\omega)$ be a stochastic process, where $x \in [0,1]$, with μ_k and $\phi_k(x)$ defined as in Example 2. However, we fix $L_p = 1/4$ and $L = 1/4$. Let $\tilde{U} = [0.3, 0.7]$ with $w(x) = \mathbb{1}_{\tilde{U}} 0.5 \exp\left(-\frac{(x-x_s)^2}{\sigma^2}\right)$, $x_s = 0.5$ and $\sigma = 10^{-3/2}$. Note that this example does not necessarily have the covariance structure as shown in Equation (8). However, these coefficients lead to a more oscillatory structure for $v_M(x,\omega)$. It is hard for the observer to distinguish $w(x)$ in $u(x,\omega)$ from $v_M(x,\omega)$ without knowledge of the location. By building the

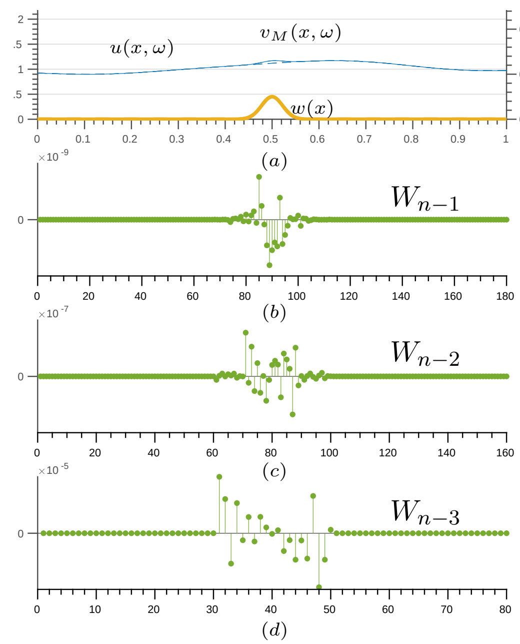


Figure 2. Example of the functional analysis approach to change detection. (a) The original signal $v_M(x,\omega)$ (dash line with left vertical axis) with the Gaussian bump $w(x)$ (solid orange line with right vertical axis) with support on the domain $\tilde{U} := [0.3, 0.7]$. Adding these two signals give $u(x,\omega)$ (solid line with left vertical axis). The multilevel basis is constructed such that W_k, $k \in \mathbb{N}_0$, is orthogonal to the signal $v_M(x,\omega)$. (b), (c) and (d) Detection of $w(x)$ at levels W_{n-1}, W_{n-2} and W_{n-3}, where $n = 6$.

multilevel spaces adapted to $v_M(x,\omega)$ the filter coefficients for levels W_{n-1}, W_{n-2} and W_{n-3} easily detect the location of $w(x)$ (See Figure 3).

4. Construction of multilevel orthogonal eigenspace

For many practical problems the domain U will be restricted to some form of a mesh. The multilevel basis (MB) of the finite dimensional spaces $V_n = V_0 \oplus W_0 \oplus \cdots \oplus W_{n-1}$ for a finite fixed $n \in \mathbb{N}_0$ can be constructed on this mesh. The construction of the MB for problems in $\mathbb{R}^3$ within the context of polynomials and integral operators was first proposed in [32]. In [6] this was modified for discrete domains in the context of Kriging and high dimensional problems from the work in [4].

Definition 2. *Suppose that $\mathcal{T}$ is a collection of N simplices in $\mathbb{R}^d$. Then $\mathcal{T}$ is a k-simplicial complex if the following properties are satisfied*

i) Every face of a simplex in $\mathcal{T}$ is also in $\mathcal{T}$.
ii) The non-empty intersection of any two simplices $\tau_1, \tau_2 \in \mathcal{T}$ is a face of both τ_1 and τ_2.
iii) The highest dimension of any simplex in $\mathcal{T}$ is $k \leq d$.

This definition allows us to construct many general domains in $\mathbb{R}^d$ formed by k-simplices. For example, in $\mathbb{R}^3$ we can build a triangulation of a surface with 2-simplices.

Definition 3. *Let $\mathbf{x}_i$ be the barycenter of any simplex $\tau_i \in \mathcal{T}$, and $\mathbb{S} := \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$.*

Assumption 3. *We make the following assumptions for $\mathcal{T}$ on the domain $U \subset \mathbb{R}^d$:*

i) $\mathcal{T}$ contains N simplices τ_i, for $i = 1, \ldots, N$, of the same order.
ii) $U = \cup_{\tau_i \in \mathcal{T}} \tau_i$.

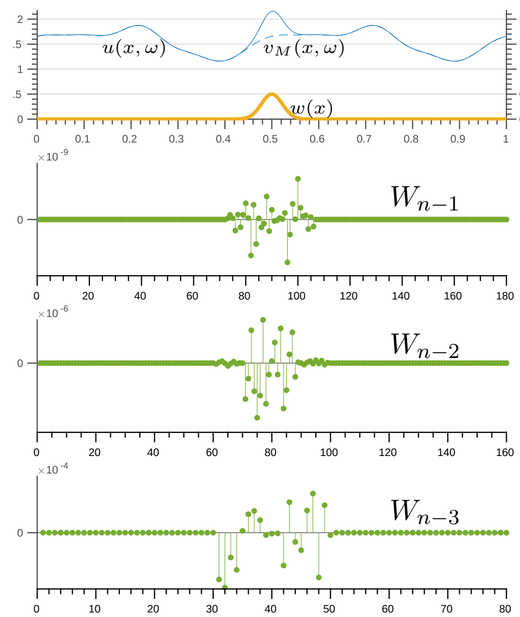


FIGURE 3. Example of change detection for $v_M(x,\omega)$ with $L = L_p = 1/4$. It is hard for the observer to distinguish $w(x)$ in $u(x,\omega)$ from $v_M(x,\omega)$ without knowledge of the location. By building the multilevel spaces adapted to $v_M(x,\omega)$ the filter coefficients for levels W_{n-1}, W_{n-2} and W_{n-3} easily detect the location of $w(x)$.

iii) For any $i = 1,\dots,N$ and for any simplex $\tau_i \in \mathcal{T}$ let $\chi_i = c_i 1_{\tau_i}$. The coefficients c_i for $i = 1,\dots,N$ are chosen such that $\chi_1,\dots,\chi_N$ form an orthonormal set in $L^2(U)$.

iv) Let $\mathcal{E} := \{\chi_1,\dots,\chi_N\}$ and $V_n = \mathcal{P}(\mathcal{E}) :=$ span$\{\chi_1,\dots,\chi_N\}$. Assume that Karhunen-Loève eigenfunctions $\phi_i \in \mathcal{P}(\mathcal{E})$ for all $i = 1,\dots,M$ where $N > M$.

With the goal of constructing a multi-level basis, the domain U is in general embedded in a kd-tree type decomposition. This will allow the MB construction algorithm to efficiently access the simplices of $\mathcal{T}$ by searching a binary tree. The approach is described in [4] in the context of discrete vector spaces. For very high dimensional domains alternative choices also include Random Projection (RP) trees [9].

Suppose that all the barycenters $\mathbf{x} \in \mathbb{S}$ are embedded in the cell $B^0_0 \subset \mathbb{R}^d$, which corresponds to the top of the binary tree. Without loss of generality it can be assumed that $B^0_0 = [0,1]^d$ and $\mathbb{S} \subset B^0_0$. Each cell B^l_k at the level l of the tree and index k contains a subset of the barycenters in $\mathbb{S}$. There is a subdivision of any cell B^l_k into two cells $B^{l-1}_{\texttt{left}}$ and $B^{l-1}_{\texttt{right}}$ according to the following rule (see Algorithm 1):

1) For each coordinate $1 \leq j \leq d$, project every barycenter $\mathbf{x}_i \in B^l_k$ onto the unit vector along coordinate k and compute the sample variance of these projection coefficients.
2) Choose the unit coordinate vector v in the direction $1 \leq j \leq d$ with the maximal sample variance for the above projection coefficients.
3) Compute the median of the projections along v and split the cell in two at this coordinate position.

The initial cell B^0_0 is subdivided in this manner until a maximum number of n_0 barycenters are located at each of the leaf cells. It is also assumed that $n_0 > M$, with M the above number of truncated KL coefficients. Let $\mathcal{B}^l$ be the collection of all the cells B^l_k at level l. Algorithm 2 below

Algorithm 1: ChooseRule($\tilde{\mathbb{S}}$) function for kD-tree where $\tilde{\mathbb{S}} \subset \mathbb{S}$.

Input: $\tilde{\mathbb{S}}$
Output: Rule, threshold, v
begin
choose a coordinate direction that has maximal variance of the projection of the points in $\tilde{\mathbb{S}}$.
Rule(x) := $x \cdot v \leq$ threshold = median

describes in more detail the construction of the binary tree. In Figure 4 an example of the kd-Tree partitioning of a triangulation $\mathcal{T}$ is shown along with the associated binary tree.

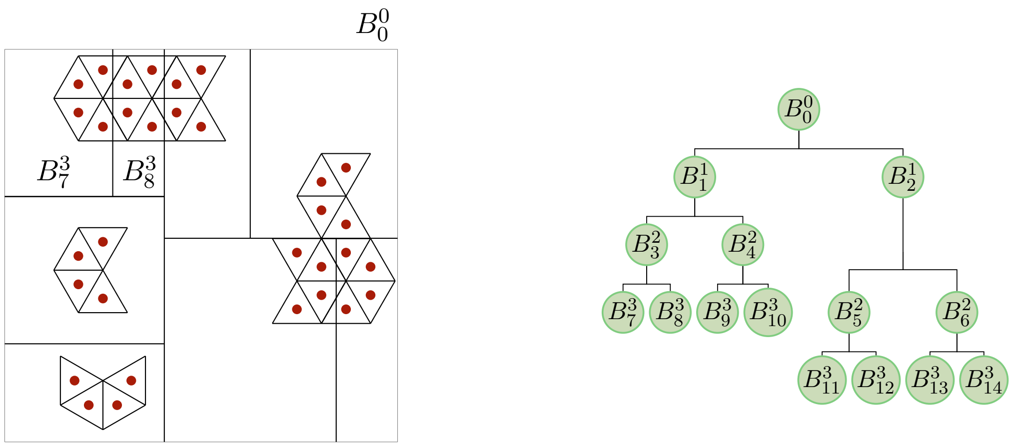


FIGURE 4. Multilevel kd-tree domain decomposition of a triangulation $\mathcal{T}$ with respective binary tree. Assume that the tree has $l = 0, \ldots, n$ levels.

Remark 7. Note that if the number of barycenters is even then the tree will end evenly at some level n. However, if the number is odd and $n_0 = 1$ then one branch can end at level n and the other at level $n - 1$.

From Assumption 3 the set of barycenter locations $\{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ have a one to one correspondence with $\mathcal{E}$, i.e. $\mathbf{x}_t \longleftrightarrow \chi_t$ for all $t = 1, \ldots, N$. The multilevel basis construction algorithm adapted to the Karhunen Loève expansion $v_M(\mathbf{x}, \omega)$ is described as follows:

(I) Start at the finest level of the tree, i.e. $q = n$.
(II) For each leaf cell $B_k^q \in \mathcal{B}^q$ assume without loss of generality that there are s barycenters $\mathbb{S}^q := \{\mathbf{x}_1, \ldots, \mathbf{x}_s\}$ with associated functions $\mathcal{E}_k^q := \{\chi_1, \ldots, \chi_s\}$. Let $\mathcal{Q}_k^q(\mathcal{E}_k^q)$ be the span of the functions in $\mathcal{E}_k^q$.
   i) Let $\boldsymbol{\phi}_j^{q,k} := \sum_{\chi_i \in \mathcal{E}_k^q} c_{i,j}^{q,k} \chi_i,\ \ j = 1, \ldots, a_{q,k};\ \ \boldsymbol{\psi}_j^{q,k} := \sum_{\chi_i \in \mathcal{E}_k^q} d_{i,j}^{q,k} \chi_i,\ \ j = a_{q,k} + 1, \ldots, s$, where $c_{i,j}^{q,k}$, $d_{i,j}^{q,k} \in \mathbb{R}$, and are undetermined at the moment, and for some yet undetermined $a_{q,k} \in \mathbb{N}_0$. The objective of the above linear combinations is to construct new functions $\boldsymbol{\psi}_j^{q,k}$ orthogonal to V_0, i.e., such that for all $\phi_i \in V_0$, $i = 1, \ldots, M$, we have that

$$\int_U \phi_i(\mathbf{x}) \boldsymbol{\psi}_j^{q,k}(\mathbf{x})\, \mathrm{d}\mathbf{x} = 0. \tag{9}$$

Algorithm 2: MakeTree($\mathbb{S}$) function

Input: $\mathbb{S}$, node, n_0
Output: Tree
begin
  if *Tree = root* **then**
    node ← 0, currentdepth ← 0 Tree ← MakeTree($\mathbb{S}$, node, currentdepth + 1, n_0)
  else
    Tree.node = node
    Tree.currentdepth = currentdepth - 1
    node ← node + 1
    if $|\tilde{\mathbb{S}}| < n_0$ **then**
      return (Leaf)
    (Rule, threshold, v) ← ChooseRule($\tilde{\mathbb{S}}$)
    (Tree.LeftTree, node) ← MakeTree($\mathbf{x} \in \tilde{\mathbb{S}}$: Rule($\mathbf{x}$) = True, node, currentdepth + 1, n_0)
    (Tree.RightTree, node) ← MakeTree($\mathbf{x} \in \tilde{\mathbb{S}}$: Rule($\mathbf{x}$) = false, node, currentdepth + 1, n_0)
    Tree.threshold = threshold
    Tree.v = v

ii) From the eigenfunctions $\phi_1, \dots, \phi_M$ of the KL expansion and $\mathcal{E}^q_k$ we can form the matrices

$$\mathbf{M}^{q,k} := \begin{bmatrix} \langle \phi_1(\mathbf{x}), \chi_1(\mathbf{x}) \rangle & \dots & \langle \phi_1(\mathbf{x}), \chi_s(\mathbf{x}) \rangle \\ \langle \phi_2(\mathbf{x}), \chi_1(\mathbf{x}) \rangle & \dots & \langle \phi_2(\mathbf{x}), \chi_s(\mathbf{x}) \rangle \\ \vdots & \ddots & \vdots \\ \langle \phi_M(\mathbf{x}), \chi_1(\mathbf{x}) \rangle & \dots & \langle \phi_M(\mathbf{x}), \chi_s(\mathbf{x}) \rangle \end{bmatrix}$$

and

$$\mathbf{N}^{q,k} := \left[\begin{array}{ccc|ccc} \langle \phi_1, \boldsymbol{\phi}^{q,k}_1 \rangle & \dots & \langle \phi_1, \boldsymbol{\phi}^{q,k}_{a_{q,k}} \rangle & \langle \phi_1, \boldsymbol{\psi}^{q,k}_{a_{q,k}+1} \rangle & \dots & \langle \phi_1, \boldsymbol{\psi}^{q,k}_s \rangle \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \langle \phi_M, \boldsymbol{\phi}^{q,k}_1 \rangle & \dots & \langle \phi_M, \boldsymbol{\phi}^{q,k}_{a_{q,k}} \rangle & \langle \phi_M, \boldsymbol{\psi}^{q,k}_{a_{q,k}+1} \rangle & \dots & \langle \phi_M, \boldsymbol{\psi}^{q,k}_s \rangle \end{array} \right]$$

where $\langle \cdot, \cdot \rangle$ is the $L^2(U)$ inner product.

iii) Apply the Singular Value Decomposition (SVD) to $\mathbf{M}^{q,k}$ i.e. $\mathbf{M}^{q,k} = \mathbf{UDV}$ where $\mathbf{U} \in \mathbb{R}^{M \times M}$, $\mathbf{D} \in \mathbb{R}^{M \times s}$, and $\mathbf{V} \in \mathbb{R}^{s \times s}$. Assume that $a_{q,k}$ is the rank of the matrix $\mathbf{M}^{q,k}$.

iv) Consider the following choice for the coefficients $c^{q,k}_{i,j}$ and $d^{q,k}_{i,j}$ from the right SVD matrix:

$$\left[\begin{array}{ccc|ccc} c^{q,k}_{0,1} & \dots & c^{q,k}_{a_{q,k},1} & d^{q,k}_{a_{q,k}+1,1} & \dots & d^{q,k}_{s,1} \\ c^{q,k}_{0,2} & \dots & c^{q,k}_{a_{q,k},2} & d^{q,k}_{a_{q,k}+1,2} & \dots & d^{q,k}_{s,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ c^{q,k}_{0,s} & \dots & c^{q,k}_{a_{q,k},s} & d^{q,k}_{a_{q,k}+1,s} & \dots & d^{q,k}_{s,s} \end{array} \right] := \mathbf{V}^{\mathrm{T}}.$$

Since the vectors (column and row) in $\mathbf{V}$ are orthonormal in an $l_2(\mathbb{R}^s)$ sense and the functions in $\mathcal{E}^q_k$ are orthonormal in $L_2(U)$ then $\boldsymbol{\phi}^{q,k}_1, \dots, \boldsymbol{\phi}^{q,k}_{a_{q,k}}$, and $\boldsymbol{\psi}^{q,k}_{a_{q,k}+1}, \dots, \boldsymbol{\psi}^{q,k}_s$ form an orthonormal basis of $\mathcal{Q}^q_k(\mathcal{E}^q_k)$. As in [32, 6] it can be shown that this choice leads to $\boldsymbol{\psi}^{q,k}_{a_{q,k}+1}, \dots, \boldsymbol{\psi}^{q,k}_s$ satisfying equation (9). From the SVD of $\mathbf{M}^{q,k}$ and this choice of coefficients

$$\mathbf{N}^{q,k} = \mathbf{M}^{q,k}\mathbf{V} = \mathbf{U}\mathbf{D}.$$

Now, decompose $\mathbf{D}$ as $\mathbf{D} = [\mathbf{\Sigma} \,|\, \mathbf{0}]$, where $\mathbf{\Sigma} \in \mathbb{R}^{M \times a_{q,k}}$ is a diagonal matrix with the non-zero singular values of $\mathbf{M}^{q,k}$ and the zero matrix $\mathbf{0} \in \mathbb{R}^{M \times (s - a_{q,k})}$. Thus $\mathbf{U}\mathbf{D} = [\mathbf{U}\mathbf{\Sigma} \,|\, \mathbf{0}]$ and $\mathbf{N}^{q,k} = \mathbf{M}^{q,k}\mathbf{V} = [\mathbf{U}\mathbf{\Sigma} \,|\, \mathbf{0}]$. It follows that SVD columns $a_{q,k}+1, \dots, s$ of $\mathbf{V}$ form an orthonormal basis of the nullspace of $\mathbf{M}^{q,k}$ and therefore $\boldsymbol{\psi}^{q,k}_{a_{q,k}+1}, \dots, \boldsymbol{\psi}^{q,k}_{s}$ satisfy equation (9), and are orthogonal to V_0 and compactly supported in the cell B^q_k.

v) Denote by $D^q_k := \{\boldsymbol{\psi}^{q,k}_{a_{q,k}+1}, \dots, \boldsymbol{\psi}^{q,k}_{s}\}$ and $C^q_k := \{\boldsymbol{\phi}^{q,k}_1, \dots, \boldsymbol{\phi}^{q,k}_{a_{q,k}}\}$.

Remark 8. Notice that the functions $\boldsymbol{\phi}^{q,k}_1, \dots, \boldsymbol{\phi}^{q,k}_{a_{q,k}}$ are *not* in general orthogonal to V_0.

(III) Let $\mathcal{D}^q := \cup_{B^q_k \in \mathcal{B}^q} D^q_k$ and $\mathcal{C}^q := \cup_{B^q_k \in \mathcal{B}^q} C^q_k$. It is not hard to see that they form an orthonormal set in $L^2(U)$. Denote by W_{q-1} the span of all the functions in $\mathcal{D}^q$ and similarly V_{q-1} with respect to $\mathcal{C}^q$.

(IV) The next step is to go to level $q-1$. For any two sibling cells denoted as $B^q_{\texttt{left}}$ and $B^q_{\texttt{right}}$ at level q denote $\mathcal{E}^{q-1}_{\tilde{k}}$, for some index $\tilde{k}$, as the collection of functions $\boldsymbol{\phi}^{q,\texttt{left}}_1, \dots, \boldsymbol{\phi}^{q,\texttt{left}}_{a_{q,\texttt{left}}}$ and $\boldsymbol{\phi}^{q,\texttt{right}}_1, \dots, \boldsymbol{\phi}^{q,\texttt{right}}_{a_{q,\texttt{right}}}$

(V) Recursively, let $q := q-1$. If $B^q_k \in \mathcal{B}^q$ is a leaf cell (which may occur as not all branches of the tree necessarily have the same numbers of levels) then go to (II). However, if $B^q_k \in \mathcal{B}^q$ is not a leaf cell, then go to (II) but replace the collection of leaf cell functions with $\mathcal{E}^q_k := \{\boldsymbol{\phi}^{q+1,\texttt{left}}_1, \dots, \boldsymbol{\phi}^{q+1,\texttt{left}}_{a_{q+1,\texttt{left}}}, \boldsymbol{\phi}^{q+1,\texttt{right}}_1, \dots, \boldsymbol{\phi}^{q+1,\texttt{right}}_{a_{q+1,\texttt{right}}}\}$.

(VI) When $q = -1$ is reached then incrementation stops.

When the algorithm terminates a series of orthogonal subspaces $W_0, \dots, W_n$ and corresponding basis functions $\mathcal{D}^0, \dots \mathcal{D}^t$ are obtained. Furthermore, it can be shown that $V_0 = \text{span}\{\phi_1, \dots \phi_M\}$ is also the span of $\{\boldsymbol{\phi}^{0,0}_1, \dots, \boldsymbol{\phi}^{0,0}_{a_{0,0}}\}$ and $a_{0,0} = M$.

Remark 9. Following the arguments in [7, 4] it can be shown that

- $V_n = \mathcal{P}(\mathcal{E}) = V_0 \oplus W_0 \oplus W_1 \oplus \cdots \oplus W_{n-1}$
- $\mathcal{C}^0$, $\mathcal{D}^0$, $\mathcal{D}^1$, $\dots$ $\mathcal{D}^{n-1}$ form an orthonormal basis for $V_0 \oplus W_0 \oplus W_1 \oplus \cdots \oplus W_{n-1}$
- At most $\mathcal{O}(Nn)$ computational steps are needed to construct the multilevel basis of V_n.
- Let $\gamma \in \mathcal{P}(\mathcal{E})$ and denote $c_1, \dots c_N$ the orthogonal projection coefficients on $\mathcal{P}(\mathcal{E})$ where

$$c_i = \int_U \gamma(\mathbf{x}) \chi_i(\mathbf{x}) \,\mathrm{d}\mathbf{x}$$

and $i = 1, \dots, N$. It can be shown that the multilevel projection coefficients d^l_k for $l = 0, \dots n-1$ can be computed in at most $\mathcal{O}(Nn)$ computational steps and memory from $c_1, \dots, c_N$.

Remark 10. For many practical situations only spatial samples of $u(\mathbf{x}, \omega)$ are available. For such cases the alternative choice for $\mathcal{E}$ is a set of unit vectors. A similar construction to the multilevel basis can be done in a vector sense (see [4, 7] for details). This is equivalent to the continuous multilevel basis construction, up to a re-scaling of the domain, by assuming that each simplex in $\mathcal{T}$ has the same unit volume measure and $u(\mathbf{x}, \omega)$ is constant on each simplex. The multilevel coefficients from Examples 2 and 3 where obtained with the discrete version of the multilevel basis using 500 equally spaced samples.

Remark 11. The algorithm is efficiently implemented in MATLAB [24] and can handle highly complex geometries. The code will be made available to the general public.

5. Spherical example

We will now demonstrate the application of the multilevel orthogonal eigenspace for the detection of signals on Spherical Fractional Brownian Motion (SFBM) defined on the unit sphere $\mathbb{S}_2$ [16]. This is a more complex scenario that shows the flexibility of this approach.

Suppose $\{P_l^m(x)\}_{l\geq 0}$ is the set of associated Legendre polynomials for $m \geq 0$. If m is negative then the associated Legendre polynomials are given by

$$P_l^{-m}(x) = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(x).$$

The coordinates of the unit sphere $\mathbb{S}_2$ are given by the colatitude $\theta \in [0,\pi)$ and longitude $\varphi \in [0, 2\pi)$. The spherical harmonics $Y_l^m(\theta,\varphi)$ on $\mathbb{S}_2$ are defined by

$$Y_l^m(\theta,\varphi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}} P_l^m(cos\,\theta)e^{im\varphi},$$

where $m = -l, \dots, l$. In [16] the author demonstrates that the Karhunen-Loève expansion of the SFBM is given by

$$v(\theta,\varphi,\omega) = \sum_{l\geq 0}\sum_{m=-l}^{l} \sqrt{-\pi d_l}\varepsilon_l^m(\omega)(Y_l^m(\theta,\varphi) - Y_l^m(0,0))$$

in the L^2 sense, where $\varepsilon_l^m \sim N(0,1)$ i.i.d. and

$$d_l := \int_{-1}^{1} \arccos(x) P_l(x)\,\mathrm{d}x.$$

Remark 12. Notice the spherical harmonics $Y_l^m(\theta,\varphi) : [0,\pi] \times [0,2\pi] \to \mathbb{C}$ are complex-valued, which however does not restrict this analysis. Although the theoretical discussion of the multilevel orthogonal eigenspaces is given for real Hilbert spaces, the method can be readily extended to the complex case. In particular, the algorithm implementation can also handle this case.

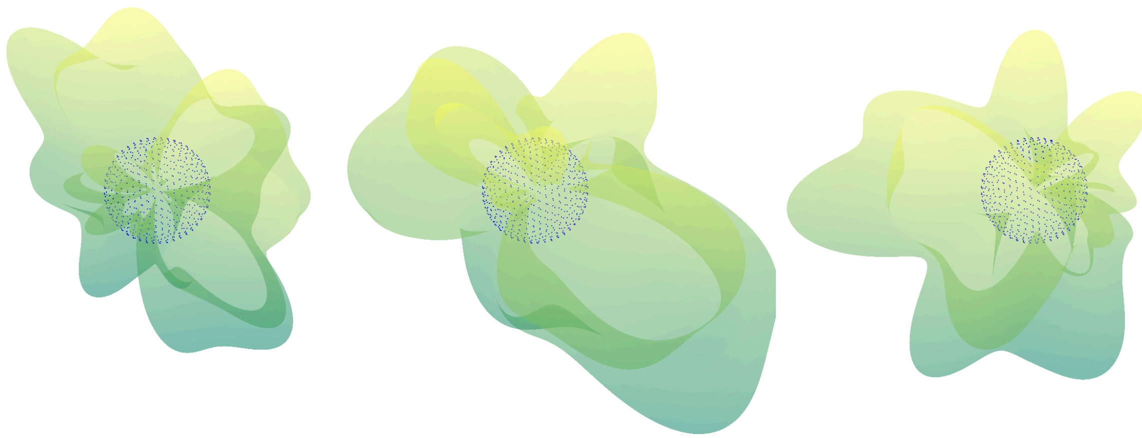
In Figure 3 three realizations from the KL expansion of the absolute value of the SFBM are shown. For this example $l = 10$, which is sufficient to capture much of the stochastic movement since it is shown in [16] that d_l decays as l^{-2}. However, since $d_l = 0$ whenever $l = 3, 5, 7, 9$ then the truncated KL expansion is reduced to $M = 56$ eigenfunctions.

Suppose we apply a perturbation to $v(\theta,\varphi,\omega)$ of the form

$$w(\theta,\varphi) = c\exp\left(-\frac{(\theta-\pi/2)^2 + (\varphi-\pi/2)^2}{\sigma^2}\right)$$

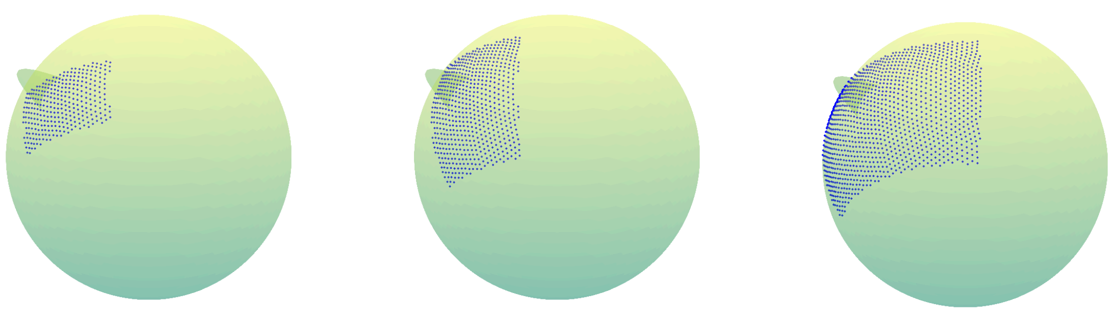
where $c = 0.5$ and $\sigma = 0.1$, i.e. $u(\theta,\varphi,\omega) = v + w$. The goal is to detect w with the ML orthogonal eigenspace. For this case we sample the sphere with 10,242 almost equally spaced barycenters [21] and construct the ML basis.

After applying the ML to the spherical signal u, we analyze the projection coefficients corresponding to the spaces W_{n-1}, W_{n-2} and W_{n-3} where $n = 8$. In Figure 6 the locations of the basis functions corresponding to the ML projection coefficients are shown for any coefficient with absolute value greater than 10^{-4}. Over imposed with on each sphere is the Gaussian perturbation $w(\phi,\vartheta)$. Notice that these coefficients essential identify the location of the Gaussian from the original signal u at different levels of resolution. Furthermore, we can use these coefficients to estimate the size of the perturbation from Theorem 3.3.



$v(\mathbf{x}, \omega_1)$ $v(\mathbf{x}, \omega_2)$ $v(\mathbf{x}, \omega_3)$

FIGURE 5. Realizations of the Spherical Fraction Brownian Motion (SFBM) from $l = 10$ ($M = 56$ eigenfunctions) Karhunen Loève expansion. The blue dots correspond to the sampling of the unit sphere.



W_{n-1} W_{n-2} W_{n-3}

FIGURE 6. Support of MB functions corresponding to projection coefficients with absolute values greater than 10^{-4}. The Gaussian perturbation $w(\theta, \varphi)$ is over imposed on the unit sphere. Notice that the support indicates the location of the detected signal $w(\phi, \varphi)$ from the $u(\phi, \varphi, \omega)$.

6. LAST COMMENTS

In this paper we have developed a new approach for change detection by applying the tools that are available to us from functional analysis. By leveraging the power of the KL and other tensor product expansions a multilevel nested functional spaces are constructed. These spaces can be used to detect extraneous signals that are orthogonal to the truncated eigenspace.

Our results show that this method is very flexible allowing the application to complex domains. Furthermore, this approach can also applied to other forms of tensor product expansions such as Polynomial Chaos Expansions (PCE). We have mostly shown examples of detecting a fixed perturbation. However, this approach can be extended to fully stochastic perturbations. Future work involves collecting samples and formulating hypothesis tests. In addition, we also envision applications to machine learning classification.

Acknowledgements: I appreciate the help and advice from Trevor Martin. In particular, for proof reading the manuscript.

References

[1] A. Aue, S. Hörmann, L. Horváth, M. Huŝková, and J. G. Steinebach. Sequential testing for the stability of high-frequency portfolio betas,. *Econometric Theory*, 28:804, 2012.

[2] A. Aue, S. Hörmann, L. Horváth, and M. Reimherr. Break detection in the covariance structure of multivariate time series models. *The Annals of Statistics*, 37:4046, 2009.

[3] A. Aue and L. Horváth. Structural breaks in time series. *Journal of Time Series Analysis*, 34:1, 2013.

[4] J. E. Castrillón-Candás. High dimensional multilevel Kriging: A computational mathematics approach. *Arxiv*, 2020. `https://arxiv.org/abs/1701.00285`.

[5] J. E. Castrillón-Candás and K. Amaratunga. Fast estimation of continuous karhunen-loeve eigenfunctions using wavelets. *IEEE Transactions on Signal Processing*, 50(1):78–86, 2002.

[6] J. E. Castrillón-Candás, M. G. Genton, and R. Yokota. Multi-level restricted maximum likelihood covariance estimation and Kriging for large non-gridded spatial datasets. *Spatial Statistics*, 18, Part A:105 – 124, 2016. Spatial Statistics Avignon: Emerging Patterns.

[7] J. E. Castrillón-Candás, J. Li, and V. Eijkhout. A discrete adapted hierarchical basis solver for radial basis function interpolation. *BIT Numerical Mathematics*, 53(1):57–86, 2013.

[8] C.-S. J. Chu, M. Stinchcombe, and H. White. Monitoring structural change. *Econometrica*, 64:1045, 1996.

[9] S. Dasgupta and Y. Freund. Random projection trees and low dimensional manifolds. In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, STOC '08, pages 537–546, New York, NY, USA, 2008. ACM.

[10] R. A. Davis, D. Huang, and Y.-C. Yao. Testing for a change in the parameter values and order of an autoregressive model. *The Annals of Statistics*, 23:282, 1995.

[11] H. Dette and J Gösmann. A likelihood ratio approach to sequential change point detection for a general class of parameters. *Journal of the American Statistical Association*, 115(531):1361–1377, 2020.

[12] S. Fremdt. Page's sequential procedure for change-point detection in time series regression. *Statistics*, 48:1, 2014.

[13] H. Harbrecht, M. Peters, and M. Siebenmorgen. Analysis of the domain mapping method for elliptic diffusion problems on random domains. *Numerische Mathematik*, 134(4):823–856, 2016.

[14] D. V. Hinkley. Inference about the change-point from cumulative sum tests. *Biometrika*, 58:509, 1971.

[15] L. Horváth, M. Huŝková, P. Kokoszka, and J. Steinebach. Monitoring changes in linear models. *Journal of Statistical Planning and Inference*, 126:225, 2004.

[16] J. Istas. Karhunen–Loeève expansion of spherical fractional brownian motions. *Statistics and Probability Letters*, 76:1578–1583, 2006.

[17] V. Jandhyala, S. Fotopoulos, I. MacNeill, and P. Liu. Inference for single and multiple change-points in time series. *Journal of Time Series Analysis*, 34:423, 2013.

[18] M. Jirak. Uniform change point tests in high dimension. *The Annals of Statistics*, 43:2451, 2015.

[19] C. Kirch and S. Weber. Modified sequential change point procedures based on estimating functions. *Electronic Journal of Statistics*, 12:1579, 2018.

[20] T. L. Lai. Sequential changepoint detection in quality control and dynamical systems. *Journal of the Royal Statistical Society, Series B*, 57:613, 1995.

[21] K. V. Laven. *Grid Sphere.* 2010. `https://www.mathworks.com/matlabcentral/fileexchange/28842-grid-sphere`.

[22] W.A. Light and E.W. Cheney. *Approximation theory in tensor product spaces.*, volume 1169 of *Lecture notes in mathematics*. Springer, New York, 1985.

[23] M. Loève. *Probability theory*, volume II. Springer-Verlag, 4^{th} edition, 1978. Graduate Texts in Mathematics.

[24] MATLAB. *version 9.4 (R2018a)*. The MathWorks Inc., Natick, Massachusetts, 2018.

[25] G. V. Moustakides. Optimal stopping times for detecting changes in distributions. *The Annals of Statistics*, 14:1379, 1986.

[26] F. Nobile and R. Tempone. Analysis and implementation issues for the numerical approximation of parabolic equations with random coefficients. *International Journal for Numerical Methods in Engineering*, 80(6-7):979–1006, 2009.

[27] E. S. Page. Continuous inspection schemes. *Biometrika*, 41:100, 1954.

[28] K. Pape, D. Wied, and P. Galeano. Monitoring multivariate variance changes. *Journal of Empirical Finance*, 39:54, 2016.

[29] C. Schwab and R. A. Todor. Karhunen–Loève approximation of random fields by generalized fast multipole methods. *Journal of Computational Physics*, 217(1):100 – 122, 2006. Uncertainty Quantification in Simulation Science.
[30] X. Shao. Self-normalization for time series: A review of recent developments. *Journal of the American Statistical Association*, 110:1797, 2015.
[31] X. Shao and X. Zhang. Testing for change points in time series. *Journal of the American Statistical Association*, 105:1228, 2010.
[32] J. Tausch and J. White. Multiscale bases for the sparse representation of boundary integral operators on complex geometry. *SIAM Journal on Scientific Computing*, 24(5):1610–1629, 2003.
[33] L. Wang. *Karhunen-Loeve expansions and their applications*. PhD thesis, London school of economics and political science, March 2008.
[34] D. Wied and P. Galeano. Monitoring correlation change in a sequence of random variables. *Journal of Statistical Planning and Inference*, 143:186, 2013.
[35] T. Zhang and L. Lavitas. Unsupervised self-normalized change-point testing for time series. *Journal of the American Statistical Association*, 113:637, 2018.