Sequential block bootstrap in a Hilbert space with application to change point analysis

Olimjon Sharipov*, Johannes Tewes† and Martin Wendler‡

December 2, 2014

Abstract

A new test for structural changes in functional data is investigated. It is based on Hilbert space theory and critical values are deduced from bootstrap iterations. Thus a new functional central limit theorem for the block bootstrap in a Hilbert space is required. The test can also be used to detect changes in the marginal distribution of random vectors, which is supplemented by a simulation study. Our methods are applied to hydrological data from Germany.

Keywords: near epoch dependence, Hilbert space, block bootstrap, functional data, change-point test

1 Introduction

Statistical methods for functional data have received great attention during the last decade and environmental observations, see Hörmann and Kokoszka [16], are one of many areas where such data appear. Due to a strong seasonal effect, for example in temperature or hydrological data, such time series are non-stationary and thus change point analysis is a complex topic. A possible solution is looking at annual curves instead of the whole time series and therefore observations become functions. Functional principal components was used by Kokoszka et.al. [19] in testing for independence in the functional linear model and by Benko, Härdle and Kneip [2] in two sample tests for $L^2[0,1]$-valued random variables, a method that was extended to change point analysis by Berkes et.al. [3]. Another approach is due to Fraiman et.al. [14] who use record functions to detect trends in functional data. In contrast to all former approaches our method takes the fully functional observation into account. Whereas the statistic of Berkes et.al. [2] is $\mathbb{R}^d$-valued, our statistic depends directly

*Institute of Mathematics, National University of Uzbekistan, 29 Dormon Yoli Str., Tashkent, 100125,Uzbekistan

†Fakultät für Mathematik, Ruhr-Universität Bochum, 44780 Bochum, Germany, Email address: Johannes.Tewes@rub.de

‡Mathematisches Institut, Universität zu Köln, 50931 Köln, Germany

on the functional or more generally Hilbert space-valued random variables. This gets clear when considering the analogue of the Cusum statistic, which takes the maximum of the norm of

$$\sum_{i=1}^{k} X_i - \frac{k}{n} \sum_{i=1}^{k} X_i \quad \text{for } k = 1, \ldots, n-1, \tag{1}$$

where X_i takes values in a Hilbert space H.
Another change point problem are changes in the marginal distribution of random variables, now taking values in $\mathbb{R}^d$. The Kolmogorov Smirnov-type change point test was used for example by Inoue [17] and its statistic reads as follow

$$\max_{1 \leq m \leq n-1} \sup_{t \in [0,1]} |F_m(t) - F_{m+1;n}(t),| \tag{2}$$

where F_m and $F_{m+1;n}$ are empirical distribution functions, based on $X_1, \ldots, X_m$ and $X_{m+1}, \ldots X_n$, respectively. Define Y_i by $Y_i(t) := 1_{\{X_i \leq t\}}$ then (2) equals

$$\max_{1 \leq m \leq n-1} \|\bar{Y}_m - \bar{Y}_{m+1;n}\|_\infty.$$

The Y_i are no longer real valued random variables, they take values in a function space. Often one uses the space $D[0,1]$ of cadlag functions but functional central limit theorems in $D[0,1]$ are difficult to obtain. Therefore in this paper we want to consider the Hilbert space L^2. Using the norm $\|\cdot\|$, induced by the inner product of the Hilbert space instead of the supremums norm, we get the statistic

$$\max_{1 \leq m \leq n-1} \|\bar{Y}_m - \bar{Y}_{m+1;n}\|_{L_2}.$$

Critical values for change point test are often deduced from asymptotics. (1) can be expressed as a functional of the partial sum process

$$\sum_{i=1}^{\lfloor nt \rfloor} X_i \quad \text{for } t \in [0,1], \tag{3}$$

whose asymptotic behavior for H-valued data was investigated by Chen and White [8] for mixingales and near epoch dependent processes. For statistical inference one needs control over the asymptotic distribution but due to dependence and the infinite dimension of the $\{X_i\}_{i \geq 1}$ this depends on an unknown infinite dimensional parameter - the covariance operator. Our solution is the bootstrap which has been successfully applied to many statistics in the case of real or $\mathbb{R}^d$-valued data. For Hilbert spaces only Politis and Romano [23] and recently Dehling, Sharipov and Wendler [12] showed the asymptotic validity of the bootstrap. The results of Politis and Romano [23] can only handle bounded random variables. Thus indicator functions and statistics of type (2) can be bootstrapped by their method, but general functional data can not.
We extend the non overlapping block bootstrap of Dehling, Sharipov and Wendler

[12] by a sequential component which is inevitable for change point statistics.
The paper is organized as follows: Sections 2.1 and 2.2 contain the main results, an invariance principle for H-valued processes and the functional central limit theorem for bootstrapped data. Section 3 describes the statistics and the bootstrap methodology for different change point tests including converging alternatives. In a small simulation study the finite sample behavior of our test both under stationarity and under structural changes in mean or skewness is investigated and compared to the performance of the classical cusum test. Moreover real life data are analyzed and finally proofs are provided in the appendix.

2 Main results

2.1 Functional Central Limit Theorem for Hilbert space-valued functionals of mixing processes

Let H be a separable Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$. Further let $(\xi_i)_{i\in\mathbb{Z}}$ be a stationary sequence of random variables, taking values in an arbitrary separable measurable space. A sequence $(X_n)_{n\in\mathbb{Z}}$ of H-valued random variables is called L_p-near epoch dependent (NED(p)) on $(\xi)_{i\in\mathbb{Z}}$, if there is a sequence $(a_k)_{k\in\mathbb{N}}$ with $a_k \to 0$ as $k \to \infty$ and

$$E\|X_0 - E[X_0|\mathcal{F}_{-k}^k]\|^p \le a_k.$$

Here $\mathcal{F}_{-k}^k$ denotes the σ-field generated by $\xi_{-k},\dots,\xi_k$. For the definition of conditional expectation in Hilbert spaces see Ledoux and Talagrand [20].
Concerning the sequence $(\xi_i)_{i\in\mathbb{Z}}$, we will assume the following notion of mixing. Define the coefficients

$$\beta(k) = \left| E \sup_{A\in\mathcal{F}_k^\infty} [P(A|\mathcal{F}_{-\infty}^0) - P(A)] \right|.$$

$(\xi)_{i\in\mathbb{Z}}$ is called absolutely regular if $\beta(k) \to 0$ as $k \to \infty$.

It is our aim to prove functional central limit theorems for H valued variables. Therefore we will use the space $D_H[0,1]$, the set of all cadlag functions mapping from $[0,1]$ to H. An H-valued function on $[0,1]$ is said to be cadlag, if it is right-continuous and the left limit exists for all $x \in [0,1]$. Analogously to the real valued case we define the Skorohod metric

$$d(f,g) = \inf_{\lambda\in\Lambda} \left\{ \sup_{t\in[0,1]} \|f(t) - g\circ\lambda(t)\| + \|id - \lambda\|_\infty \right\} \quad f,g \in D_H[0,1],$$

where Λ is defined as usual, $\|\cdot\|$ is the Hilbert space norm and $\|\cdot\|_\infty$ the uniform norm.
Most properties from the well known space $D[0,1]$ carry over to this space, so

equipped with the Skorohod metric $D_H[0,1]$ becomes a separable Banach space.

The first result states convergence of the partial sum process. Such a result was given by Walk [26] for martingale difference sequences and by Chen and White [8] in the near epoch dependent case. They assume strong mixing, which is more general than absolute regularity. Then again we require L_1-near epoch dependence, while they use L_2-near epoch dependence, which implies our conditions.

Theorem 2.1. *Let $(X_n)_{n\in\mathbb{Z}}$ be L_1-near epoch dependent on a stationary, absolutely regular sequence $(\xi_n)_{n\in\mathbb{Z}}$ with $EX_1 = \mu$ and assume that the following conditions hold for some $\delta > 0$*

1. $E\|X_1\|^{4+\delta} < \infty$,
2. $\sum_{m=1}^{\infty} m^2 (a_m)^{\delta/(\delta+3)} < \infty$,
3. $\sum_{m=1}^{\infty} m^2 (\beta(m))^{\delta/(\delta+4)} < \infty$.

Then

$$\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor nt\rfloor}(X_i - \mu)\right)_{t\in[0,1]} \Rightarrow (W_t)_{t\in[0,1]} \tag{4}$$

where $(W_t)_{t\in[0,1]}$ is a Brownian motion in H and W_1 has the covariance operator $S \in \mathcal{S}(H)$, defined by

$$\langle Sx, y\rangle = \sum_{i=-\infty}^{\infty} E[\langle X_0 - \mu, x\rangle\langle X_i - \mu, y\rangle], \quad \text{for } x, y \in H. \tag{5}$$

Furthermore the series in (5) converges absolutely.

2.2 Sequential Bootstrap for H-valued random variables.

When using the result of the previous section in applications, for example change point test (see section 3), one is confronted with the problem, that the limiting distribution may be unknown, or even if it is known it depends on an infinite dimensional parameter, in our case the covariance operator S.
To circumvent this problem, we will use the non overlapping block bootstrap of Carlstein [7] to establish a process with the same limiting distribution as $\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor nt\rfloor}(X_i - \mu)$.
For a block length $p(n)$ consider the $k = \lfloor n/p\rfloor$ blocks $I_1, \ldots, I_k$, defined by

$$I_j = (X_{(j-1)p+1}, \ldots, X_{jp}).$$

Then we choose randomly and independently k blocks, so that the bootstrap sample $X_1^*, \ldots, X_{kp}^*$ satisfies

$$P\left((X^*_{(j-1)p+1}, \ldots, X^*_{jp}) = I_i\right) = \frac{1}{k} \quad \text{for} \quad i, j = 1, \ldots, k,$$

Now we can define a bootstrapped version of the partial sum process, which is given by

$$W_{n,p}^*(t) = \frac{1}{\sqrt{kp}} \sum_{i=1}^{\lfloor kpt \rfloor} (X_i^* - E^* X_i^*). \tag{6}$$

As usual E^* and P^* denote conditional expectation and probability, respectively, given $\sigma(X_1, \ldots, X_n)$. Further $\Rightarrow_*$ denotes weak convergence with respect to P^*. The next Theorem establishes the asymptotic distribution of (6).

Theorem 2.2. *Let $(X_n)_{n\in\mathbb{Z}}$ be L_1-near epoch dependent on a stationary, absolutely regular sequence $(\xi_n)_{n\in\mathbb{Z}}$ with $EX_1 = \mu$ and assume that the following conditions hold for some $\delta > 0$*

1. $E\|X_1\|^{4+\delta} < \infty$,
2. $\sum_{m=1}^{\infty} m^2 (a_m)^{\delta/(\delta+3)} < \infty$,
3. $\sum_{m=1}^{\infty} m^2 (\beta(m))^{\delta/(\delta+4)} < \infty$.

Further let the block length be nondecreasing, $p = O(n^{1-\epsilon})$ for some ϵ and $p_n = p_{2^l}$ for $n = 2^{l-1} + 1, \cdots, 2^l$. Then

$$\left(W_{n,p}^*(t)\right)_{t\in[0,1]} \Rightarrow_* (W_t)_{t\in[0,1]} \quad a.s. \tag{7}$$

where $(W_t)_{t\in[0,1]}$ is a Brownian motion in H and W_1 has the covariance operator $S \in \mathcal{S}(H)$, defined in Theorem 2.1

3 Application to change point tests

3.1 Change in the mean of H-valued data

Let us consider the following change point problem. Given $X_1, \ldots, X_n$, we want to test the Hypothesis

$$H\colon \qquad EX_1 = \cdots = EX_n$$

against the Alternative

$$A\colon \qquad EX_1 = \cdots = EX_k \neq EX_{k+1} = \cdots = EX_n,$$

for some $k \in \{1, \ldots, n\}$.

For real-valued variables, asymptotics of cusum-type tests have been intensively studied by Csörgő and Horvath [9]. They investigated test for i.i.d data, weakly dependent data and for long range dependent processes. The third case was extended by Dehling, Rooch and Taqqu [10].

For functional data, Berkes et.al. [3] have developed estimators and tests for a change point in the mean, which is extended by Hörmann and Kokoszka [16] and Aston and

Kirch [1] to weakly dependent data. They use functional principal components, while - motivated by Theorems 2.1 and 2.2 - we bootstrap the complete functional data. Consider the test statistic

$$T_n = \max_{1\le m\le n-1} \frac{1}{\sqrt{n}} \left\| \sum_{i=1}^{m} X_i - \frac{m}{n} \sum_{i=1}^{n} X_i \right\|$$

and its bootstrap analogue

$$T_n^* = \max_{1\le m\le kp-1} \frac{1}{\sqrt{kp}} \left\| \sum_{i=1}^{m} X_i^* - \frac{m}{kp} \sum_{i=1}^{kp} X_i^* \right\|.$$

The next result states that T_n and T_n^* have the same limiting distribution, which is a direct consequence of Theorem 2.1 and Theorem 2.2 and the continuity of both the maximum function and the Hilbert space norm.

Corollary 3.1. *(i) Under the conditions of Theorem 2.1*

$$T_n \Rightarrow \max_{t\in[0,1]} \|W(t) - tW(1)\|, \tag{8}$$

where $(W(t))_{t\in[0,1]}$ *is the Brownian motion defined in Theorem 2.1.*
(ii) Under the conditions of Theorem 2.2

$$T_n^* \Rightarrow_* \max_{t\in[0,1]} \|W(t) - tW(1)\| \quad a.s. \tag{9}$$

Next we want to work out the asymptotic distribution of the (bootstrapped) change-point statistic under a sequence of converging alternatives. Therefore define the triangular array of H-valued random variables

$$Y_{n,i} = \begin{cases} X_i & \text{if } i \le \lfloor n\tau \rfloor \\ X_i + \Delta_n & \text{if } i > \lfloor n\tau \rfloor \end{cases}$$

for $n \in \mathbb{N}$ and $i \le n$. Here $\lfloor n\tau \rfloor$ is the unknown change-point for some $\tau \in (0,1)$ and $(\Delta_n)_n$ is an H-valued deterministic sequence with

$$\|\sqrt{n}\Delta_n - \Delta\| \to 0,$$

for $n \to \infty$ and some $\Delta \in H$.
Now we want to test the Hypothesis $\Delta_n = 0$ against the sequence of Alternatives where $\Delta, \Delta_{n\in\mathbb{N}} \in H \setminus \{0\}$.
Note that a bootstrap sample $(Y_{n,i}^*)_{i\le kp, n\ge 1}$ can be created analogously to $(X_i^*)_{i\le kp}$. Then we can define the statistics T_n and T_n^*, now based on $Y_{n,i}$ and $Y_{n,i}^*$, respectively.

Corollary 3.2. *(i) Consider an array* $(Y_{n,i})_{n\in\mathbb{N}, i\le n}$*. If the conditions of Theorem 2.1 hold for* $(X_i)_{i\ge1}$*, then under the sequence of local alternatives*

$$T_n \Rightarrow \max_{t\in[0,1]} \|W(t) - tW(1) + \phi_\tau(t)\Delta\|, \tag{10}$$

where $(W(t))_{t\in[0,1]}$ is the Brownian Motion defined in Theorem 2.1 and the function $\phi_\tau\colon [0,1] \to \mathbb{R}$ is defined by

$$\phi_\tau(t) = \begin{cases} t(1-\tau) & \textit{if } t \le \tau \\ (1-t)\tau & \textit{if } t > \tau. \end{cases}$$

(ii) If the conditions of Theorem 2.2 are satisfied, then under the sequence of local alternatives

$$T_n^* \Rightarrow_* \max_{t\in[0,1]} \|W(t) - tW(1)\| \quad a.s. \tag{11}$$

The Corollaries motivate the following test procedure, which is typically for bootstrap tests:

(i) Compute T_n.

(ii) Simulate $T_{j,n}^*$ for $j = 1, \dots, J$.

(iii) Based on the independent (conditional on $X_1, \dots, X_n$) random variables $T_{n,1}^*, \dots, T_{n,J}^*$, compute the empirical $(1-\alpha)$- quantile $q_{n,J}(\alpha)$.

(iv) If $T_n > q_{n,J}(\alpha)$ reject the Hypothesis.

By Corollary 3.1 and the Glivenko-Cantelli Theorem the proposed test has asymptotically significance level α, whereas by Corollary 3.2 it has asymptotically nontrivial power.

3.2 Change in the marginal distribution

We will now apply the results to random variables, whose realizations are not truly functional. Consider for example real valued random variables $X_1, \dots, X_n$ and the test problem of no change in the underlying distribution, in detail

$$H\colon \qquad P(X_1 \le t) = \dots = P(X_n \le t) \ \ \forall t \in \mathbb{R}$$

against

$$A\colon \qquad P(X_1 \le t) = \dots = P(X_k \le t) \ne P(X_{k+1} \le t) = \dots = P(X_n \le t),$$

for some $k \in \{1, \dots, n\}$ and $t \in \mathbb{R}$.
Asymptotic tests have been investigated by Csörgő and Horvath [9] and Szyszkowicz [25] in the independent case, by Inoue [17] for strong mixing data and by Giraitis, Leipus and Surgailis [15] for long-memory linear processes. Common test statistics depend on the empirical distribution function and therefore the indicators

$$1_{\{X_i \le t\}}, \quad t \in \mathbb{R}. \tag{12}$$

Those can be interpreted as random functions and therefore random elements of the Hilbert space of functions $f\colon \mathbb{R} \to \mathbb{R}$, equipped with the inner product

$$\langle f, g \rangle = \int_{\mathbb{R}} f(t) g(t) w(t) \; dt$$

for some positive, bounded weight function with $\int w(t) \; dt < \infty$.
By Fubini's Theorem we have

$$E\left[\int_{\mathbb{R}} 1_{\{X_i \le t\}} h(t) w(t) \; dt\right] = \int_{\mathbb{R}} F(t) h(t) w(t) \; dt \quad \forall h \in H,$$

so by the definition it follows that the mean of (12) is just the distribution function of X. So the change in the mean-problem (in H) becomes a change in distribution-problem (in $\mathbb{R}$).
Furthermore the arithmetic mean becomes the empirical distribution function. Note, that this still holds, when we consider $\mathbb{R}^d$-valued data which leads to the following test statistic

$$T_{n,w} = \max_{1 \le m \le n-1} \frac{1}{n} \int_{\mathbb{R}^d} \left(\sum_{i=1}^{m} 1_{\{X_i \le t\}} - \frac{m}{n} \sum_{i=1}^{n} 1_{\{X_i \le t\}} \right)^2 w(t) \; dt, \tag{13}$$

which can be described as Cramér-von Mises change-point statistic. In the $\mathbb{R}^d$-valued case the weight function becomes a positive function $w\colon \mathbb{R}^d \to \mathbb{R}$ with

$$\int_{\mathbb{R}^d} w(t) \; dt < \infty.$$

The empirical process has been bootstrapped by Bülmann [6] and Naik-Nimabalkar and Rajarshi [22] and recently by Doukhan et.al. [13] using the wild bootstrap and by Kojadinovic and Yan [18] using weighted bootstrap.
Our bootstrapped version of (13) reads as follows

$$T^*_{n,w} = \max_{1 \le m \le kp-1} \frac{1}{kp} \int_{\mathbb{R}^d} \left(\sum_{i=1}^{m} 1_{\{X^*_i \le t\}} - \frac{m}{kp} \sum_{i=1}^{kp} 1_{\{X^*_i \le t\}} \right)^2 w(t) \; dt, \tag{14}$$

where the sample $X^*_1, \dots, X^*_{kp}$ is produced by the non overlapping block bootstrap.
We will now state conditions, for which the bootstrap holds in this scenario.

Corollary 3.3. *Let $(X_n)_{n \in \mathbb{N}}$ be $\mathbb{R}^d$ valued random variables, L_1- near epoch dependent on a stationary, absolutely regular sequence $(\xi_n)_{n \in \mathbb{Z}}$ such that for some $\delta > 0$*

1. $E\|X_1\| < \infty$
2. $\sum_{m=1}^{\infty} m^2 (a_m)^{\delta/(1+4\delta)} < \infty$,
3. $\sum_{m=1}^{\infty} m^2 (\beta_m)^{\delta/(\delta+3)} < \infty$.

Let the block lenght p be nondecreasing with $p(n) = O(n^{1-\epsilon})$ and $p(n) = p(2^l)$ for $n = 2^{l-1}+1, \dots, 2^l$.
Then T_n^* has almost surely the same limiting distribution as T_n .

Note, that producing a bootstrap sample $X_1^*, \dots, X_{kp}^*$ first, and then treating the indicators

$$1_{\{X_1^* \le \cdot\}}, \dots, 1_{\{X_{kp}^* \le \cdot\}},$$

is the same as if we first look upon the indicators as H-valued random variables $Y_1, \dots, Y_n$ and then generate $Y_1^*, \dots, Y_{kp}^*$.
Now we can apply Corollary 3.1 and therefore we have to verify the conditions of Theorems 2.1 and 2.2, respectively.
The moment condition is automatically satisfied, due to the definition of $w(t)$ and the dependence conditions are satisfied because of Lemma 2.2 in Dehling, Sharipov and Wendler [12] and the Lipschitz-continuity of the mapping $x \mapsto 1_{\{x \le \cdot\}}$.

Remark 3.4. Instead of the inner product we have defined one can use

$$\langle f, g\rangle_{id} = \int f(t)g(t)\ dt \quad \text{ or } \quad \langle f, g\rangle_F = \int f(t)g(t)\ dF(t),$$

which lead to well known change point statistics. Note, that in the first case the norm of the indicator $1_{\{X_1 \le \cdot\}}$ is infinite, but we can consider $1_{\{X_1 \le \cdot\}} - F(\cdot)$. Further we have to make additional moment assumptions on the X_i to make Corollary 3.3 still hold.

3.3 Data Examples

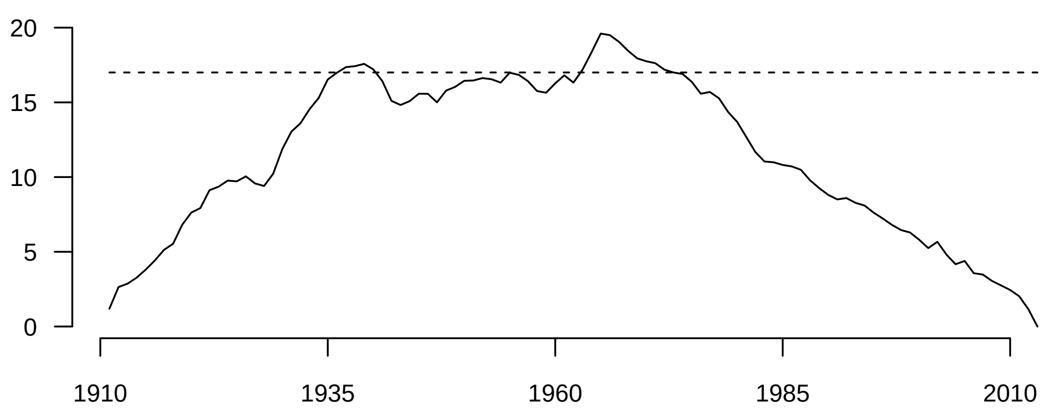

Figure 1: Process $\frac{1}{\sqrt{n}}\|\bar{X}_k - k/n\bar{X}_n\|$ (black line) computed from 103 annual flow curves of the river Chemnitz and 0.95 level of significance (dashed line) computed from 999 Bootstrap iterations.

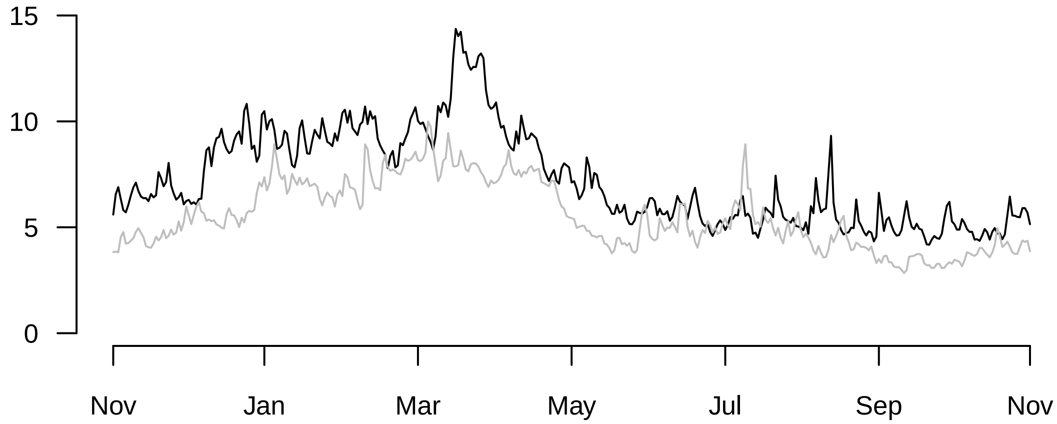

Figure 2: Average annual flow curves of the time period 1910 - 1964 (grey line) and the time period 1965 - 2012 (black line).

To illustrate our methods we apply the tests, described in the previous subsections to hydrological observations.
The first data set contains average daily flows of the river Chemnitz for the time period 1910 - 2012. Thus one gets 103 annual flow curves which can be interpreted as realizations of $\mathbb{R}^{365}$-valued random variables. Alternatively one could smoothen the curves and hence get functional data. Let X_i be the ith annual curve, than Figure 1 shows the process

$$\frac{1}{\sqrt{n}}\left\|\sum_{i=1}^{k} X_i - \frac{k}{n}\sum_{i=1}^{n} X_i\right\| \qquad k = 1, \dots n-1.$$

The valued of the test statistic is the maximum of this process, which is attained 1964. Because it is larger than the bootstrapped 5% level of significance, the test indicates that there has been a change in structure of the annual flow curves.
Figure 2 illustrates the character of this change by comparing the average flow curves based on the data before and after 1964.

As a second example, we look at annual maximum flows of the river Elbe for the time period 1850 - 2012. We apply the test for distributional change to this $\mathbb{R}$-valued observations and therefore compute (13) and 999 iterations of (14). Figure 3 shows the process

$$\frac{1}{n}\int\left(\sum_{i=1}^{k} 1_{\{X_i \le x\}} - \frac{k}{n}\sum_{i=1}^{n} 1_{\{X_i \le x\}}\right)^2 \phi(x)dx \qquad k = 1, \dots, n-1$$

where we have used the probability density of the $N(2000, 2000^2)$ distribution as weight function $\phi(x)$. The value of the test statistic equals the maximum of this process, which is larger than the bootstrapped level of significance and therefore a change is detected.
Finally Figure 4 compares the empirical distribution functions based on the data before and after 1900, which is where the maximum is attained.

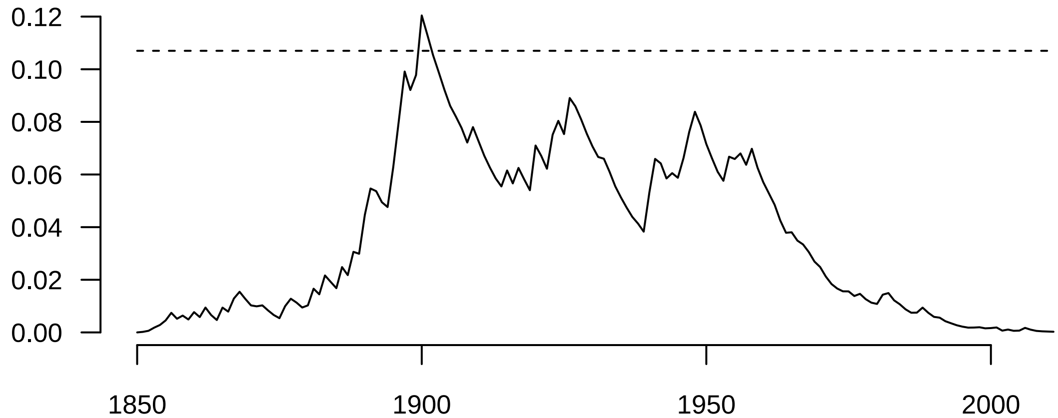

Figure 3: Process $\frac{1}{n}\int(F_k(x)-k/nF_n(x)^2\phi(x)dx$ (black line) computed from 163 annual maximum flows of the river Elbe and 0.95 level of significance (dashed line) computed from 999 Bootstrap iterations.

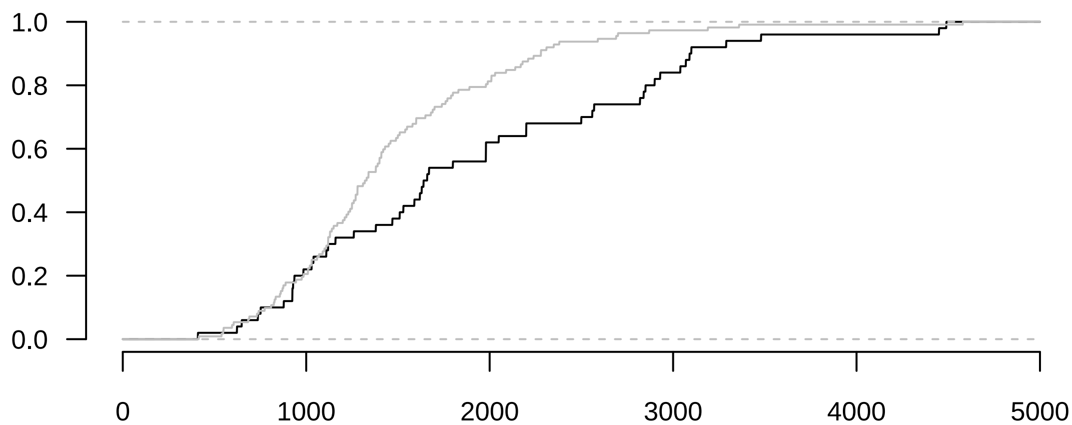

Figure 4: Empirical distribution functions of the first 50 observations (black line) and the last 113 observations (grey line).

3.4 Simulation Study

Corollary 3.3 give a wide range of block lengths and dependency conditions, that enables us to derive asymptotic properties of the change-in-distribution-test.
In a small simulation study we investigate the finite sample performance considering different block lengths and three kinds of dependencies. The data generating process is an AR(1)-process, in detail

$$X_t = a_1 X_{t-1} + \epsilon_t,$$

with $a_1 \in \{0.2, 0.5, 0.8\}$ and $\epsilon_t \sim N(0, 1-a_1^2)$, independent of each other. In all situations we have calculated critical values by $J = 999$ bootstrap-iterations and

Table 1: Empirical size (No change)

		$a_1 = 0.2$	$a_1 = 0.5$	$a_1 = 0.8$
$n = 50$	$l = 4$	0.05	0.127	0.23
	$l = 5$	0.044	0.085	0.212
	$l = 7$	0.046	0.076	0.155
$n = 100$	$l = 6$	0.035	0.082	0.254
	$l = 8$	0.056	0.059	0.171
	$l = 10$	0.047	0.072	0.131
	$l = 12$			0.122
$n = 200$	$l = 8$	0.061	0.091	0.201
	$l = 10$	0.04	0.064	0.149
	$l = 12$	0.055	0.067	0.137
	$l = 15$	0.042	0.066	0.1

Table 2: Empirical size ($\mu = 0.5$, change at 1/2 of length)

	$a_1 = 0.2$	$a_1 = 0.5$	$a_1 = 0.8$
	$l = 4$	$l = 7$	$l = 7$
$n = 50$	0.23	0.161	0.207
	$l = 10$	$l = 8$	$l = 12$
$n = 100$	0.302	0.28	0.206
	$l = 12$	$l = 12$	$l = 15$
$n = 200$	0.669	0.456	0.258

Table 3: Empirical size of Cusum-test($\mu = 1$, change at 1/2 of length)

	$a_1 = 0.2$	$a_1 = 0.5$	$a_1 = 0.8$
	$l = 4$	$l = 7$	$l = 7$
$n = 50$	0.678	0.431	0.335
	$l = 10$	$l = 8$	$l = 12$
$n = 100$	0.851	0.708	0.373
	$l = 12$	$l = 12$	$l = 15$
$n = 200$	0.998	0.945	0.630

Table 4: Empirical size ($\mu = 1$, change at 1/2 of length)

	$a_1 = 0.2$	$a_1 = 0.5$	$a_1 = 0.8$
	$l = 4$	$l = 7$	$l = 7$
$n = 50$	0.686	0.313	0.351
	$l = 10$	$l = 8$	$l = 12$
$n = 100$	0.847	0.695	0.419
	$l = 12$	$l = 12$	$l = 15$
$n = 200$	0.995	0.937	0.64

Table 5: Empirical size (change in skewness at 1/2 of length)

	$a_1 = 0.2$	$a_1 = 0.5$	$a_1 = 0.8$
	$l = 4$	$l = 7$	$l = 7$
$n = 50$	0.323	0.244	0.196
	$l = 10$	$l = 8$	$l = 12$
$n = 100$	0.546	0.461	0.223
	$l = 12$	$l = 12$	$l = 15$
$n = 200$	0.945	0.846	0.375

Table 6: Empirical size of Cusum-test (change in skewness at 1/2 of length)

	$a_1 = 0.2$	$a_1 = 0.5$	$a_1 = 0.8$
	$l = 4$	$l = 7$	$l = 7$
$n = 50$	0.047	0.062	0.08
	$l = 10$	$l = 8$	$l = 12$
$n = 100$	0.033	0.045	0.076
	$l = 12$	$l = 12$	$l = 15$
$n = 200$	0.04	0.045	0.065

empirical size and power by $m = 1000$ iterations of the test.
For the empirical size under the hypothesis of no change see table 1. For rather weakly dependent variables ($a_1 = 0.2$) the performance is quite good, even for small sample sizes like $n = 50$. Whereas for an AR-coefficient $a_1 = 0.8$ the empirical size is drastically larger than the nominal one. This is typically for bootstrap tests due to an underestimation of covariances, see for example Doukhan et. al. [13].
Regarding the power of our test we choose for given sample size and AR-coefficient the block length, that provides the best empirical power under this circumstances. We start with the following change-in-mean model.

$$Y_t = \begin{cases} X_t & \text{for } t \leq t^* \\ X_t + \mu & \text{for } t > t^*. \end{cases}$$

Table 2 and table 4 give an overview of the empirical size under this alternative for $\mu = 0.5$ and $\mu = 1$, respectively. We see that a level shift of height $\mu = 0.5$ in an AR-process with $a_1 = 0.8$ is to small to be detected. However for larger shifts ($\mu = 1$) the power of our test is notable good.
There are several tests to detect a mean-shift, such as the Cusum-test or the Wilcoxon-change-point-test. If critical values can deduce from a known asymptotic distribution, the Cusum test is supposed to have greater power then our test. However if critical values are investigated by the bootstrap, tables 3 and 4 indicate that both tests have similar power properties.
To illustrate the power of our test against several alternatives, consider a change in skewness of a process. Therefore we need a second DGP $X'_t = a_1 X'_{t-1} + \epsilon'_t$, indepen-

dent of the first one, and define

$$Y_t = \begin{cases} X_t^2 + X_t'^2 & \text{for } t \leq t^* \\ 4 - (X_t^2 + X_t'^2) & \text{for } t > t^*. \end{cases}$$

Table 5 shows that against this alternative the power is excellent for $n = 200$ and coefficients $a_1 \leq 0.5$. Table 6 illustrates the power of the Cusum-test, where critical values are also computed by 999 bootstrap iterations. The study shows that this test does not see changes in the skewness when the mean is unmodified.

Acknowledgements

The research was supported by the DFG Sonderforschungsbereich 823 (Collaborative Research Center) Statistik Nichtlinearer Dynamischer Prozesse.
The authors are grateful to Svenja Fischer and Andreas Schumann from the faculty of civil engineering, Ruhr-Universität Bochum, for providing hydrological data.

References

[1] J. Aston and C. Kirch. Detecting and estimating changes in dependent functional data. *Journal of Multivariate Analysis*, 109:204–220, August 2012.

[2] M. Benko, W. Härdle, and A. Kneip. Common functional principal components. *The Annals of Statistics*, 37(1):1–34, 2009.

[3] I. Berkes, R. Gabrys, L. Horváth, and P. Kokoszka. Detecting changes in the mean of functional observations. *Journal of the Royal Statistical Society: Series B*, 71(5):927–946, 2009.

[4] P. Billingsley. *Convergence of Probability Measures*. New York:Wiley, 1968.

[5] S. Borovkova, R. Burton, and H. Dehling. Limit theorems for functionals of mixing processes with applications to U-statistics and dimension estimation. *Transactions of the American Mathematical Society*, 353:4261–4318, 2001.

[6] P. Bühlmann. Blockwise bootstrapped empirical process for stationary sequences. *The Annals of Statistics*, 22(2):995–1012, 1994.

[7] E. Carlstein. The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *The Annals of Statistics*, 14(3):1171–1179, 1986.

[8] X. Chen and H. White. Central limit and functional central limit theorems for hilbert-valued dependent heterogeneous arrays with applications. *Econometric theory*, 1998.

[9] M. Csörgő and L. Horvath. *Limit Theorems in in Change-Point Analysis*. John Wiley and Sons, 1997.

[10] H. Dehling, A. Rooch, and M. Taqqu. Nonparametric change-point test for long range dependent data. *Scandinavian Journal of Statistics*, 40:153–173, 2012.

[11] H. Dehling, A. Rooch, and M. Taqqu. Power of change-point tests for long-range dependent data. Discussion Paper, June 2013.

[12] H. Dehling, O. Sharipov, and M. Wendler. Bootstrap for dependent Hilbert space-valued random variables with application to von Mises statistics. *Journal of Multivariate Analysis*, 133:200–215, 2015.

[13] P. Doukhan, G. Lang, A. Leucht, and M. Neumann. Dependent wild bootstrap for the empirical process. *Accepted for Journal of Time Series Analysis.*, 2014+.

[14] R. Fraiman, A. Justel, R. Liu, and P. Llop. Detecting trends in time series of functional data: A study of antarctic climate change. *The Canadian Journal of Statistics*, pages 1–13, 2014.

[15] L. Girairis, R. Leipus, and D. Surgailis. The change-point problem for dependent observations. *Journal of Statistical Planing and Inference*, 53(297-310), 1996.

[16] S. Hörmann and P. Kokoszka. Weakly dependent functional data. *The Annals of Statistics*, 38:1848–1884, 2010.

[17] A. Inoue. Testing for distributional change in time series. *Econometric Theory*, 17:156–187, 2001.

[18] I. Kojadinovic and J.Yan. Goodness-of-fit testing based on a weighted bootstrap: A fast large-sample alternative to the parametric bootstrap. *The Canadian Journal of Statistics*, 40(3):480–500, 2012.

[19] P. Kokoszka, I. Maslova, J. Sojka, and L. Zhu. Testing for lack of dependence in the functional linear model. *The Canadian Journal of Statistics*, 36(2):207–222, 2008.

[20] M. Ledoux and M. Talagrand. *Probability in Banach Spaces*. Springer, 1991.

[21] F. Móricz. Moment inequalities and the strong laws of large numbers. *Probability Theory and related fields*, 35(299-314), 1976.

[22] U. V. Naik-Nimbalkar and M. B. Rajarshi. Validity of blockwise bootstrap for empirical processes with stationary observations. *The Annals of Statistics*, 22(2):980–994, 1994.

[23] D. Politis and J. Romano. Limit theorems for weakly dependent hilbert space valued random variables with applications to the stationary bootstrap. *Statistica Sinca*, 4:461–476, 1994.

[24] H. Rosenthal. On the subspaces of L^p ($p > 2$) spanned by independent random variables. *Israel Journal of Mathematics*, 8:273–303, 1970.

[25] B. Szyszkowicz. Weak convergence of weighted empirical type processes under contiguous and changepoint alternatives. *Stochastic Processes and their Applications*, 50(2):281–313, April 1994.

[26] H. Walk. An invariance principle for the Robbins-Monro process in a Hilbert space. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 39:135–150, 1977.

[27] J. Wooldrige and H. White. Some invariance principles and central limit theorems for dependent heterogeneous processes. *Econometric Theory*, 4(210-230), 1988.

A Prelimanary Results

Theorem A.1. *Let $\{W_n\}_{n\geq 1}$ be a sequence of $D_H[0,1]$- valued random functions with $W_n(0) = 0$. Then $\{W_n\}_{n\geq 1}$ is tight in $D_H[0,1]$ if the following condition is satisfied:*

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\frac{1}{\delta}P\left(\sup_{s\leq t\leq s+\delta}\|W_n(t)\| > \epsilon\right) = 0, \tag{15}$$

for each positive ϵ and each $s \in [0,1]$.
Furthermore the weak limit of any convergent subsequence of $\{W_n\}$ is in $C_H[0,1]$, almost surely.

For real valued random variables this is Theorem 8.3 of Billingsley [4], which carries over to $D[0,1]$. The proof still holds for Hilbert space valued functions.

Lemma A.2. *Let $(X_n)_{n\geq 1}$ be H-valued, stationary and L_1-near epoch dependent on an absolutely regular process with mixing coefficients $(\beta(m))_{m\geq 1}$ and approximation constants $(a_m)_{m\geq 1}$. If $EX_1 = 0$ and*

1. $E\|X_1\|^{4+\delta} < \infty$,
2. $\sum_{m=1}^{\infty} m^2 (a_m)^{\delta/(\delta+3)} < \infty$,
3. $\sum_{m=1}^{\infty} m^2 (\beta(m))^{\delta/(\delta+4)} < \infty$,

holds for some $\delta > 0$, then

$$E\|X_1 + X_2 + \cdots + X_n\|^4 \leq Cn^2 \left(E\|X_1\|^{4+\delta}\right)^{\frac{1}{1+\delta}}. \tag{16}$$

The result follows from the proof of Lemma 2.24 of Borovkova, Burton and Dehling [5], which is also valid for Hilbert spaces.

Lemma A.3. *Let $(X_n)_{n\geq 1}$ be a stationary sequence of H-valued random variables such that $EX_1 = 0$, $E\|X_1\|^4 < \infty$ and for some $C > 0$*

$$E\|X_1 + X_2 + \cdots + X_n\|^4 \leq Cn^2.$$

Then

$$E\max_{k\leq n}\|X_1 + X_2 + \cdots + X_k\|^4 \leq Cn^2.$$

This Lemma is a special case of Theorem 1 of Mórizc [21]. The proof carries over directly to Hilbert spaces.

B Proofs of the main results

Proof of Theorem 2.1. We start with the special case $H = \mathbb{R}$. Let $EX_1 = 0$, then by Lemma 2.23 of Borovkova , Burton and Dehling [5] we have

$$\frac{1}{n}E\left(\sum_{i=1}^{n} X_i\right)^2 \to \sigma^2, \tag{17}$$

where $\sigma^2 = \sum_{i=-\infty}^{\infty} E(X_0 X_i)$ and this series converges absolutely.
Furthermore by Lemma 2.4. of Dehling et. al. [12]we have

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(X_i - \mu) \Rightarrow N(0, \sigma^2). \tag{18}$$

Now convergence of finite dimensional distributions follows in the same way as in the proof of Theorem 21.1 in Billingsley [4], where functionals of ϕ-mixing sequences are considered. However concerning the mixing property the crucial line in obtaining convergence of more than one dimension is (21.29) (in [4] on page 187). But this converges to 0 even if the sequence is strong mixing and strong mixing is implied by our condition - absolute regularity.
If we can show, that the set

$$\left\{\max_{s\leq t\leq s+\delta}\frac{1}{\delta}(W_n(t) - W_n(s))^2 \mid 0 \leq s \leq 1, 0 \leq \delta \leq 1, n \leq N(s,\delta)\right\} \tag{19}$$

is uniformly integrable, then according to Lemma 2.2 in Wooldrige and White [27] W_n is tight in $D[0,1]$ equipped with the Skorohod topology and further the weak limit is almost surely in $C[0,1]$.
So fix $s \in [0,1]$ and $\delta \in [0,1]$. By the proof of Lemma 2.24 in Borovkova, Burton and Dehling [5] we obtain

$$E\left(\sum_{i=\lfloor ns\rfloor+1}^{\lfloor n(\delta+s)\rfloor} X_i\right)^4 \leq C(\lfloor n(\delta+s)\rfloor - \lfloor ns\rfloor)^2$$

Next Theorem 1 of Móricz [21] together with the moment inequality stated above implies

$$E\left(\max_{s\le t\le s+\delta}\left|\sum_{i=\lfloor ns\rfloor+1}^{\lfloor nt\rfloor}X_i\right|\right)^4 \le C(\lfloor n(\delta+s)\rfloor-\lfloor ns\rfloor)^2. \tag{20}$$

Now we will show uniform integrability of 19. Using first Hölder- and Markov inequality and then (20) one obtains

$$\begin{aligned}
&E\left(\max_{s\le t\le s+\delta}\frac{1}{\delta}(W_n(t)-W_n(s))^2 1_{\{\max\frac{1}{\delta}(W_n(t)-W_n(s))^2\ge K\}}\right)\\
\le\ &\frac{1}{K}\frac{1}{\delta^2}E\left(\max_{s\le t\le s+\delta}|W_n(t)-W_n(s)|\right)^4\\
\le\ &\frac{1}{K}\frac{1}{n^2\delta^2}E\left(\max_{s\le t\le s+\delta}\left|\sum_{i=\lfloor ns\rfloor+1}^{\lfloor nt\rfloor}X_i\right|\right)^4\\
\le\ &C\frac{1}{K}\frac{(\lfloor n(\delta+s)\rfloor-\lfloor ns\rfloor)^2}{n^2\delta^2}.
\end{aligned}$$

Because the last term tends to 0 as K tends to ∞, (19) is uniformly integrable and the partial sum process converges in $D[0,1]$ towards a Brownian Motion W with

$$W(1)=_{\mathcal{D}}N(0,\sigma^2).$$

Now we are able to treat the general case, where we still assume $EX_1=0$. We will prove the weak convergence by verifying the three conditions of Lemma 4.1 in Chen and White [8], which carries over from $C_H[0,1]$ to $D_H[0,1]$.
For fixed $h\in H\setminus\{0\}$ consider the sequence $(\langle X_i,h\rangle)_{i\in\mathbb{N}}$ of real valued random variables.
The mapping $x\mapsto\langle x,h\rangle$ is Lipschitz-continuous with constant $\|h\|$ and therefore by Lemma 2.2 of Dehling et.al. [12], $(\langle X_i,h\rangle)_{i\in\mathbb{N}}$ is L_1-near epoch dependent on an absolute regular process with approximation constants $(\|h\|a_m)_{m\in\mathbb{N}}$ and has further $4+\delta$-moments, because

$$E|\langle X_1,h\rangle|^{4+\delta}\le\|h\|E\|X_1\|^{4+\delta}<\infty.$$

Thus we can apply our sequential central limit theorem in $D[0,1]$ and get

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor nt\rfloor}\langle X_i,h\rangle\Rightarrow W_h(t),$$

where W_h is a Brownian motion with $EW_h(1)^2=\sigma^2(h)$ and

$$\sigma^2(h)=\sum_{i=-\infty}^{\infty}E(\langle X_0,h\rangle\langle X_i,h\rangle).$$

Define the covariance operator $S \in \mathcal{S}(H)$ as in (5), than $\langle Sh, h\rangle = \sigma^2(h)$ holds for all $h \in H \setminus \{0\}$.
By the arguments, used in the proofs of Theorems 4.3 and 4.6 in Chen and White [8]

$$(P_k W_n(t))_{t\in[0,1]} \Rightarrow \left(W^k(t)\right)_{t\in[0,1]} \quad \text{in } D_{H_k}[0,1].$$

Here H_k is the closed linear span of the first k elements of an arbitrary complete orthonormal basis of H. $P_k\colon H \to H_k$ is the orthonormal projection operator and X^k is a Brownian motion in $C_{H_k}[0,1]$, where $X^k(1)$ has the covariance operator $S_k = P_k S P_k$ (see Chen and White).
Thus condition (a) of Lemma 4.1 in Chen and White [8] is satisfied. For condition (b) we need $W^k \Rightarrow W$ in distribution, as k goes to ∞. But this holds, because S_k converges in the trace norm (see [8]) towards S.
Thus it remains to prove condition (c). We will show

$$\lim_{k\to\infty} \limsup_{n\to\infty} E\left(\sup_{t\in[0,1]} \|W_n(t) - P_k W_n(t)\|^4\right) = 0. \tag{21}$$

This is slightly different from the condition of Chen and White (we use fourth moments, while they consider second moments), but the Lemma remains true.
Define the operator $A_k\colon H \to H$ by $A_k = I - P_k$, where I is the identity operator on H, and note that the mapping $h \mapsto A_k(h)$ is Lipschitz-continuous with constant 1. Thus $(A_k(X_i))_{i\in\mathbb{N}}$ is a 1-approximating functional with the same constants as (X_i). From Lemma 2 it follows

$$E\|A_k(X_1) + \cdots + A_k(X_n)\|^4 \le Cn^2 \left(E\|A_k(X_1)\|^{4+\delta}\right)^{\frac{1}{1+\delta}}. \tag{22}$$

Now consider

$$E\left(\sup_{t\in[0,1]} \|W_n(t) - P_k W_n(t)\|^4\right) = \frac{1}{n^2} E\left(\max_{1\le m\le n} \left\|\sum_{i=1}^m A_k(X_i)\right\|^4\right)$$

and note that the term on the right hand side is bounded by $C\left(E\|A_k(X_1)\|^{4+\delta}\right)^{\frac{1}{1+\delta}}$, due to (22) and Lemma A.3. The constant C does not depend on k so it suffices to show

$$E\|A_k(X_1)\|^{4+\delta} \xrightarrow{k\to\infty} 0. \tag{23}$$

But because of the Hilbert space property

$$\|X_1 - \sum_{i=1}^k \langle X_1, e_i\rangle e_i\| \xrightarrow{k\to\infty} 0 \quad \text{a.s.}$$

Further $\|A_k(X_1)\|^{4+\delta} \le \|X_1\|^{4+\delta} < \infty$ almost surly and thus, by dominated convergence, (23) holds. But this implies (21) and therefore finishes the proof. $\square$

Proof of Theorem 2.2. Assume $EX_1 = 0$ and define

$$S^*_{n,i} := \frac{1}{\sqrt{p}} \sum_{j=(i-1)p+1}^{ip} (X^*_j - E^* X^*_j)$$

$$\text{and} \quad R^*_{n,kp}(t) := \frac{1}{\sqrt{kp}} \sum_{j=\lfloor kt \rfloor p+1}^{\lfloor kpt \rfloor} (X^*_j - E^* X^*_j).$$

Consider the following decomposition of the process $W_{n,kp}$ into the partial sum process of the independent blocks and the remainder

$$W^*_{n,kp}(t) = \frac{1}{\sqrt{k}} \sum_{i=1}^{\lfloor kt \rfloor} S^*_{n,i} + R^*_{n,kp}(t). \tag{24}$$

We start by proving that $R^*_{n,kp}$ is negligible, i.e.

$$R^*_{n,kp}(\cdot) \xrightarrow{P^*} 0 \quad \text{a.s.} \tag{25}$$

uniformly as $n \to \infty$. Note, that $R^*_{n,kp}(t)$ is the sum over the first l variables of a randomly generated block, where $l = l(k,p,t) = \lfloor kpt \rfloor - \lfloor kt \rfloor p$. Thus, for fixed t we have

$$\|R^*_{n,kp}(t)\| \leq \frac{1}{\sqrt{kp}} \max_{1 \leq j \leq k} \left\| \sum_{i=j(p-1)+1}^{j(p-1)+l} (X_i - E^* X^*_i) \right\|.$$

Taking the supremum over t, we get

$$\begin{aligned} \sup_{t \in [0,1]} \|R^*_{n,kp}(t)\| \leq & \frac{1}{\sqrt{kp}} \max_{1 \leq j \leq k} \max_{1 \leq l \leq p} \left\| \sum_{i=j(p-1)+1}^{j(p-1)+l} (X_i - E^* X^*_i) \right\| \\ =: & \, Y_{n,kp}. \end{aligned}$$

We will show, that $Y_{n,kp}$ converges to 0, almost surely.
For $n \in \{2^{l-1}+1, \cdots, 2^l\}$ observe

$$\begin{aligned} Y_n \leq & \frac{2}{\sqrt{2^l}} \max_{j \leq k(2^l)} \max_{1 \leq m \leq p(2^l)} \left\| \sum_{i=j(p-1)+1}^{j(p-1)+m} (X_i - E^* X^*_i) \right\| \\ =: & \, Y'_l. \end{aligned}$$

Taking the sum instead of the maximum, we obtain for the fourth moments of Y_l'

$$\begin{aligned}
E|Y_l'|^4 &\leq \frac{16}{2^{2l}} E\left(\max_{j\leq k(2^l)} \max_{m\leq p(2^l)} \left\| \sum_{i=j(p-1)+1}^{j(p-1)+m} (X_i - E^* X_i^*) \right\| \right)^4 \\
&\leq \frac{16}{2^{2l}} \sum_{j=1}^{k(2^l)} E\left(\max_{m\leq p(2^l)} \left\| \sum_{i=j(p-1)+1}^{j(p-1)+m} (X_i - E^* X_i^*) \right\| \right)^4 . \\
&= \frac{16k(2^l)}{2^{2l}} E\left(\max_{m\leq p(2^l)} \left\| \sum_{i=1}^{m} (X_i - E^* X_i^*) \right\| \right)^4 .
\end{aligned}$$

The last line holds since $(X_i)_{i\in\mathbb{N}}$ and $E^*X_i^*$ does not depend on the block in which X_i^* is, but only on the position of X_i^* in this block. We want to make use of Lemma A.3. For $p = p(2^l)$ and $k = k(2^l)$ we obtain using the Minkowski inequality

$$\begin{aligned}
E\left\| \sum_{i=1}^{p} (X_i - E^* X_i^*) \right\|^4 =& E\left\| \sum_{i=1}^{p} X_i - \frac{1}{k} \sum_{i=1}^{kp} X_i \right\|^4 \\
\leq& \left\{ \left(E\left\| \sum_{i=1}^{p} X_i \right\|^4 \right)^{1/4} + \left(E\left\| \frac{1}{k} \sum_{i=1}^{kp} X_i \right\|^4 \right)^{1/4} \right\}^4 \\
=& O(p^2).
\end{aligned}$$

In the last line we have used Lemma A.2 and the fact, that the first summand dominates.
Next by virtue of Lemma A.3 we obtain

$$E\left(\max_{m\leq p(2^l)} \left\| \sum_{i=1}^{m} (X_i - E^* X_i^*) \right\| \right)^4 = O(p^2)$$

Thus $E|Y_l'|^4 = O(\frac{p(2^l)}{2^l}) = O((2^{-\epsilon})^l)$, because of $p(n) = O(n^{1-\epsilon})$, see the definition of the block length. Now an application of the Markov inequality and Borel-Cantelli Lemma implies

$$Y_l' \xrightarrow{l\to\infty} 0 \quad \text{a.s.}$$

Now $Y_n \leq Y_l'$ for $n \in \{2^{l-1}, \cdots, 2^l\}$ and thus Y_n converges almost surely to 0 as n tends to infinity. Finally this leads to

$$E^*(\sup_{t\in[0,1]} \|R_n^*(t)\|) \leq E^* Y_n = Y_n \to 0 \quad \text{a.s.}$$

and thus we have proved (25).
So in order to verify convergence of the Bootstrap process in $D_H[0,1]$ it suffices to

show that

$$V^*_{n,kp}(t) = \frac{1}{\sqrt{k}} \sum_{i=1}^{\lfloor kt \rfloor} S^*_{n,i} \tag{26}$$

converges to the desired Gaussian process.

We start with the finite dimensional convergence. For $0 \leq t_1 < \cdots < t_l \leq 1$ and $l \geq 1$ consider the increments

$$(V^*_{n,kp}(t_1), V^*_{n,kp}(t_2) - V^*_{n,kp}(t_1), \cdots, V^*_{n,kp}(t_l) - V^*_{n,kp}(t_{l-1})). \tag{27}$$

Note, that the random variables $S^*_{n,i}$ are independent, conditional on $(X_i)_{i\in\mathbb{Z}}$, so it is enough to treat $V^*_{n,kp}(t_i) - V^*_{n,kp}(t_{i-1})$ for some $i \leq l$. By the consistency of the bootstrapped sample mean of H-valued data (see Dehling, Sharipov and Wendler [12]), there is a subset $A \subset \Omega$ with $P(A) = 1$, so that for all $\omega \in A$ the Central Limit Theorem holds, that is

$$\frac{1}{\sqrt{k}} \sum_{i=1}^{k} S^*_{n,i} \Rightarrow N, \tag{28}$$

where N is a Gaussian H-valued random variable with mean 0 and covariance operator $S \in \mathcal{S}$ defined by

$$\langle Sx, y\rangle = \sum_{i=-\infty}^{\infty} E[\langle X_0, x\rangle\langle X_i, y\rangle], \quad \text{for } x, y \in H.$$

For $\omega \in A$ and arbitrary $t_i > t_{i-1}$ it follows

$$V^*_{n,kp}(t_i) - V^*_{n,kp}(t_{i-1}) = \frac{1}{\sqrt{k}} \sum_{i=\lfloor kt_{i-1}\rfloor+1}^{\lfloor kt_i \rfloor} S^*_{n,i} \tag{29}$$

$$= \frac{\sqrt{\lfloor kt_i\rfloor - \lfloor kt_{i-1}\rfloor}}{\sqrt{k}} \frac{1}{\lfloor kt_i\rfloor - \lfloor kt_{i-1}\rfloor} \sum_{i=\lfloor kt_{i-1}\rfloor+1}^{\lfloor kt_i \rfloor} S^*_{n,i} \tag{30}$$

$$\Rightarrow \sqrt{t_i - t_{i-1}} N, \tag{31}$$

thus the one dimensional distributions converge almost surely. But because of the conditional independence this is enough for the finite dimensional convergence.

By Theorem A.1, tightness will follow if we can show

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \frac{1}{\delta} P^*\left(\sup_{0\leq t\leq\delta} \|V^*_{n,kp}(t)\| > \epsilon\right) = 0 \ \text{ a.s.} \tag{32}$$

for all $\epsilon > 0$.
Using first Chebychev's inequality and then Rosenthal's inequality (see Rosenthal [24] and Ledoux and Talagrand [20] for the validity in Hilbert spaces) we obtain

$$\begin{aligned}
&\frac{1}{\delta}P^*\left(\sup_{0\le t\le\delta}\frac{1}{\sqrt{k}}\left\|\sum_{i=1}^{\lfloor kt\rfloor}S^*_{n,i}\right\|>\epsilon\right)\\
&\le\frac{1}{\delta}\frac{1}{k^2\epsilon^4}E^*\left(\max_{1\le j\le\lfloor k\delta\rfloor}\left\|\sum_{i=1}^{j}S^*_{n,i}\right\|^4\right)\\
&\le\frac{1}{\delta}\frac{1}{k^2\epsilon^4}C\left\{\lfloor k\delta\rfloor E^*\|S^*_{n,1}\|^4+\left(\lfloor k\delta\rfloor E^*\|S^*_{n,1}\|^2\right)^2\right\}\\
&\le C\frac{1}{\delta}\frac{k\delta}{k^2\epsilon^4}E^*\|S^*_{n,1}\|^4+C\frac{1}{\delta}\frac{k^2\delta^2}{k^2\epsilon^4}(E^*\|S^*_{n,1}\|^2)^2\\
&=C(I_1+I_2).
\end{aligned}$$

By the construction of the bootstrap sample we get

$$\begin{aligned}
I_1=&\frac{1}{k\epsilon^4}\frac{1}{k}\sum_{i=1}^{k}\left(\frac{1}{\sqrt{p}}\left\|\sum_{j\in B_i}(X_j-\bar{X}_{n,kp})\right\|\right)^4\\
=&\frac{1}{\epsilon^4}\frac{1}{k}\sum_{i=1}^{k}\left(\frac{1}{k^{1/4}p^{1/2}}\left\|\sum_{j\in B_i}(X_j-\bar{X}_{n,kp})\right\|\right)^4.
\end{aligned}$$

Now by a strong Law of Large numbers (see Lemma 2.7 in Dehling et.al. [12])

$$\frac{1}{k^{1/4}p^{1/2}}\left\|\sum_{j\in B_i}(X_j-\bar{X}_{n,kp})\right\|\to 0 \quad \text{a.s.}$$

and thus I_1 goes almost surely to 0 as n goes to ∞.
Furthermore in Dehling, Sharipov and Wendler [12] it is shown that $E^*\|S^*_{n,i}\|^2$ converges almost surely to $E^*\|N\|^2$, where N is Gaussian with the covariance operator defined above. Therefore $E^*\|N\|^2$ is a.s. bounded and thus we obtain for the second term

$$I_2=\frac{\delta}{\epsilon^4}(E^*\|S^*_{n,1}\|^2)^2\xrightarrow{n\to\infty}\frac{\delta}{\epsilon^4}(E^*\|N\|^2)^2\xrightarrow{\delta\to\infty}0 \quad \text{a.s.}$$

which implies (32) and therefore finishes the proof. □

Proof of Corollary 3.2. Part (i) can be obtained by arguments, similar to the case of real-valued random variables, see Theorem 2.1 in Dehling, Rooch and Taqqu [11].

To verify part (ii) define random variables $U_1,\dots U_k$, where U_i is the number of the ith drawn block. Clearly the U_i are all independent and uniformly distributed on $\{1,\dots,k\}$.

Note, that the random variables in the blocks $B_1, \ldots, B_{\lfloor k\tau \rfloor}$ are of the form X_i and the variables of the blocks $B_{\lfloor k\tau\rfloor+2}, \ldots, B_k$ are of the form $X_i + \Delta_n$. The change point occurs in the block $B_{\lfloor k\tau\rfloor+1}$, so this block contains shifted and non-shifted variables. This subdivision in different types of blocks leads to the following decomposition of the process

$$\frac{1}{\sqrt{kp}}\left(\sum_{i=1}^{\lfloor kpt\rfloor} Y^*_{n,i} - \frac{\lfloor kpt\rfloor}{kp}\sum_{i=1}^{kp} Y^*_{n,i}\right) = \frac{1}{\sqrt{kp}}\left(\sum_{i=1}^{\lfloor kpt\rfloor} X^*_{i} - \frac{\lfloor kpt\rfloor}{kp}\sum_{i=1}^{kp} X^*_{i}\right) + \sqrt{kp}\Delta_n R_{n,k,p}(t),$$

where

$$\begin{aligned}
R_{n,k,p}(t) = {} & \frac{1}{kp}p\sum_{i=1}^{\lfloor kt\rfloor} 1_{\{U_i > \lfloor k\tau\rfloor+1\}} && (33)\\
& - \frac{1}{kp}p\frac{\lfloor kpt\rfloor}{kp}\sum_{i=1}^{k} 1_{\{U_i > \lfloor k\tau\rfloor+1\}} && (34)\\
& + \frac{1}{kp}((\lfloor kt\rfloor)+1)p - \lfloor n\tau\rfloor)\sum_{i=1}^{\lfloor kt\rfloor} 1_{\{U_i = \lfloor k\tau\rfloor+1\}} && (35)\\
& - \frac{1}{kp}((\lfloor kt\rfloor)+1)p - \lfloor n\tau\rfloor)\frac{\lfloor kpt\rfloor}{kp}\sum_{i=1}^{k} 1_{\{U_i = \lfloor k\tau\rfloor+1\}} && (36)\\
& + 1_{\{U_{\lfloor kt\rfloor+1} > \lfloor k\tau\rfloor+1\}}\frac{1}{kp}(\lfloor kpt\rfloor - \lfloor kt\rfloor p) && (37)\\
& + 1_{\{U_{\lfloor kt\rfloor+1} = \lfloor k\tau\rfloor+1\}}\frac{1}{kp}\max\{(\lfloor kpt\rfloor - \lfloor n\tau\rfloor p), 0\}. && (38)
\end{aligned}$$

By part (ii) of Corollary 3.1 and $\sqrt{kp}\Delta_n \to \Delta$ it remains to show

$$R_{n,k,p} \xrightarrow{P^*} 0 \quad \text{a.s.}$$

uniformly as $n \to \infty$. But this holds because $R_{n,k,p}$ is independent of the X_i and due to the following facts: (33) + (34) and (35) + (36) are $o_P(1)$, each. To see this observe

$$\frac{1}{k}\sum_{i=1}^{\lfloor kt\rfloor} 1_{\{U_i > \lfloor k\tau\rfloor+1\}} \xrightarrow{P} t(1-\tau),$$

uniformly in t as $k \to \infty$.
(37) is $o_P(1)$ because $(\lfloor kpt\rfloor - \lfloor kt\rfloor p)/(kp) \to 0$ and finally (38) is $o_P(1)$ because $P(U_{\lfloor kt\rfloor+1} = \lfloor k\tau\rfloor + 1) = k^{-1}$.

□