TI 2026-055/III
Tinbergen Institute Discussion Paper

Confidence Sets for the Date of a Weak Mean Break in Functional Data

Yicong Lin[1]

1 Vrije Universiteit Amsterdam and Tinbergen Institute

Confidence Sets for the Date of a Weak Mean Break in Functional Data

Yicong Lin*

Vrije Universiteit Amsterdam and Tinbergen Institute

August, 2026

Abstract

We develop confidence sets for the date of a single mean break in functional data when the break may be too weak to be consistently detected. Under each maintained null, segmentwise demeaning removes the unknown mean and break functions, so valid inference does not require consistent detection of the break. We select among invariant tests by maximizing their weighted average local power over alternative break dates and directions. In infinite dimensions, the resulting covariance perturbation can render the null and alternative measures mutually singular, causing the usual likelihood-based derivation of a test that maximizes weighted average power to break down. We characterize the weights under which likelihood-based comparison remains valid and show that covariance-squared weighting yields a simple locally best invariant statistic. Under mild conditions, we establish the asymptotic validity of the procedure and characterize its local power under weak breaks. Simulations show that the resulting confidence sets achieve accurate empirical coverage, whereas a competing interval designed for consistently detectable breaks exhibits severe undercoverage when the break magnitude is small.

Keywords: confidence set; functional data; infinite dimensional inference; locally best invariant test; weak mean break.

*Department of Econometrics and Data Science, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV, Amsterdam, the Netherlands. E-mail address: yc.lin@vu.nl.

1 Introduction

A growing literature has studied the detection and dating of breaks in functional observations. A particularly relevant contribution is Aue et al. (2018), which develops a fully functional procedure for detecting and dating a mean break without first projecting the observations onto a finite dimensional space. Recent contributions, among many others, have studied heterogeneous mean breaks in high dimensional functional time series (Li et al., 2024), mean breaks and more general distributional changes using empirical energy distance (Boniece et al., 2025), and mean breaks under local stationarity and partial measurement errors (Bai et al., 2026). These advances have extended the literature to more complex data structures, but existing results for inference on break dates typically require the breaks to be sufficiently large for consistent detection.

In practice, a mean break may be difficult to distinguish from sampling variation when the sample is short, the break is small relative to the noise, or the observations are strongly dependent. Yet quantifying uncertainty about the break date remains essential, since the estimated date is often used to link the underlying change to an external event. Existing asymptotic theories typically assume either a fixed break magnitude or one that shrinks sufficiently slowly for the accumulated signal to diverge, so that the break eventually becomes consistently detectable. The confidence interval of Aue et al. (2018), for example, is constructed under such a framework. We instead consider local breaks from the outset, for which the signal remains comparable to sampling variation, so that the break cannot be consistently detected and uncertainty about its date persists even asymptotically. To address this problem, we develop confidence sets for the date of a single mean break without requiring consistent detection of the break.

We construct confidence sets by inverting a sequence of invariant tests of candidate break dates. Specifically, for each candidate date k, we test the null hypothesis that the break occurred at k against the alternative that it occurred elsewhere, and retain all dates that are not rejected, as in Elliott and Müller (2007) for finite dimensional structural break models. For each maintained date, we demean the observations separately within the two segments determined by k. When the maintained date is correct, this transformation eliminates both the unknown mean and the unknown break function exactly. The null distribution is therefore free of the break magnitude, so valid inference does not require consistent detection of the break.

Invariance alone does not determine a unique test. We therefore compare invariant tests using weighted average local power over alternative break dates and break directions, as in the general optimal testing framework of Andrews and Ploberger (1994). A test that maximizes this criterion is locally best invariant (LBI). Although this comparison is straightforward in finite dimensions, extending it to an infinite dimensional space is not. Averaging over break directions changes the covariance structure of the alternative distribution. In finite dimensions, a small covariance perturbation leaves the alternative sufficiently close to the null for the two distributions to be compared through a likelihood ratio. In infinite dimensions, however, the effects of even a small change can accumulate across infinitely many directions, causing the likelihood ratio to cease to exist. The usual local power argument then breaks down altogether. We overcome this difficulty by characterizing the weights over break directions for which likelihood-based comparison remains valid. Covariance-squared weighting satisfies this requirement and yields a simple LBI statistic based directly on the Hilbert space norm, without requiring covariance inversion or the associated tuning choices.

Beyond this optimality result, we establish the asymptotic theory under mild conditions. The null theory accommodates arbitrary deterministic break sequences and allows the long-run covariance operators to differ across the two maintained segments. It yields feasible critical values and establishes asymptotically valid coverage of the inverted confidence set without requiring consistent detection of the break. We also derive local power under weak breaks, thereby showing how incorrect dates are excluded and how informative the resulting confidence set is. The simulations show that the LBI-based confidence sets attain empirical coverage close to the nominal level, albeit with greater length. By contrast, the interval proposed by Aue et al. (2018) understates uncertainty about the break date under weak breaks, thereby producing shorter intervals with substantial undercoverage in finite samples.

The remainder of the paper is organized as follows. Section 2 formulates the testing problem and introduces the weighted average power criterion. Section 3 derives the locally best invariant test under a functional Gaussian benchmark. Section 4 establishes the limiting distributions under the null and local alternatives, together with the coverage properties of the inverted confidence set. Section 5 discusses implementation and presents the simulation results. Section 6 concludes. Proofs and additional simulation results are provided in the Online Appendix.

2 The functional break date problem

Let $\mathcal{H}$ be a real separable Hilbert space equipped with the norm $\|\cdot\|_{\mathcal{H}}$, and consider the following model with a single change in mean:

$$X_s = \mu + \delta\, \mathbb{1}\{s > k_0\} + \varepsilon_s, \qquad s = 1, \dots, n, \tag{2.1}$$

where $\mu, \delta \in \mathcal{H}$, $\mathbb{1}\{\cdot\}$ is the indicator function, and $k_0 = \lfloor n\tau_0 \rfloor$ for some $\tau_0 \in (0, 1)$. The functional observations $X_s \in \mathcal{H}$ are the variables of interest, while $\{\varepsilon_s\}$ is a mean-zero error process satisfying Assumption 1 below. The true break date k_0 is unknown and must be inferred from the data. In this paper, we focus on constructing confidence sets for k_0 when the change in mean δ is small, a setting in which uncertainty about the existence of a break and uncertainty about its location are intrinsically related. Specifically, we allow the break function $\delta = \delta_n$ to depend on n and, in particular, consider the local specification $\delta_n = d/\sqrt{n}$ for a fixed $d \in \mathcal{H}$. This differs from the shrinking break framework considered by Aue et al. (2018, Theorem 4), which allows $\|\delta_n\|_{\mathcal{H}} \to 0$ while requiring $n\|\delta_n\|^2_{\mathcal{H}} \to \infty$. The latter condition places the break in a regime in which it can be detected consistently. Under the local specification considered here, by contrast, $n\|\delta_n\|^2_{\mathcal{H}} = \|d\|^2_{\mathcal{H}} < \infty$, so the break signal remains of local order and cannot be detected with probability approaching one as $n \to \infty$.

The difficulty of dating a small break motivates an inference procedure that does not rely on first obtaining a precise estimate of k_0. A natural approach is to construct a confidence set by inverting a sequence of hypothesis tests; see, for example, Beutner et al. (2023) for an overview of different confidence set constructions for trend breaks in scalar time series. Specifically, for each candidate date $k \in K_n := \{2, \dots, n-2\}$ we consider the test

$$H_0(k) : k_0 = k \qquad \text{against} \qquad H_1(k) : k_0 \neq k. \tag{2.2}$$

A confidence set for k_0 is then obtained by collecting the candidate dates for which $H_0(k)$ is not rejected. Under $H_0(k)$, both the pre-break mean μ and the break function δ are nuisance parameters. Since the object of interest is the break date rather than these nuisance parameters, we adopt the invariance approach of King (1987) and King and Wu (1997), and restrict attention to tests invariant to the transformation $X_s \mapsto X_s + a + b\,\mathbb{1}\{s > k\}$ for arbitrary $a, b \in \mathcal{H}$; see also Elliott and Müller (2006, 2007). Under the maintained null, this transformation changes only the mean functions of the two segments while leaving the break date unchanged. Consequently, an invariant test can eliminate dependence on the nuisance parameters (μ, δ), which is particularly important in the local break setting considered here.

To implement this invariance restriction, we remove the sample mean separately from each of the two segments determined by the maintained date k. Let $\overline{X}_{1:k} = k^{-1} \sum_{s=1}^{k} X_s$ and $\overline{X}_{k+1:n} = (n-k)^{-1} \sum_{s=k+1}^{n} X_s$, and define

$$\widehat{\varepsilon}_{k,s} = \begin{cases} X_s - \overline{X}_{1:k}, & s \leq k, \\ X_s - \overline{X}_{k+1:n}, & s > k. \end{cases} \tag{2.3}$$

The residual sequence in (2.3) is unchanged under the transformation $X_s \mapsto X_s + a + b\,\mathbb{1}\{s > k\}$ and forms a maximal invariant: two samples give the same residual sequence if and only if they differ by separate constant shifts on the two segments. Hence any invariant test of $H_0(k)$ can be constructed as a function of these residuals. When the maintained date coincides with the true break date, $k = k_0$, the nuisance parameters (μ, δ) are removed exactly, so the null distribution of any invariant statistic constructed from (2.3) is free of (μ, δ).

Having removed the nuisance mean and break function under the maintained null, we next turn to the choice among invariant tests. The alternative is composite: while the maintained break date k is fixed under $H_0(k)$, under $H_1(k)$ the true break may occur at any $k_0 = t$ with $t \neq k$ and with any break function $\delta \in \mathcal{H}$. Consequently, there is generally no uniformly most powerful invariant test over the entire collection of alternatives. Moreover, under the invariance requirement imposed above, the break function δ is not identified under the null in the invariant experiment, since the distribution of the maximal invariant does not depend on δ when $k_0 = k$, whereas it generally does under the alternative. This gives rise to a nonstandard testing problem involving a nuisance parameter that is not identified under the null, closely related to the class of problems studied by Andrews and Ploberger (1994) and Andrews (1999). We follow Andrews and Ploberger (1994) and select among invariant tests using a weighted average power criterion.

To this end, for each maintained date k, let $\varphi_{n,k}$ denote an invariant test of $H_0(k)$ with null rejection probability α, where $\varphi_{n,k} = 1$ indicates rejection of $H_0(k)$ and $\varphi_{n,k} = 0$ otherwise. Throughout, sums over $t \neq k$ are understood to run over $t \in K_n$, i.e., $\sum_{t\neq k} = \sum_{t\in K_n, t\neq k}$. Let $w_{k,t} \geq 0$ denote weights assigned to the alternative dates, normalized such that $\sum_{t\neq k} w_{k,t} = 1$. The weighted average power of $\varphi_{n,k}$ is then defined as

$$P_{n,k}(\varphi_{n,k}; b^2) = \sum_{t\neq k} w_{k,t} \int_{\mathcal{H}} \mathbb{E}_{t,\delta}\big(\varphi_{n,k}\big)\, \mathrm{d}\nu_{k,t,b^2}(\delta), \tag{2.4}$$

where ν_{k,t,b^2} is a probability measure on $\mathcal{H}$ used to weight possible break functions when the alternative break date is t, and $\mathbb{E}_{t,\delta}$ denote expectation under the model with $k_0 = t$ and break function δ. The parameter $b^2 \geq 0$ indexes the overall scale of the break functions under this weighting distribution. We require $\nu_{k,t,0}$ to place all its mass on the zero function, so that the weighted alternatives collapse to the null as $b^2 \downarrow 0$; see, e.g., Eq. (3.1) in the next section.

Eq. (2.4) averages power over the two dimensions of the alternative. Specifically, (1) for a fixed alternative date $t \neq k$, $\mathbb{E}_{t,\delta}(\varphi_{n,k})$ is the rejection probability when the true break occurs at t with break function $\delta \in \mathcal{H}$, and hence the power of the maintained date test against the particular alternative (t, δ). The integral $\int_{\mathcal{H}} \mathbb{E}_{t,\delta}\big(\varphi_{n,k}\big)\, \mathrm{d}\nu_{k,t,b^2}(\delta)$ then averages this power over possible break functions according to ν_{k,t,b^2}. (2) The outer sum then averages these date-dependent powers across all alternative dates $t \neq k$, with $w_{k,t}$ determining the relative importance assigned to each date. Note that the measures ν_{k,t,b^2} and the date weights $w_{k,t}$ serve only to define the power criterion and need not be interpreted as priors on the break function or break date. Different choices induce different notions of average power and may therefore lead to different optimal tests; see, for example, Kurozumi and Yamamoto (2015) in the context of a finite dimensional break date problem. In addition, for an invariant test with null rejection probability α, the degeneration of ν_{k,t,b^2} at $b^2 = 0$ implies $\int_{\mathcal{H}} \mathbb{E}_{t,\delta}\big(\varphi_{n,k}\big)\, \mathrm{d}\nu_{k,t,0}(\delta) = \mathbb{E}_{t,0}\big(\varphi_{n,k}\big) = \alpha$. Hence, $P_{n,k}(\varphi_{n,k}; 0) = \sum_{t\neq k} w_{k,t}\alpha = \alpha$, where the last equality follows from the normalization $\sum_{t \neq k} w_{k,t} = 1$. Thus, all invariant tests with null rejection probability α have the same weighted average power at $b^2 = 0$.

Among invariant tests with null rejection probability α, a locally best invariant test is defined as the test whose weighted average power increases most rapidly as the scale of the weighted break functions moves away from zero. Formally, it maximizes

$$\frac{\partial}{\partial b^2} P_{n,k}(\varphi_{n,k}; b^2) \Big|_{b^2=0+}. \tag{2.5}$$

Implementing this criterion requires a specific choice of the weighting measures ν_{k,t,b^2} for the break functions. This choice requires particular care, because weighting schemes that are well behaved in finite dimensions may fail to admit a likelihood ratio with respect to the null distribution in an infinite dimensional space, preventing the usual likelihood ratio expansion around $b^2 = 0$. Section 3 therefore introduces a weighting scheme that avoids this difficulty and derives the corresponding locally best invariant test.

3 Constructing locally best invariant tests

Let $\mathcal{N}_{\mathcal{H}}(\mu, \mathcal{C})$ denote the Gaussian probability measure on $\mathcal{H}$ with mean $\mu \in \mathcal{H}$ and covariance operator $\mathcal{C}$. We derive locally best invariant (LBI) tests in a Gaussian benchmark experiment in which $\varepsilon_t \sim \mathcal{N}_{\mathcal{H}}(0, \mathcal{C}_\varepsilon)$ independently for $t = 1, \ldots, n$. Hence, $\mathbb{E}\,\|\varepsilon_t\|_{\mathcal{H}}^2 = \operatorname{tr}(\mathcal{C}_\varepsilon) < \infty$, where $\operatorname{tr}(\cdot)$ denotes the operator trace (see, e.g., Da Prato and Zabczyk, 2014, Proposition 2.16). Moreover, for every $f \in \mathcal{H}$, $\langle \varepsilon_t,\, f \rangle_{\mathcal{H}}$ is Gaussian with mean zero and variance $\langle \mathcal{C}_\varepsilon f,\, f \rangle_{\mathcal{H}}$. Note that Gaussianity and independence are imposed only for the optimality derivation; the validity of the resulting tests will be established under weaker assumptions in Section 4, without relying on Gaussianity and allowing for serial dependence.

Under the weighted average power criterion (2.4), averaging over the alternative break dates and break functions defines a single averaged alternative distribution. Whenever this distribution is absolutely continuous with respect to the maintained null distribution, the weighted average power of a test can be expressed as the null expectation of the test function multiplied by the corresponding likelihood ratio. Provided that this likelihood ratio is right differentiable at $b^2 = 0$ and that differentiation can be interchanged with expectation, the LBI problem (2.5) reduces to maximizing the null expectation of the test function multiplied by the right derivative of the likelihood ratio; see Eq. (A.6) in Online Appendix A for details. A Neyman–Pearson-type argument then implies rejection for large values of this derivative; see Lehmann and Romano (2005, Theorem 3.2.1).

3.1 The choice of weights

We now implement the weighted average power criterion in (2.4). We make two choices. First, we assign equal weight to all alternative break dates, setting $w_{k,t} = (n-4)^{-1}$ for $t \in K_n$ and $t \neq k$. We maintain the same weighting criterion throughout the paper and leave the investigation of alternative date-weighting schemes for future research. Second, for each alternative date $t \neq k$, we choose the break function weighting measure as

$$\nu_{k,t,b^2} = \mathcal{N}_{\mathcal{H}}\big(0, b^2\, a_{k,t}\, \mathcal{C}_\delta\big), \qquad a_{k,t} = \begin{cases} k^{-2}, & t < k, \\ (n-k)^{-2}, & t > k. \end{cases} \tag{3.1}$$

The scaling $a_{k,t}$ depends on which of the two segments induced by the maintained date contains the alternative break. If $t < k$, then under the alternative the true break occurs within observations $1, \ldots, k$. The transformation in (2.3) therefore does not eliminate the break, and its effect on the corresponding residual partial sums is of order $k\delta$. Scaling the covariance of δ by k^{-2} makes its standard deviation proportional to k^{-1}, so the accumulated shift remains of the same order across maintained dates. The same argument yields the factor $(n-k)^{-2}$ when $t > k$. The covariance operator $\mathcal{C}_\delta$, in turn, determines how different directions of the break function are weighted.

To see further the role of the Gaussian weighting in (3.1), fix an alternative date $t \neq k$, so that $k_0 = t$ under the corresponding alternative, and define $S_{k,t} = \sum_{s=1}^{t} \widehat{\varepsilon}_{k,s}$ if $t < k$ and $S_{k,t} = \sum_{s=k+1}^{t} \widehat{\varepsilon}_{k,s}$ if $t > k$, where $\widehat{\varepsilon}_{k,s}$ is defined in (2.3). Note that, for a fixed alternative date $t \neq k$, the break induces a single piecewise constant mean pattern within the segment containing t, namely either $\{1, \ldots, k\}$ or $\{k+1, \ldots, n\}$. One can show that, whenever it is well defined, the likelihood ratio of the Gaussian averaged alternative with respect to the null depends on $\{\widehat{\varepsilon}_{k,s},\ s = 1, \ldots, n\}$ only through $S_{k,t}$. Under the alternative with break function $\delta \in \mathcal{H}$,

$$S_{k,t} \mid \delta \sim \mathcal{N}_{\mathcal{H}}\big(- q_{k,t}\, \delta,\, q_{k,t}\, \mathcal{C}_{\varepsilon} \big), \qquad q_{k,t} = \begin{cases} \dfrac{t(k-t)}{k}, & t < k, \\ \dfrac{(t-k)(n-t)}{n-k}, & t > k. \end{cases} \tag{3.2}$$

Thus, relative to the maintained null, conditioning on δ changes only the mean of $S_{k,t}$, from zero to $-q_{k,t}\,\delta$, while its covariance remains $q_{k,t}\,\mathcal{C}_{\varepsilon}$. Averaging over the centered Gaussian weighting distribution in (3.1) then restores the zero mean but adds the variation of the random mean shift to the covariance. More specifically, $S_{k,t}$ has unconditional mean zero and covariance operator $q_{k,t}\,\mathcal{C}_{\varepsilon} + b^2\, a_{k,t}\, q_{k,t}^2\, \mathcal{C}_{\delta}$ by the law of total linear covariance; see, e.g., Klebanov et al. (2021, Theorem 4.5(g)). Rescaling $S_{k,t}$ by $q_{k,t}^{-1/2}$ yields a $\mathcal{H}$-valued Gaussian random element with covariance operator $\mathcal{C}_{\varepsilon} + \eta_{k,t}\, \mathcal{C}_{\delta}$, where $\eta_{k,t} = b^2\, a_{k,t}\, q_{k,t}$.

As noted above, the likelihood ratio of the Gaussian-averaged alternative with respect to the null depends on the residuals defined in (2.3) only through $S_{k,t}$. Gaussian averaging therefore affects the likelihood ratio solely through the covariance perturbation in the distribution of $S_{k,t}$. Consequently, the derivative of the weighted average power in (2.4) is given by the null expectation of the test function multiplied by the weighted sum of the derivatives at $b^2 = 0$ of the corresponding likelihood ratios. The LBI criterion in (2.5) can thus be analyzed by examining how the distribution of $S_{k,t}$ changes as its covariance operator is perturbed away from its null value. However, by the fundamental Feldman–Hájek dichotomy (e.g., Da Prato and Zabczyk, 2014, Theorem 2.25), the perturbed Gaussian measure need not admit a likelihood ratio with respect to the null measure for an arbitrary choice of $\mathcal{C}_{\delta}$ in (3.1) in an infinite dimensional space; see the further discussion below. This issue arises from the infinite dimensionality of the problem and does not occur in finite dimensional settings. A likelihood expansion around $\eta_{k,t} = 0$, or equivalently as $b^2 \downarrow 0$, is therefore substantially more delicate in an infinite dimensional space than in a finite dimensional setting; compare Elliott and Müller (2007). It requires carefully formulated restrictions on $\mathcal{C}_{\delta}$, which we introduce below.

3.2 Covariance weighting in infinite dimensions

The relevant restriction concerns not only the magnitude of $\mathcal{C}_\delta$, but also the covariance perturbation it induces relative to the null covariance operator. To make this restriction explicit, we consider weighting covariance operators of the form:

$$\mathcal{C}_\delta = \mathcal{C}_\varepsilon^{1/2}\,\mathcal{K}\,\mathcal{C}_\varepsilon^{1/2}, \tag{3.3}$$

where $\mathcal{K}$ is a nonnegative self-adjoint operator on $\mathcal{H}$. Under this representation, the covariance operator $\mathcal{C}_\varepsilon + \eta_{k,t}\mathcal{C}_\delta$ of the rescaled partial sum can be written as $\mathcal{C}_\varepsilon^{1/2}(\mathcal{I} + \eta_{k,t}\mathcal{K})\mathcal{C}_\varepsilon^{1/2}$ with $\eta_{k,t} = b^2 a_{k,t} q_{k,t}$ as previously defined and $\mathcal{I}$ denotes the identity operator on $\mathcal{H}$. We require $\mathcal{K}$ to be Hilbert–Schmidt; see, for example, Da Prato and Zabczyk (2014, Appendix C) for the definition. By the equivalence criterion for Gaussian measures in Bogachev (1998, Corollary 6.4.11), the Hilbert–Schmidt condition on $\mathcal{K}$ ensures that the perturbed and null Gaussian measures are equivalent; see also Da Prato and Zabczyk (2014, Theorem 2.25). Their equivalence, in turn, implies the existence of the corresponding likelihood ratio by the Radon–Nikodym theorem. The Hilbert–Schmidt condition ensures equivalence of the perturbed and null Gaussian measures and hence the existence of the corresponding likelihood ratio.

The same representation also identifies choices of $\mathcal{C}_\delta$ that are not admissible in an infinite dimensional space. Consider $\mathcal{C}_\delta = \mathcal{C}_\varepsilon$. For every eigenfunction ϕ_j of $\mathcal{C}_\varepsilon$ associated with a positive eigenvalue λ_j, any operator $\mathcal{K}$ satisfying $\mathcal{C}_\varepsilon^{1/2}\,\mathcal{K}\,\mathcal{C}_\varepsilon^{1/2} = \mathcal{C}_\varepsilon$ must satisfy $\langle \mathcal{K}\phi_j,\, \phi_j\rangle_{\mathcal{H}} = 1$. Consequently, if $\mathcal{C}_\varepsilon$ has infinite rank, then the Cauchy–Schwarz inequality gives $\sum_{j=1}^{\infty} \|\mathcal{K}\phi_j\|_{\mathcal{H}}^2 \geq \sum_{j=1}^{\infty} \left|\langle \mathcal{K}\phi_j,\, \phi_j\rangle_{\mathcal{H}}\right|^2 = \sum_{j=1}^{\infty} 1 = \infty$. Thus, $\mathcal{K}$ cannot be Hilbert–Schmidt. Under this specification, the covariance operator of the rescaled partial sum $q_{k,t}^{-1/2} S_{k,t}$ is $(1+\eta_{k,t})\mathcal{C}_\varepsilon$. By the Feldman–Hájek dichotomy, whenever $\eta_{k,t} > 0$, the corresponding Gaussian measure is singular with respect to the null measure, so the likelihood ratio required for the expansion in (2.5) does not exist.

One spectral family is $\mathcal{C}_\delta = \mathcal{C}_\varepsilon^{1+\gamma}$ for $\gamma > 0$, which corresponds to $\mathcal{K} = \mathcal{C}_\varepsilon^{\gamma}$ in (3.3). With the eigenpairs introduced above, $\mathcal{K}\phi_j = \lambda_j^{\gamma}\phi_j$, so $\mathcal{K}$ is Hilbert–Schmidt if and only if $\sum_{\lambda_j>0} \|\mathcal{K}\phi_j\|_{\mathcal{H}}^2 = \sum_{\lambda_j>0} \lambda_j^{2\gamma} < \infty$. Since every summable nonnegative sequence is p-summable for every $p \geq 1$, the condition $\operatorname{tr}(\mathcal{C}_\varepsilon) < \infty$ guarantees the required Hilbert–Schmidt property for every $\gamma \geq 1/2$. Different values of γ generally lead to different local scores and hence to different LBI tests.

We set $\gamma = 1$, so that $\mathcal{C}_\delta = \mathcal{C}_\varepsilon^2$. As shown in the proof of Theorem 1 (see (A.9) in Online Appendix A), for each alternative date $t \neq k$, the data-dependent component of the right derivative of the corresponding likelihood ratio with respect to b^2, evaluated at $b^2 = 0$, is $a_{k,t} \|S_{k,t}\|_{\mathcal{H}}^2 /2$. Consequently, choosing $\gamma = 1$ yields a statistic based on the usual Hilbert norm and requires neither inversion of the covariance operator nor the associated selection of a functional principal component truncation level or regularization parameter. Combining this weighting scheme with the date weights and scaling specified in Section 3.1 yields the functional locally best invariant (fLBI) test defined by the statistic in Eq. (3.4) below.

Theorem 1 (Functional locally best invariant test)**.** *Fix $k \in K_n$. In the Gaussian benchmark experiment, consider invariant tests $\varphi_{n,k}$ of $H_0(k)$ with null rejection probability α under the weighted average power criterion in (2.4). Suppose that $w_{k,t} = (n-4)^{-1}$ for every $t \in K_n$ and $t \neq k$, and set $\mathcal{C}_\delta = \mathcal{C}_\varepsilon^2$ in (3.1). Then the right derivative in (2.5) is maximized by a test that rejects $H_0(k)$ for large values of*

$$U_n(k) = \frac{1}{k^2} \sum_{t=2}^{k-1} \|S_{k,t}\|_{\mathcal{H}}^2 + \frac{1}{(n-k)^2} \sum_{t=k+1}^{n-2} \|S_{k,t}\|_{\mathcal{H}}^2 . \tag{3.4}$$

Here, $S_{k,t} = \sum_{s=1}^{t} \widehat{\varepsilon}_{k,s}$ if $t < k$ and $S_{k,t} = \sum_{s=k+1}^{t} \widehat{\varepsilon}_{k,s}$ if $t > k$, where $\widehat{\varepsilon}_{k,s}$ is defined in (2.3).

The fLBI test statistic in (3.4) is somewhat related to the functional CUSUM statistic of Aue et al. (2018) in that both are constructed from squared Hilbert norms of tied-down partial sum processes. The processes themselves, however, are different. Their CUSUM process is demeaned over the full sample and is tied down at the two sample endpoints. In contrast, for a maintained date k, our residual partial sums are demeaned separately within the two maintained segments and are therefore tied down at the endpoints of their respective segments. Moreover, their procedure maximizes the squared norm over candidate dates, whereas (3.4) aggregates the date-specific squared norms.

4 Asymptotic inference and confidence sets

The Gaussianity and independence assumptions in Section 3 are used to construct the fLBI test and establish its exact optimality; they are not required for its asymptotic validity. Next we establish the asymptotic validity under non-Gaussian and weak dependence, allowing the long-run covariance operator to differ before and after the mean break. Under these broader conditions, the *exact* fLBI optimality result need not continue to hold, and a procedure tailored to the full dependence structure may have different power properties. For real constants $c_1 < c_2$, let $\mathcal{D}_{\mathcal{H}}[c_1, c_2]$ denote the space of $\mathcal{H}$-valued càdlàg functions on $[c_1, c_2]$.

Assumption 1. *Let $k_0 = \lfloor n\tau_0 \rfloor$ for a fixed $\tau_0 \in (0, 1)$. As $n \to \infty$,*

$$\left(\left\{ \frac{1}{\sqrt{n}} \sum_{s=1}^{\lfloor nr \rfloor} \varepsilon_s \right\}_{0 \le r \le \tau_0}, \left\{ \frac{1}{\sqrt{n}} \sum_{s=k_0+1}^{\lfloor nr \rfloor} \varepsilon_s \right\}_{\tau_0 \le r \le 1} \right) \Rightarrow \left(\{\mathcal{W}_1(r)\}_{0 \le r \le \tau_0}, \{\mathcal{W}_2(r) - \mathcal{W}_2(\tau_0)\}_{\tau_0 \le r \le 1} \right). \tag{4.1}$$

Here, $\Rightarrow$ denotes weak convergence in $\mathcal{D}_{\mathcal{H}}[0, \tau_0] \times \mathcal{D}_{\mathcal{H}}[\tau_0, 1]$, equipped with the product Skorokhod topology. The processes $\mathcal{W}_1$ and $\mathcal{W}_2$ are independent $\mathcal{H}$-valued Brownian motions on $[0, 1]$ with covariance operators $\mathcal{C}_1$ and $\mathcal{C}_2$, respectively, satisfying $\operatorname{tr}(\mathcal{C}_1) < \infty$ and $\operatorname{tr}(\mathcal{C}_2) < \infty$.

Assumption 1 imposes a high-level joint functional central limit theorem (FCLT) for the pre- and post-break partial sums under full sample normalization, allowing the long-run covariance operators to differ across the two segments. Primitive sufficient conditions for FCLTs of this type are available under near-epoch dependence on mixing processes and under Bernoulli-shift approximability; see, for instance, Chen and White (1998) and Berkes et al. (2013), respectively. The independence of $\mathcal{W}_1$ and $\mathcal{W}_2$ is an asymptotic condition and does not require the observations on opposite sides of the break to be independent in finite samples. Under suitable short-range dependence conditions, the normalized pre- and post-break partial sum processes become asymptotically uncorrelated. Since their joint limit is Gaussian, the limiting processes are then independent.

Theorem 2 (Null limiting distribution)**.** *Let $\mathcal{B}_\ell(u) = \mathcal{W}_\ell(u) - u\,\mathcal{W}_\ell(1)$ for $\ell = 1, 2$.*

Under Assumption 1,

$$U_n(k_0) \xrightarrow{d} \int_0^1 \left\|\mathcal{B}_1(u)\right\|_{\mathcal{H}}^2 \mathrm{d}u + \int_0^1 \left\|\mathcal{B}_2(u)\right\|_{\mathcal{H}}^2 \mathrm{d}u, \qquad n \to \infty, \tag{4.2}$$

where $\xrightarrow{d}$ *denotes convergence in distribution.*

The limiting distribution in (4.2) is independent of μ, δ_n, and τ_0, but depends on the eigenvalue sequences of $\mathcal{C}_1$ and $\mathcal{C}_2$. Specifically, let $\{\lambda_{\ell,j},\, j \geq 1\}$ denote the eigenvalues of $\mathcal{C}_1$ for $\ell = 1$ and of $\mathcal{C}_2$ for $\ell = 2$. The corresponding Karhunen–Loève representation (Lemma B.1, Online Appendix B) is given by

$$\int_0^1 \left\|\mathcal{B}_1(u)\right\|_{\mathcal{H}}^2 \mathrm{d}u + \int_0^1 \left\|\mathcal{B}_2(u)\right\|_{\mathcal{H}}^2 \mathrm{d}u \overset{d}{=} \sum_{\ell=1}^{2}\sum_{j=1}^{\infty} \lambda_{\ell,j} \sum_{m=1}^{\infty} \frac{Z_{\ell,j,m}^2}{\pi^2 m^2}, \tag{4.3}$$

where $\overset{d}{=}$ denotes equality in distribution, and the $Z_{\ell,j,m}$ are independent standard normal random variables. Thus, once the eigenvalues of the two long-run covariance operators have been obtained, the limiting distribution can be approximated by truncating the infinite sums in (4.3).

For feasible inference, the nuisance long-run covariance operators $\mathcal{C}_1$ and $\mathcal{C}_2$ must be estimated. Estimation of long-run covariance operators has been studied extensively in the functional time series literature, and existing procedures can be employed in the present setting; see, for example, Horváth et al. (2016), Rice and Shang (2017), and the references therein. Let $\widehat{\mathcal{C}}_{1,n}(k)$ and $\widehat{\mathcal{C}}_{2,n}(k)$ denote estimators of the long-run covariance operators associated with the two segments determined by the maintained date. For each $k \in K_n$, let $\widehat{c}_{1-\alpha}(k)$ denote the $(1-\alpha)$ quantile obtained from (4.3) after replacing the eigenvalue sequences of $\mathcal{C}_1$ and $\mathcal{C}_2$ by those of $\widehat{\mathcal{C}}_{1,n}(k)$ and $\widehat{\mathcal{C}}_{2,n}(k)$, respectively. Allowing the pre- and post-break long-run covariance operators to differ requires computing $\widehat{c}_{1-\alpha}(k)$ separately for each $k \in K_n$. For the coverage result, it is sufficient that the estimated critical value be consistent at the true break date. We define the fLBI-based confidence sets by inverting the sequence of tests:

$$\mathbb{C}_{n,1-\alpha} = \left\{k \in K_n : U_n(k) \leq \widehat{c}_{1-\alpha}(k)\right\}. \tag{4.4}$$

Corollary 1. *Suppose that Assumption 1 holds and that* $\mathrm{tr}(\mathcal{C}_1) + \mathrm{tr}(\mathcal{C}_2) > 0$. *If* $\left\|\widehat{\mathcal{C}}_{1,n}(k_0) - \mathcal{C}_1\right\|_1 + \left\|\widehat{\mathcal{C}}_{2,n}(k_0) - \mathcal{C}_2\right\|_1 \xrightarrow{p} 0$, *where* $\|\cdot\|_1$ *denotes the trace norm and*

$\xrightarrow{p}$ *denotes convergence in probability, then* $\lim_{n\to\infty} \mathbb{P}\big(k_0 \in \mathbb{C}_{n,1-\alpha}\big) = 1-\alpha$.

The proof of Corollary 1 is straightforward. The condition $\mathrm{tr}(\mathcal{C}_1) + \mathrm{tr}(\mathcal{C}_2) > 0$ ensures that the limiting distribution in (4.2) is nondegenerate and continuous. The result then follows from Theorem 2 and the assumed trace-norm consistency of the long-run covariance operator estimators. Importantly, the coverage result holds for every deterministic sequence $\delta \in \mathcal{H}$, including both fixed breaks and shrinking breaks.

The preceding consistency condition is sufficient for coverage, but does not control the behavior of the covariance estimators at incorrect maintained dates. This distinction can be important for the power of the fLBI tests and, consequently, for the informativeness of the inverted confidence set in (4.4). Estimating the long-run covariance operators from residuals constructed under an incorrect maintained date may leave the true break in the residuals, inflate the covariance estimates, and result in nonmonotonic power; see the discussions for finite dimensional models in Perron (2006, Section 4.2) and Yamamoto (2018). We shall return to this issue in Section 5.1.

We next characterize the statistic at an incorrect maintained date $k = \lfloor n\tau \rfloor$, where $\tau \neq \tau_0$, under a local break magnitude. Under an incorrect maintained date, exactly one of the two candidate segments contains the true break. After rescaling the candidate segment containing the true break to $[0, 1]$, let $\rho_0 \in (0, 1)$ denote the location of the true break relative to the left endpoint of that segment. Specifically, $\rho_0 = (\tau_0 - \tau)/(1-\tau)$ if $\tau < \tau_0$, whereas $\rho_0 = \tau_0/\tau$ if $\tau > \tau_0$. Let $u \in [0, 1]$ index time within the rescaled candidate segment. The true break occurs at $u = \rho_0$. Consequently, the subinterval $[0, \rho_0]$ corresponds to the pre-break regime, whereas $(\rho_0, 1]$ corresponds to the post-break regime. Let $\widetilde{\mathcal{W}}_1$ and $\widetilde{\mathcal{W}}_2$ be independent $\mathcal{H}$-valued Brownian motions with covariance operators $\mathcal{C}_1$ and $\mathcal{C}_2$, respectively. As in Assumption 1, define

$$\widetilde{\mathcal{W}}_{\rho_0}(u) = \begin{cases} \widetilde{\mathcal{W}}_1(u), & 0 \le u \le \rho_0, \\ \widetilde{\mathcal{W}}_1(\rho_0) + \widetilde{\mathcal{W}}_2(u) - \widetilde{\mathcal{W}}_2(\rho_0), & \rho_0 < u \le 1. \end{cases} \tag{4.5}$$

Let $\widetilde{\mathcal{B}}_{\rho_0}(u) = \widetilde{\mathcal{W}}_{\rho_0}(u) - u\widetilde{\mathcal{W}}_{\rho_0}(1)$ denote the corresponding tied-down process. The increments of $\widetilde{\mathcal{W}}_{\rho_0}$ over $[0, \rho_0]$ are governed by $\mathcal{C}_1$, whereas those over $(\rho_0, 1]$ are governed by $\mathcal{C}_2$. Because the candidate segment containing the true break is demeaned according to (2.3) using a single segment mean, the local mean change is not completely removed. Its remaining contribution accumulates in the residual partial sum process, producing a deterministic component whose dependence on the rescaled time index u, up to sign and scale, is given by

$$g_{\rho_0}(u) = u\rho_0 - (u \wedge \rho_0), \qquad u \in [0,1]. \tag{4.6}$$

The following theorem establishes the local power result at an incorrect maintained break date.

Theorem 3 (Local break asymptotics when $\tau \neq \tau_0$). *Suppose that Assumption 1 holds and that $\delta = \delta_n = d/\sqrt{n}$, where $d \in \mathcal{H}$. (i) If $\tau < \tau_0$, then*

$$U_n(k) \xrightarrow{d} \int_0^1 \left\| \mathcal{B}_1^\circ(u) \right\|_{\mathcal{H}}^2 \mathrm{d}u + \int_0^1 \left\| \widetilde{\mathcal{B}}_{\rho_0}(u) + \sqrt{1-\tau}\, g_{\rho_0}(u)\, d \right\|_{\mathcal{H}}^2 \mathrm{d}u, \tag{4.7}$$

where $\mathcal{B}_1^\circ$ is a Brownian bridge with covariance operator $\mathcal{C}_1$ and is independent of $\widetilde{\mathcal{B}}_{\rho_0}$. (ii) If $\tau > \tau_0$, then

$$U_n(k) \xrightarrow{d} \int_0^1 \left\| \widetilde{\mathcal{B}}_{\rho_0}(u) + \sqrt{\tau}\, g_{\rho_0}(u)\, d \right\|_{\mathcal{H}}^2 \mathrm{d}u + \int_0^1 \left\| \mathcal{B}_2^\circ(u) \right\|_{\mathcal{H}}^2 \mathrm{d}u, \tag{4.8}$$

where $\mathcal{B}_2^\circ$ is a Brownian bridge with covariance operator $\mathcal{C}_2$ and is independent of $\widetilde{\mathcal{B}}_{\rho_0}$.

Theorem 3 characterizes the contribution of the local mean break to the limiting statistic and shows how its magnitude depends on the distance between the maintained and true break fractions. In particular, the drift $g_{\rho_0}(\cdot)$ appears only in the candidate segment containing the true break. Since $\int_0^1 \left[g_{\rho_0}(u)\right]^2 \mathrm{d}u = \rho_0^2(1-\rho_0)^2/3$, the effect of the local break on the limiting statistic vanishes as τ approaches τ_0 as expected.

5 Implementation and simulation study

We evaluate the proposed fLBI-based confidence sets in terms of their empirical coverage probability and empirical length. Before presenting the simulation study (Section 5.3), we first describe two practical aspects of implementing the fLBI-based confidence sets: the estimation of long-run covariance operators (Section 5.1) and the computation of critical values (Section 5.2).

5.1 Estimation of long-run covariance operators

All feasible procedures discussed below use the same kernel estimator of the long-run covariance operator and differ only in the residual sequences used to construct it. Let $\widehat{\varepsilon}_1, \ldots, \widehat{\varepsilon}_m$ denote a generic residual sequence used to estimate one of the two segment-specific long-run covariance operators, where m denotes the length of the sequence. For the residual sequences corresponding to the first and second maintained segments, respectively, $m = k$ and $m = n - k$. In either case, the residual sequence satisfies $\sum_{s=1}^{m} \widehat{\varepsilon}_s = 0$. For $\ell = 0, \ldots, m-1$, define the sample lag-ℓ autocovariance operator by $\widehat{\Gamma}_\ell = m^{-1} \sum_{s=\ell+1}^{m} \widehat{\varepsilon}_s \otimes \widehat{\varepsilon}_{s-\ell}$, where $(x \otimes y)f = \langle y,\, f \rangle_{\mathcal{H}}\, x$, and let $\widehat{\Gamma}_\ell^*$ denote the adjoint of $\widehat{\Gamma}_\ell$. For a kernel function K and a bandwidth $h_m > 0$, a generic long-run covariance operator $\mathcal{C}$ is estimated by

$$\widehat{\mathcal{C}} = \widehat{\Gamma}_0 + \sum_{\ell=1}^{m-1} K\left(\frac{\ell}{h_m + 1}\right)\left(\widehat{\Gamma}_\ell + \widehat{\Gamma}_\ell^*\right). \tag{5.1}$$

In the simulations below, we follow Horváth et al. (2014) and use the flat-top kernel advocated by Politis and Romano (1996, 1999), $K(x) = \mathbb{1}\{|x| < 0.1\} + (1.1 - |x|)\mathbb{1}\{0.1 \leq |x| < 1.1\}$, with bandwidth $h_m = \lfloor m^{1/2} \rfloor$. Alternative choices of the kernel and bandwidth are possible; see, for example, Horváth et al. (2016) for a detailed investigation of bandwidth selection for functional time series.

We consider two feasible versions of the fLBI-based confidence set, denoted by H_0–fLBI and H_1–fLBI which differ only in the residual sequences used to estimate the two segment-specific long-run covariance operators. The H_0–fLBI procedure constructs the residuals under the maintained null hypothesis $H_0(k)$. Specifically, for each maintained break date k, the observations are demeaned separately over $\{1, \ldots, k\}$ and $\{k+1, \ldots, n\}$, as in Eq. (2.3), and then (5.1) is applied separately to the two resulting residual sequences. When $k = k_0$, this residualization removes the mean break exactly. When $k \neq k_0$, however, one of the two maintained segments contains the true break, which is not removed by demeaning the entire segment using a single sample mean. The resulting residual sequence therefore retains a level shift. A kernel long-run covariance estimator may interpret this remaining deterministic component as persistent stochastic variation, thereby inflating the estimated long-run covariance operator. Since the critical value is computed from the eigenvalues of the estimated long-run covariance operators, such contamination may affect the power of the test and, consequently, the length of the confidence set. This issue is not specific to the LBI test proposed in this paper, but arises more generally in structural break inference. It is well documented in the finite dimensional time series literature, where consistent long-run covariance estimation under both the null and alternative hypotheses is crucial for inference; see Perron (2006) and Yamamoto (2018).

The H_1–fLBI procedure therefore adapts the modification proposed by Yamamoto (2018) by using auxiliary residuals constructed under a specification that permits an additional break at a date distinct from the maintained date. We first employ the break date estimator proposed by Aue et al. (2018, Eq. (2.3)):

$$\widehat{k} = \min \underset{j \in K_n}{\operatorname{argmax}} \left\| \frac{1}{\sqrt{n}} \left(\sum_{s=1}^{j} X_s - \frac{j}{n} \sum_{s=1}^{n} X_s \right) \right\|_{\mathcal{H}} . \tag{5.2}$$

For each maintained date k, if $\widehat{k} \neq k$, the observations are demeaned separately over the three index sets $\{1, \ldots, k \wedge \widehat{k}\}$, $\{(k \wedge \widehat{k}) + 1, \ldots, k \vee \widehat{k}\}$, and $\{(k \vee \widehat{k}) + 1, \ldots, n\}$. The resulting auxiliary residuals with indices in $\{1, \ldots, k\}$ and $\{k + 1, \ldots, n\}$ are then collected into residual sequences of lengths k and $n - k$, respectively. We apply (5.1) separately to these two sequences to estimate the long-run covariance operators for the first and second maintained segments. If $\widehat{k} = k$, the construction reduces to demeaning separately over the two maintained segments. In Section 5.3, we assess whether the H_1–fLBI modification reduces empirical confidence set length relative to H_0–fLBI while preserving the empirical coverage of H_0–fLBI.

5.2 Obtaining critical values

For both procedures in Section 5.1, the statistic $U_n(k)$ is computed from the maintained date residuals defined in (2.3). Thus, the two procedures differ only in the estimated long-run covariance operators used to approximate the limiting distribution in (4.3) and compute the corresponding critical values.

Inspired by Horváth et al. (2014, Section 4.1), we use a cumulative explained long-run variation rule to reduce the finite sample effect of inaccurately estimated tail eigenvalues. For each $k \in K_n$ and $\ell = 1, 2$, let $\widehat{\lambda}_{\ell,1}(k) \geq \widehat{\lambda}_{\ell,2}(k) \geq \cdots \geq 0$ denote the ordered eigenvalues of the estimated long-run covariance operator for the ℓth maintained segment. We retain the largest $q_\ell(k)$ eigenvalues such that

$$q_\ell(k) = \min\left\{q \geq 1 : \frac{\sum_{j=1}^{q} \widehat{\lambda}_{\ell,j}(k)}{\sum_{j=1}^{m_\ell(k)-1} \widehat{\lambda}_{\ell,j}(k)} \geq \mathcal{T}\right\}, \tag{5.3}$$

where $m_1(k) = k$ and $m_2(k) = n - k$. We set $\mathcal{T} = 0.99$ in our simulations. The rule is applied separately to the two segment-specific long-run covariance operator estimates. The results are insensitive to using $\mathcal{T} \in \{0.95, 0.975, 0.999\}$. By the representation in (4.3), the critical values used to construct the fLBI-based confidence sets are then obtained from 5,000 simulations of

$$\sum_{j=1}^{q_1(k)} \widehat{\lambda}_{1,j}(k) \sum_{r=1}^{1000} \frac{Z_{1,j,r}^2}{\pi^2 r^2} + \sum_{j=1}^{q_2(k)} \widehat{\lambda}_{2,j}(k) \sum_{r=1}^{1000} \frac{Z_{2,j,r}^2}{\pi^2 r^2}, \tag{5.4}$$

where the $Z_{\ell,j,r}$ are independent standard normal random variables. Given a nominal level α, the critical value is the empirical $(1-\alpha)$-quantile of the simulated values.

5.3 Simulation study

Following Aue et al. (2018), we set $\mathcal{H} = L^2([0,1])$ and generate the functional observations using the first $J = 21$ elements of the orthonormal Fourier basis: $\phi_1(u) = 1$, $\phi_{2j}(u) = \sqrt{2}\sin(2\pi j u)$, and $\phi_{2j+1}(u) = \sqrt{2}\cos(2\pi j u)$ for $j = 1, \ldots, 10$. For each error design described below, observations are generated from (2.1) with $\mu = 0$, $n \in \{200, 400\}$, $k_0 = \lfloor n\tau_0 \rfloor$, and $\tau_0 \in \{0.3, 0.5\}$. The two break fractions represent a central break and an earlier, asymmetric break. The break function is $\delta = \delta_n = d/\sqrt{n}$, where $d = \frac{\Delta}{\sqrt{5}} \sum_{j=1}^{5} \phi_j$ and $\Delta \in \{5, 6, 7, 8\}$. Hence, $n\|\delta_n\|_{\mathcal{H}}^2 = \Delta^2$. The break is distributed equally across the first five basis directions so that the results are not driven by a change along a single functional principal component.

To generate the error processes, let $\lambda_{1,j} = j^{-2}/\sum_{r=1}^{J} r^{-2}$ and $\lambda_{2,j} = j^{-1}/\sum_{r=1}^{J} r^{-1}$ for $j = 1, \ldots, J$. The corresponding covariance operators satisfy $\mathcal{C}_\ell \phi_j = \lambda_{\ell,j}\phi_j$ for $\ell = 1, 2$, with $\mathrm{tr}(\mathcal{C}_1) = \mathrm{tr}(\mathcal{C}_2) = 1$. Thus, the heterogeneous designs change the spectral shape of the long-run covariance operator while holding its trace fixed. We consider four error designs, denoted by IND–COM, IND–HET, FAR–COM, and FAR–HET, respectively. The first component indicates whether the errors are independent or follow a functional autoregressive (FAR) process, while COM and HET indicate whether the true long-run covariance operator is common or differs across the two regimes. For each $s \in \mathbb{Z}$ and $j = 1, \ldots, J$, let $\xi_{s,j}$ be independent $\mathcal{N}(0,1)$ random variables. For $\ell = 1, 2$, define $\eta_{\ell,s} = \sum_{j=1}^{J} \lambda_{\ell,j}^{1/2} \xi_{s,j}\, \phi_j$. (i) Under the IND–COM design $\varepsilon_s = \eta_{1,s}$. (ii) Under the IND–HET design $\varepsilon_s = \eta_{1,s}$ for $s \leq k_0$ and $\varepsilon_s = \eta_{2,s}$ for $s > k_0$. (iii) Under the FAR–COM design, $\varepsilon_s = \psi\, \varepsilon_{s-1} + (1-\psi)\, \eta_{1,s}$, where $\psi = 0.3$. The initial value ε_0 is drawn from the centered stationary Gaussian distribution with covariance operator $(1-\psi)(1+\psi)^{-1}\mathcal{C}_1$. (iv) Under the FAR–HET design, we initialize ε_0 as in FAR–COM and generate $\varepsilon_s = \psi\, \varepsilon_{s-1} + (1-\psi)\, \eta_{\ell,s}$, where $\ell = 1$ if $s \leq k_0$ and $\ell = 2$ if $s > k_0$. Thus, the FAR recursion continues across the break date, while the covariance operator of $\eta_{\ell,s}$ changes from $\mathcal{C}_1$ to $\mathcal{C}_2$. The normalization by $1 - \psi$ ensures that the stationary long-run covariance operator of ε_s in each FAR regime coincides with that in its independent counterpart. The paired designs therefore isolate the finite sample effect of serial dependence while holding fixed both the trace and the spectral shape of the long-run covariance operators.

We compare four confidence sets. The infeasible Oracle–fLBI confidence set uses the population long-run covariance operators and their full eigenvalue sequences. The H_0–fLBI and H_1–fLBI confidence sets use the long-run covariance operator estimates described above. Finally, we consider the confidence interval proposed by Aue et al. (2018, Section 3.3), which we refer to as the ARS interval. Additional simulations in which we replace the flat-top kernel with the Bartlett kernel and vary the bandwidth yield qualitatively similar results for the ARS interval. In the reported simulations, we therefore implement the ARS interval using the same flat-top kernel and bandwidth as the feasible LBI procedures, thereby maintaining a common long-run covariance estimation scheme across all feasible methods. Because the resulting ARS interval may extend beyond the candidate date set, we report its intersection with K_n for comparability with the fLBI-based confidence sets.

Unlike the fLBI-based confidence sets, the asymptotic justification of the ARS interval requires $\|\delta_n\|_{\mathcal{H}} \to 0$ and $n\|\delta_n\|_{\mathcal{H}}^2 \to \infty$. Its performance under the local break designs considered here therefore provides a benchmark for assessing the effects of weak mean breaks. The IND–COM and FAR–COM designs permit a direct comparison between the fLBI-based confidence sets and the ARS interval in this setting. The heterogeneous designs allow for a change in the long-run covariance operator, which is accommodated by the LBI procedures but lies outside the theoretical framework

Table 1: Empirical coverage and length under the *independent* error settings with $\tau_0 = 0.3$. The break magnitude satisfies $n\|\delta_n\|_{\mathcal{H}}^2 = \Delta^2$. The Oracle–fLBI, H_0–fLBI, and H_1–fLBI procedures use, respectively, population covariance operators, $H_0(k)$ residuals, and $H_1(k)$ residuals; ARS denotes the interval of Aue et al. (2018).

			Empirical coverage				Empirical length			
Setting	n	Δ	Oracle–fLBI	H_0–fLBI	H_1–fLBI	ARS	Oracle–fLBI	H_0–fLBI	H_1–fLBI	ARS
IND–COM	200	5	0.947	0.958	0.921	0.828	0.608	0.646	0.425	0.220
		6	0.945	0.955	0.927	0.860	0.419	0.496	0.303	0.160
		7	0.948	0.959	0.944	0.881	0.294	0.380	0.227	0.121
		8	0.949	0.961	0.948	0.910	0.230	0.312	0.186	0.092
	400	5	0.950	0.960	0.929	0.846	0.606	0.614	0.463	0.229
		6	0.948	0.954	0.934	0.873	0.415	0.452	0.323	0.166
		7	0.951	0.959	0.943	0.889	0.291	0.339	0.241	0.124
		8	0.954	0.957	0.944	0.907	0.229	0.274	0.195	0.096
IND–HET	200	5	0.953	0.968	0.944	0.825	0.495	0.509	0.311	0.154
		6	0.953	0.967	0.947	0.848	0.330	0.381	0.230	0.114
		7	0.949	0.963	0.950	0.878	0.242	0.308	0.184	0.087
		8	0.954	0.963	0.953	0.889	0.195	0.263	0.152	0.068
	400	5	0.949	0.956	0.939	0.835	0.483	0.459	0.332	0.159
		6	0.949	0.961	0.948	0.861	0.321	0.340	0.244	0.118
		7	0.952	0.962	0.953	0.884	0.238	0.270	0.192	0.090
		8	0.947	0.956	0.948	0.893	0.193	0.228	0.158	0.070

of Aue et al. (2018), and therefore provide a further stress test of the finite sample performance of the competing confidence sets.

We report the empirical coverage and length of the confidence sets. The empirical length is computed as the number of candidate dates contained in the confidence set divided by the total number of candidate dates, $|K_n| = n-3$, and then averaged across Monte Carlo replications. Each simulated functional observation is evaluated at 101 equally spaced points on $[0, 1]$. Finally, all methods have a nominal coverage probability of 95%, and the reported results are based on 5,000 Monte Carlo replications.

Tables 1 and 2 report the results for the independent and FAR error settings with $\tau_0 = 0.3$. In each table, the error covariance operator is either common across the two regimes (IND–COM, FAR–COM) or changes at the mean break (IND–HET, FAR–HET). The results for $\tau_0 = 0.5$ are qualitatively similar and are reported in Online Appendix C. Two main findings emerge. First, the fLBI-based confidence sets achieve good empirical coverage both with and without serial dependence. In particular, the coverage probabilities of Oracle–fLBI and H_0–fLBI remain close to the nominal level of 95% across the considered break magnitudes. For $\Delta = 5$, H_1–fLBI exhibits some undercoverage, but its average length is approximately 30% shorter than that of H_0–fLBI. This finding supports the use of residuals constructed under the

Table 2: Empirical coverage and length under the *serially dependent* error settings with $\tau_0 = 0.3$. The break magnitude satisfies $n\|\delta_n\|^2_{\mathcal{H}} = \Delta^2$. The Oracle–fLBI, H_0–fLBI, and H_1–fLBI procedures use, respectively, population covariance operators, $H_0(k)$ residuals, and $H_1(k)$ residuals; ARS denotes the interval of Aue et al. (2018).

			Empirical coverage				Empirical length			
Setting	n	Δ	Oracle–fLBI	H_0–fLBI	H_1–fLBI	ARS	Oracle–fLBI	H_0–fLBI	H_1–fLBI	ARS
FAR–COM	200	5	0.954	0.957	0.920	0.846	0.642	0.619	0.398	0.215
		6	0.953	0.952	0.924	0.874	0.447	0.471	0.285	0.156
		7	0.955	0.959	0.940	0.890	0.309	0.363	0.215	0.118
		8	0.956	0.960	0.945	0.920	0.239	0.301	0.178	0.089
	400	5	0.951	0.961	0.929	0.849	0.627	0.598	0.447	0.226
		6	0.951	0.953	0.933	0.878	0.431	0.438	0.313	0.164
		7	0.957	0.956	0.941	0.889	0.299	0.331	0.235	0.123
		8	0.956	0.955	0.945	0.909	0.233	0.269	0.191	0.095
FAR–HET	200	5	0.959	0.964	0.940	0.834	0.534	0.483	0.290	0.152
		6	0.961	0.963	0.944	0.862	0.353	0.362	0.217	0.112
		7	0.955	0.957	0.944	0.896	0.255	0.295	0.174	0.085
		8	0.962	0.963	0.954	0.905	0.204	0.255	0.146	0.066
	400	5	0.953	0.957	0.939	0.842	0.507	0.445	0.320	0.159
		6	0.952	0.961	0.946	0.871	0.334	0.331	0.237	0.118
		7	0.956	0.961	0.952	0.897	0.245	0.265	0.188	0.089
		8	0.949	0.957	0.949	0.911	0.197	0.224	0.155	0.069

alternative to improve the empirical power of the underlying LBI tests. At the same time, the performance of H_1–fLBI depends on the accuracy of the preliminary break date estimator and therefore improves as the break magnitude increases. Second, the ARS intervals are generally the shortest but exhibit severe undercoverage across all settings considered, although the extent of undercoverage decreases as Δ increases. Both findings are consistent with the theoretical distinction between the weak break framework underlying the LBI procedures and the strong break asymptotics underlying the ARS interval.

6 Conclusion

We developed confidence sets for the date of a weak mean break in Hilbert-valued data by inverting a sequence of tests of maintained break dates. The construction eliminated the unknown means and break magnitude under each maintained null, allowing valid inference without requiring consistent detection of the break date. The main theoretical challenge arose from the infinite dimensional nature of the problem. In finite dimensions, averaging over unknown break directions changes the covariance matrix while preserving absolute continuity between the null and alternative distributions. In infinite dimensions, however, the analogous covariance change may render the null and alternative distributions mutually singular, making the likelihood-based local power analysis needed to derive a locally best invariant (LBI) test impossible. We resolved this difficulty by restricting the covariance perturbation to an appropriate Hilbert–Schmidt class and adopting covariance-squared weighting. This restored absolute continuity and yielded a simple LBI statistic based directly on the Hilbert-space norm, avoiding covariance inversion and the associated tuning choices. We derived the null asymptotic properties for arbitrary deterministic break sequences and characterized local power under weak breaks. The simulations confirmed the expected coverage and length properties under both independent and serially correlated errors. Future research could investigate alternative weights over break dates and break directions to improve power against selected alternatives. The proposed approach could also be extended to trend breaks and structural changes in functional regression models.

References

Andrews, D. W. (1999). Estimation when a parameter is on a boundary. *Econometrica 67*(6), 1341–1383.

Andrews, D. W. K. and W. Ploberger (1994). Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica 62*(6), 1383–1414.

Aue, A., G. Rice, and O. Sönmez (2018). Detecting and dating structural breaks in functional data without dimension reduction. *Journal of the Royal Statistical Society Series B: Statistical Methodology 80*(3), 509–529.

Bai, L., Q. Hu, and W. Wu (2026). Inference for structural changes in nonstationary functional time series with partial measurement error. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, qkag072.

Berkes, I., L. Horváth, and G. Rice (2013). Weak invariance principles for sums of dependent random functions. *Stochastic Processes and their Applications 123*(2), 385–403.

Beutner, E., Y. Lin, and S. Smeekes (2023). GLS estimation and confidence sets for the date of a single break in models with trends. *Econometric Reviews 42*(2), 195–219.

Bogachev, V. I. (1998). *Gaussian Measures.* American Mathematical Society.

Boniece, B. C., L. Horváth, and L. Trapani (2025). On changepoint detection in functional data using empirical energy distance. *Journal of Econometrics 250*, 106023.

Chen, X. and H. White (1998). Central limit and functional central limit theorems for Hilbert-valued dependent heterogeneous arrays with applications. *Econometric Theory 14*(2), 260–284.

Da Prato, G. and J. Zabczyk (2014). *Stochastic Equations in Infinite Dimensions* (2 ed.). Encyclopedia of Mathematics and its Applications. Cambridge University Press.

Elliott, G. and U. K. Müller (2006). Efficient tests for general persistent time variation in regression coefficients. *Review of Economic Studies 73*, 907–940.

Elliott, G. and U. K. Müller (2007). Confidence sets for the date of a single break in linear time series regressions. *Journal of Econometrics 141*(2), 1196–1218.

Horváth, L., P. Kokoszka, and G. Rice (2014). Testing stationarity of functional time series. *Journal of Econometrics 179*(1), 66–82.

Horváth, L., G. Rice, and S. Whipple (2016). Adaptive bandwidth selection in the long run covariance estimator of functional time series. *Computational Statistics & Data Analysis 100*, 676–693.

King, M. L. (1987). Towards a theory of point optimal testing. *Econometric Reviews 6*, 169–218.

King, M. L. and P. X. Wu (1997). Locally optimal one-sided tests for multiparameter hypotheses. *Econometric Reviews 16*(2), 131–156.

Klebanov, I., B. Sprungk, and T. J. Sullivan (2021). The linear conditional expectation in Hilbert space. *Bernoulli 27*(4), 2267–2299.

Kurozumi, E. and Y. Yamamoto (2015). Confidence sets for the break date based on optimal tests. *Econometrics Journal 18*, 412–435.

Lehmann, E. L. and J. P. Romano (2005). *Testing Statistical Hypotheses* (3 ed.). New York: Springer.

Li, D., R. Li, and H. L. Shang (2024). Detection and estimation of structural breaks in high-dimensional functional time series. *Annals of Statistics 52*(4), 1716–1740.

Perron, P. (2006). Dealing with structural breaks. *Palgrave Handbook of Econometrics 1*(2), 278–352.

Politis, D. N. and J. P. Romano (1996). On flat-top kernel spectral density estimators for homogeneous random fields. *Journal of Statistical Planning and Inference 51*(1), 41–53.

Politis, D. N. and J. P. Romano (1999). Multivariate density estimation with general flat-top kernels of infinite order. *Journal of Multivariate Analysis 68*(1), 1–25.

Rice, G. and H. L. Shang (2017). A plug-in bandwidth selection procedure for long-run covariance estimation with stationary functional time series. *Journal of Time Series Analysis 38*, 591–609.

Yamamoto, Y. (2018). A modified confidence set for the structural break date in linear regression models. *Econometric Reviews 37*, 974–999.

Online Appendix to:

Confidence Sets for the Date of a Weak Mean Break in Functional Data

Yicong Lin

Vrije Universiteit Amsterdam and *Tinbergen Institute*

Contents

A Main proofs

Proof of Theorem 1. We first establish in *Step 1* that, for each fixed alternative date $k_0 = t$ with $t \neq k$, the likelihood ratio based on the residual sequence $\{\widehat{\varepsilon}_{k,s},\ s = 1, \ldots, n\}$ defined in (2.3) depends on the residuals only through $S_{k,t}$, where $S_{k,t} = \sum_{s=1}^{t} \widehat{\varepsilon}_{k,s}$ if $t < k$ and $S_{k,t} = \sum_{s=k+1}^{t} \widehat{\varepsilon}_{k,s}$ if $t > k$; see also Section 3.1. We then establish the existence of the corresponding likelihood ratio and derive its explicit form in *Step 2*, and obtain the LBI statistic in *Step 3*. Recall that we use the shorthand $\sum_{t\neq k} = \sum_{t\in K_n, t\neq k}$ throughout.

Step 1 For each alternative date $t \neq k$, let $\mathcal{P}_{k,t,b^2}$ denote the probability measure on $\mathcal{H}^n$ obtained by averaging the conditional distribution of the full residual sequence $\{\widehat{\varepsilon}_{k,s},\ s = 1, \ldots, n\}$ with respect to the Gaussian weighting measure ν_{k,t,b^2}. Specifically, let $\mathcal{B}(\mathcal{H}^n)$ denote the Borel σ-algebra on the product Hilbert space $\mathcal{H}^n$. For every event $E \in \mathcal{B}(\mathcal{H}^n)$,

$$\mathcal{P}_{k,t,b^2}(E) = \int_{\mathcal{H}} \mathbb{P}\Big(\big(\widehat{\varepsilon}_{k,1}, \ldots, \widehat{\varepsilon}_{k,n}\big) \in E \,\Big|\, k_0 = t,\ \delta = u\Big)\, \mathrm{d}\nu_{k,t,b^2}(u). \tag{A.1}$$

Since $\nu_{k,t,0}$ is concentrated at $\delta = 0$, we have $\mathcal{P}_{k,t,0}(E) = \mathbb{P}\big(\big(\widehat{\varepsilon}_{k,1}, \ldots, \widehat{\varepsilon}_{k,n}\big) \in E \mid k_0 = t, \delta = 0\big)$. This probability does not depend on t, and the resulting null distribution is therefore denoted by $\mathcal{P}_{k,0}$.

Consider first $t < k$. Since $\sum_{s=1}^{k} \widehat{\varepsilon}_{k,s} = 0$ and $S_{k,t} = \sum_{s=1}^{t} \widehat{\varepsilon}_{k,s}$, we have $\sum_{s=t+1}^{k} \widehat{\varepsilon}_{k,s} = -S_{k,t}$. Hence, the residual averages over $\{1, \ldots, t\}$ and $\{t+1, \ldots, k\}$ are $S_{k,t}/t$ and $-S_{k,t}/(k-t)$, respectively. It follows that $\big(\widehat{\varepsilon}_{k,1}, \ldots, \widehat{\varepsilon}_{k,n}\big)$ is uniquely determined by $S_{k,t}$, the residuals centered separately within these two subsegments, and the residuals over $\{k+1, \ldots, n\}$. Specifically, given $S_{k,t}$, the residuals for $s \in \{1, \ldots, t\}$ are recovered by adding $S_{k,t}/t$ to the centered residuals $\widehat{\varepsilon}_{k,s} - S_{k,t}/t$, whereas those for $s \in \{t+1, \ldots, k\}$ are recovered by subtracting $S_{k,t}/(k-t)$ from the centered residuals $\widehat{\varepsilon}_{k,s} + S_{k,t}/(k-t)$.

Moreover, straightforward algebra yields the following explicit expressions for the (centered) residuals: (i) for $s = 1, \ldots, t$, $\widehat{\varepsilon}_{k,s} - S_{k,t}/t = \varepsilon_s - t^{-1} \sum_{j=1}^{t} \varepsilon_j$; (ii) for $s = t+1, \ldots, k$, $\widehat{\varepsilon}_{k,s} + S_{k,t}/(k-t) = \varepsilon_s - (k-t)^{-1} \sum_{j=t+1}^{k} \varepsilon_j$; and (iii) for $s = k+1, \ldots, n$, $\widehat{\varepsilon}_{k,s} = \varepsilon_s - (n-k)^{-1} \sum_{j=k+1}^{n} \varepsilon_j$. Thus, none of these components depends on the break function δ. Moreover, recall $q_{k,t}$ from (3.2). Note that $S_{k,t} = -q_{k,t}\delta + (k-t)k^{-1} \sum_{j=1}^{t} \varepsilon_j - (t/k) \sum_{j=t+1}^{k} \varepsilon_j$. For every $s \in \{1, \ldots, t\}$ and $\ell \in$

$\{t+1,\dots,k\}$,

$$\begin{aligned}\operatorname{cov}\left(\sum_{j=1}^{t}\varepsilon_j,\,\varepsilon_s-\frac{1}{t}\sum_{j=1}^{t}\varepsilon_j\right)&=\mathcal{C}_\varepsilon-\mathcal{C}_\varepsilon=0,\\ \operatorname{cov}\left(\sum_{j=t+1}^{k}\varepsilon_j,\,\varepsilon_\ell-\frac{1}{k-t}\sum_{j=t+1}^{k}\varepsilon_j\right)&=\mathcal{C}_\varepsilon-\mathcal{C}_\varepsilon=0.\end{aligned}\tag{A.2}$$

Since these random elements are jointly Gaussian, each subsegment sum is therefore independent of the corresponding centered innovations. Independence of the innovations across disjoint time intervals, together with independence between δ and the innovations, therefore implies that $S_{k,t}$ and

$$\left(\left\{\widehat{\varepsilon}_{k,s}-\frac{S_{k,t}}{t}\right\}_{s=1}^{t},\left\{\widehat{\varepsilon}_{k,s}+\frac{S_{k,t}}{k-t}\right\}_{s=t+1}^{k},\{\widehat{\varepsilon}_{k,s}\}_{s=k+1}^{n}\right)\tag{A.3}$$

are independent. As shown above, the joint distribution of the random vector in (A.3) depends only on the innovations and therefore does not depend on b^2. Consequently, its joint distribution is the same under $\mathcal{P}_{k,t,b^2}$ and $\mathcal{P}_{k,0}$ for every $b^2\geq 0$.

If $t>k$, the same argument applies after partitioning the second maintained segment into $\{k+1,\dots,t\}$ and $\{t+1,\dots,n\}$. Hence, in either case, the residual sequence $\{\widehat{\varepsilon}_{k,s},\ s=1,\dots,n\}$ admits a one-to-one representation in terms of $S_{k,t}$ and the remaining components, which are independent of $S_{k,t}$ and have the same joint distribution under $\mathcal{P}_{k,t,b^2}$ and $\mathcal{P}_{k,0}$.

Let $\mathcal{P}^S_{k,t,b^2}$ and $\mathcal{P}^S_{k,t,0}$ denote the distributions of $S_{k,t}$ under $\mathcal{P}_{k,t,b^2}$ and $\mathcal{P}_{k,0}$, respectively. Under the one-to-one measurable representation established above, there exists a probability measure $\mathcal{M}_{k,t}$, not depending on b^2, such that the joint distributions of $S_{k,t}$ and the remaining components under $\mathcal{P}_{k,t,b^2}$ and $\mathcal{P}_{k,0}$ are, respectively, $\mathcal{P}^S_{k,t,b^2}\otimes\mathcal{M}_{k,t}$ and $\mathcal{P}^S_{k,t,0}\otimes\mathcal{M}_{k,t}$, where $\otimes$ denotes the product of probability measures. Consequently, whenever $\mathcal{P}^S_{k,t,b^2}\ll\mathcal{P}^S_{k,t,0}$, i.e., $\mathcal{P}^S_{k,t,b^2}$ is absolutely continuous with respect to $\mathcal{P}^S_{k,t,0}$, then the likelihood ratio of $\mathcal{P}_{k,t,b^2}$ with respect to $\mathcal{P}_{k,0}$, evaluated at the full residual sequence $\{\widehat{\varepsilon}_{k,s},\ s=1,\dots,n\}$, satisfies

$$\frac{\mathrm{d}\mathcal{P}_{k,t,b^2}}{\mathrm{d}\mathcal{P}_{k,0}}\left(\widehat{\varepsilon}_{k,1},\dots,\widehat{\varepsilon}_{k,n}\right)=\frac{\mathrm{d}\mathcal{P}^S_{k,t,b^2}}{\mathrm{d}\mathcal{P}^S_{k,t,0}}\left(S_{k,t}\right),\qquad \mathcal{P}_{k,0}\text{-a.s.}\tag{A.4}$$

Hence, the likelihood ratio based on the complete residual sequence is a function of the residuals only through $S_{k,t}$.

Step 2 Under $\mathcal{P}_{k,t,b^2}$, $S_{k,t} \sim \mathcal{N}_{\mathcal{H}}\big(0, q_{k,t}\,\mathcal{C}_\varepsilon + b^2 a_{k,t}\, q_{k,t}^2\,\mathcal{C}_\delta\big)$, whereas, under $\mathcal{P}_{k,0}$, $S_{k,t} \sim \mathcal{N}_{\mathcal{H}}\big(0, q_{k,t}\,\mathcal{C}_\varepsilon\big)$. It follows from (A.4) that, for every $t \neq k$, the likelihood ratio of $\mathcal{P}_{k,t,b^2}$ with respect to $\mathcal{P}_{k,0}$ is equal to the likelihood ratio of $\mathcal{P}^S_{k,t,b^2}$ with respect to $\mathcal{P}^S_{k,t,0}$, evaluated at $S_{k,t}$. Equivalently, applying the one-to-one transformation $x \mapsto q_{k,t}^{-1/2}x$, this likelihood ratio is equal to the likelihood ratio between the induced distributions of $q_{k,t}^{-1/2}S_{k,t}$, evaluated at $q_{k,t}^{-1/2}S_{k,t}$. These induced distributions are, respectively, $\mathcal{N}_{\mathcal{H}}\big(0, \mathcal{C}_\varepsilon + \eta_{k,t}\mathcal{C}_\delta\big)$ and $\mathcal{N}_{\mathcal{H}}\big(0, \mathcal{C}_\varepsilon\big)$, where $\eta_{k,t} = b^2 a_{k,t} q_{k,t}$.

Recall from (3.3) and the discussion in Section 3.2 that we set $\mathcal{K} = \mathcal{C}_\varepsilon$, so that $\mathcal{C}_\delta = \mathcal{C}_\varepsilon^2$. Following the notation of Chapter 1.3.3 in Da Prato and Zabczyk (2002), let $Q = \mathcal{C}_\varepsilon$, $R = \mathcal{C}_\varepsilon + \eta_{k,t}\,\mathcal{C}_\varepsilon^2$, and $S = -\eta_{k,t}\,\mathcal{C}_\varepsilon$. Then $R = Q^{1/2}\big(\mathcal{I} - S\big)Q^{1/2}$. Since $\mathcal{C}_\varepsilon$ is a nonnegative trace-class operator, both Q and R are nonnegative trace-class operators, while S is a self-adjoint trace-class operator. Moreover, $\eta_{k,t} \geq 0$ and $\mathcal{C}_\varepsilon$ is nonnegative. Hence, $S < \mathcal{I}$, and the conditions of Proposition 1.3.11 in Da Prato and Zabczyk (2002) are satisfied, yielding the Radon–Nikodym derivative of $\mathcal{N}_{\mathcal{H}}(0, R)$ with respect to $\mathcal{N}_{\mathcal{H}}(0, Q)$.

Specifically, let (λ_j, ϕ_j) denote the eigenpairs satisfying $\mathcal{C}_\varepsilon\phi_j = \lambda_j\phi_j$, where $j \in \mathbb{Z}^+$. Note that $\big(\mathcal{I} + \eta_{k,t}\mathcal{C}_\varepsilon\big)^{-1}\phi_j = (1 + \eta_{k,t}\lambda_j)^{-1}\phi_j$, and moreover, if $\lambda_j = 0$, then $\langle S_{k,t},\, \phi_j\rangle_{\mathcal{H}} = 0$ $\mathcal{P}_{k,0}$-almost surely. Expressing the quadratic term in Eq. (1.3.12) of Da Prato and Zabczyk (2002, Proposition 1.3.11) in the eigenbasis of $\mathcal{C}_\varepsilon$, evaluating it at $q_{k,t}^{-1/2}S_{k,t}$, and using $\eta_{k,t} = b^2 a_{k,t} q_{k,t}$ together with (A.4) yields the following $\mathcal{P}_{k,0}$-almost-sure expression for the likelihood ratio of $\mathcal{P}_{k,t,b^2}$ with respect to $\mathcal{P}_{k,0}$, evaluated at $\big\{\widehat{\varepsilon}_{k,s},\ s = 1, \ldots, n\big\}$:

$$
\begin{aligned}
L_{k,t}(b^2) &= \frac{\mathrm{d}\mathcal{P}_{k,t,b^2}}{\mathrm{d}\mathcal{P}_{k,0}}\big(\widehat{\varepsilon}_{k,1}, \ldots, \widehat{\varepsilon}_{k,n}\big) \\
&= \big[\det{}_F\big(\mathcal{I} + \eta_{k,t}\mathcal{C}_\varepsilon\big)\big]^{-1/2} \exp\left(\frac{\eta_{k,t}}{2q_{k,t}} \sum_{\lambda_j > 0} \frac{\langle S_{k,t},\, \phi_j\rangle_{\mathcal{H}}^2}{1 + \eta_{k,t}\lambda_j}\right) \\
&= \big[\det{}_F\big(\mathcal{I} + \eta_{k,t}\mathcal{C}_\varepsilon\big)\big]^{-1/2} \exp\left(\frac{b^2 a_{k,t}}{2} \Big\langle S_{k,t},\, \big(\mathcal{I} + \eta_{k,t}\mathcal{C}_\varepsilon\big)^{-1} S_{k,t}\Big\rangle_{\mathcal{H}}\right),
\end{aligned}
\tag{A.5}
$$

where $\det_F$ denotes the Fredholm determinant for trace class operators (Simon, 2005, Chapter 3).

Step 3 By the invariance reduction established above, $\varphi_{n,k}$ can be represented as a measurable function of $\big(\widehat{\varepsilon}_{k,1}, \ldots, \widehat{\varepsilon}_{k,n}\big)$. We use the same notation $\varphi_{n,k}$ for this function. For any integrable function $g : \mathcal{H} \to \mathbb{R}$, let $\mathbb{E}_{\delta\sim\nu_{k,t,b^2}}\big(g(\delta)\big) = \int_{\mathcal{H}} g(\delta)\, \mathrm{d}\nu_{k,t,b^2}(\delta)$. By the definition of $P_{n,k}\big(\varphi_{n,k}; b^2\big)$ in (2.4), the definition of $\mathcal{P}_{k,t,b^2}$ in (A.1), the equivalence of $\mathcal{P}_{k,t,b^2}$ and $\mathcal{P}_{k,0}$ established above, the definition of $L_{k,t}(b^2)$ in (A.5) and the change-of-measure formula, we have

$$\begin{aligned}
P_{n,k}\big(\varphi_{n,k}; b^2\big) &= \sum_{t\neq k} w_{k,t}\, \mathbb{E}_{\delta\sim\nu_{k,t,b^2}}\bigg(\mathbb{E}_{t,\delta}\Big[\varphi_{n,k}\big(\widehat{\varepsilon}_{k,1},\ldots,\widehat{\varepsilon}_{k,n}\big)\Big]\bigg)\\
&= \sum_{t\neq k} w_{k,t} \int_{\mathcal{H}^n} \varphi_{n,k}\, \mathrm{d}\mathcal{P}_{k,t,b^2}\\
&= \sum_{t\neq k} w_{k,t} \int_{\mathcal{H}^n} \varphi_{n,k} L_{k,t}(b^2)\, \mathrm{d}\mathcal{P}_{k,0}\\
&= \mathbb{E}_{k,0}\bigg(\varphi_{n,k} \sum_{t\neq k} w_{k,t} L_{k,t}(b^2)\bigg), \qquad\qquad (A.6)
\end{aligned}$$

where $\mathbb{E}_{k,0}$ denotes expectation with respect to $\mathcal{P}_{k,0}$.

To continue, we shall apply the dominated convergence theorem to justify differentiation under the null expectation. To this end, for each fixed $t \neq k$, define $\Phi_{k,t}(u) = \exp\Big(\frac{u a_{k,t}}{2}\, \langle S_{k,t},\, (\mathcal{I} + u\, a_{k,t}\, q_{k,t}\, \mathcal{C}_\varepsilon)^{-1} S_{k,t}\rangle_{\mathcal{H}}\Big)$ and $\Xi_{k,t}(u) = \det_F\big(\mathcal{I} + u\, a_{k,t}\, q_{k,t}\, \mathcal{C}_\varepsilon\big)$. Then, $L_{k,t}(u) = \Xi_{k,t}^{-1/2}(u)\Phi_{k,t}(u)$, as shown in (A.5). Note that

$$\begin{aligned}
\frac{\partial\big[\Xi_{k,t}(u)\big]^{-1/2}}{\partial u} &= -\frac{a_{k,t}\, q_{k,t}}{2}\big[\Xi_{k,t}(u)\big]^{-1/2} \sum_{j=1}^{\infty} \frac{\lambda_j}{1 + u\, a_{k,t}\, q_{k,t}\, \lambda_j}\\
&= -\frac{a_{k,t}\, q_{k,t}}{2}\big[\Xi_{k,t}(u)\big]^{-1/2} \operatorname{tr}\Big(\mathcal{C}_\varepsilon\big(\mathcal{I} + u\, a_{k,t}\, q_{k,t}\, \mathcal{C}_\varepsilon\big)^{-1}\Big),\\
\frac{\partial\Phi_{k,t}(u)}{\partial u} &= \Phi_{k,t}(u)\Bigg(\frac{a_{k,t}}{2}\Big\langle S_{k,t},\, \big(\mathcal{I} + u\, a_{k,t}\, q_{k,t}\, \mathcal{C}_\varepsilon\big)^{-1} S_{k,t}\Big\rangle_{\mathcal{H}}\\
&\quad - \frac{u\, a_{k,t}^2\, q_{k,t}}{2}\Big\langle S_{k,t},\, \big(\mathcal{I} + u\, a_{k,t}\, q_{k,t}\, \mathcal{C}_\varepsilon\big)^{-1}\mathcal{C}_\varepsilon\big(\mathcal{I} + u\, a_{k,t}\, q_{k,t}\, \mathcal{C}_\varepsilon\big)^{-1} S_{k,t}\Big\rangle_{\mathcal{H}}\Bigg)\\
&= \frac{a_{k,t}}{2}\,\Phi_{k,t}(u)\Big\langle S_{k,t},\, \big(\mathcal{I} + u\, a_{k,t}\, q_{k,t}\, \mathcal{C}_\varepsilon\big)^{-2} S_{k,t}\Big\rangle_{\mathcal{H}},
\end{aligned}$$

where the final equality follows from the identity $\big(\mathcal{I} + u\, a_{k,t}\, q_{k,t}\, \mathcal{C}_\varepsilon\big)^{-1} - u\, a_{k,t}\, q_{k,t}\big(\mathcal{I} + u\, a_{k,t}\, q_{k,t}\, \mathcal{C}_\varepsilon\big)^{-1}\mathcal{C}_\varepsilon\big(\mathcal{I} + u\, a_{k,t}\, q_{k,t}\, \mathcal{C}_\varepsilon\big)^{-1} = \big(\mathcal{I} + u\, a_{k,t}\, q_{k,t}\, \mathcal{C}_\varepsilon\big)^{-2}$. Putting together, by (A.5), we obtain

$$\frac{\partial L_{k,t}(u)}{\partial u} = \frac{a_{k,t}}{2} L_{k,t}(u)\bigg(\Big\langle S_{k,t},\, \big(\mathcal{I} + u\, a_{k,t}\, q_{k,t}\, \mathcal{C}_\varepsilon\big)^{-2} S_{k,t}\Big\rangle_{\mathcal{H}} - q_{k,t} \operatorname{tr}\Big(\mathcal{C}_\varepsilon\big(\mathcal{I} + u\, a_{k,t}\, q_{k,t}\, \mathcal{C}_\varepsilon\big)^{-1}\Big)\bigg). \qquad (A.7)$$

Since $0 \leq \left(\mathcal{I} + u\, a_{k,t}\, q_{k,t}\, \mathcal{C}_\varepsilon\right)^{-1} \leq \mathcal{I}$, for some constant $b_U > 0$ there exists a finite constant $\zeta_{k,t} > 0$, independent of u and $S_{k,t}$, such that

$$\sup_{0 \leq u \leq b^2} \left| \frac{\partial L_{k,t}(u)}{\partial u} \right| \leq \zeta_{k,t} \left(1 + \|S_{k,t}\|_{\mathcal{H}}^2\right) \exp\left(\frac{b_U^2\, a_{k,t}}{2} \|S_{k,t}\|_{\mathcal{H}}^2 \right). \tag{A.8}$$

Fix n and choose b_U sufficiently small such that $b_U^2 a_{k,t} q_{k,t} \|\mathcal{C}_\varepsilon\|_{\mathrm{op}} < 1$ for every $t \in K_n$ and $t \neq k$, where $\|\mathcal{C}_\varepsilon\|_{\mathrm{op}} := \sup_{\|f\|_{\mathcal{H}}=1} \|\mathcal{C}_\varepsilon f\|_{\mathcal{H}}$ denotes the operator norm. Such a constant b_U exists because, for each fixed n, the set of alternative dates is finite and $\|\mathcal{C}_\varepsilon\|_{\mathrm{op}} \leq \mathrm{tr}(\mathcal{C}_\varepsilon) < \infty$ (Da Prato and Zabczyk, 2014, Proposition 2.16). The choice of b_U^2 allows one to choose $\varrho_{k,t}$ such that

$$\frac{b_U^2 a_{k,t}}{2} < \varrho_{k,t} < \frac{1}{2 q_{k,t} \|\mathcal{C}_\varepsilon\|_{\mathrm{op}}}.$$

Moreover, it is not hard to show that there exists some constant $M > 0$ such that $(1 + x) \exp\left(b_U^2 a_{k,t}\, x/2\right) \leq M \exp(\varrho_{k,t} x)$ for $x \geq 0$. Therefore, by (A.8), we obtain $\left|\partial L_{k,t}(u)/\partial u\right| \leq M \zeta_{k,t} \exp(\varrho_{k,t} \|S_{k,t}\|_{\mathcal{H}}^2)$. Since $S_{k,t}$ is Gaussian with covariance operator $q_{k,t}\, \mathcal{C}_\varepsilon$ under $\mathcal{P}_{k,0}$, by the spectral representation of $S_{k,t}$ and $\mathrm{tr}(\mathcal{C}_\varepsilon) < \infty$, it is not hard to show that $\mathbb{E}_{k,0}\Big(\exp\left(\varrho_{k,t}\, \|S_{k,t}\|_{\mathcal{H}}^2\right) \Big) < \infty$. Hence, by the mean value theorem (Rudin, 1976, Theorem 5.10), for every $0 < b^2 \leq b_U^2$,

$$\mathbb{E}_{k,0}\left(\left| \frac{L_{k,t}(b^2) - L_{k,t}(0)}{b^2} \right| \right) \leq \mathbb{E}_{k,0}\left(\sup_{0 \leq u \leq b^2} \left| \frac{\partial L_{k,t}(u)}{\partial u} \right| \right) < \infty.$$

Dominated convergence therefore permits differentiation under the null expectation.

Therefore, by (A.6), (A.7), and the identity $\mathbb{E}_{k,0}(\varphi_{n,k}) = \alpha$, which follows from the imposed null rejection probability α, we obtain

$$\frac{\partial}{\partial b^2} P_{n,k}\left(\varphi_{n,k}; b^2\right) \bigg|_{b^2=0+} = \frac{1}{2}\, \mathbb{E}_{k,0}\left(\varphi_{n,k} \mathscr{S}_k\right) - \frac{\alpha}{2}\, \mathrm{tr}(\mathcal{C}_\varepsilon) \sum_{t \neq k} w_{k,t}\, a_{k,t}\, q_{k,t}, \tag{A.9}$$

where $\mathscr{S}_k := \sum_{t \neq k} w_{k,t}\, a_{k,t}\, \|S_{k,t}\|_{\mathcal{H}}^2$. The second term in (A.9) does not depend on the test. Hence, among tests satisfying $\mathbb{E}_{k,0}\left(\varphi_{n,k}\right) = \alpha$, maximizing the local slope is equivalent to maximizing $\mathbb{E}_{k,0}\left(\varphi_{n,k} \mathscr{S}_k\right)$.

The following Neyman–Pearson-type argument shows that this criterion is maximized by rejecting for large values of $\mathscr{S}_k$. Specifically, let $\varphi_{n,k}^*$ denote the test that rejects $H_0(k)$ for large values of $\mathscr{S}_k$, with critical value c^* chosen so that its null rejection probability equals α. For any other test $\varphi_{n,k}$ with the same null rejection probability, the definition of $\varphi^*_{n,k}$ implies that, for every realization of the data, $\left(\varphi_{n,k} - \varphi^*_{n,k}\right)\left(\mathscr{S}_k - c^*\right) \le 0$. Consequently,

$$\begin{aligned}
&\mathbb{E}_{k,0}\left[\left(\varphi_{n,k} - \varphi^*_{n,k}\right)\mathscr{S}_k\right] \\
&\quad = \mathbb{E}_{k,0}\left[\left(\varphi_{n,k} - \varphi^*_{n,k}\right)\left(\mathscr{S}_k - c^*\right)\right] + c^*\mathbb{E}_{k,0}\left[\varphi_{n,k} - \varphi^*_{n,k}\right] \\
&\quad = \mathbb{E}_{k,0}\left[\left(\varphi_{n,k} - \varphi^*_{n,k}\right)\left(\mathscr{S}_k - c^*\right)\right] \le 0,
\end{aligned}$$

where the second equality follows because the two tests have the same null rejection probability. Thus, $\varphi^*_{n,k}$ maximizes $\mathbb{E}_{k,0}\left(\varphi_{n,k}\mathscr{S}_k\right)$ and hence the local slope. Under the equal date-weighting scheme in Section 3.1, $\mathscr{S}_k$ differs from the statistic defined in (3.4) only by a positive multiplicative constant. The LBI test therefore rejects $H_0(k)$ for large values of the statistic in (3.4). □

Proof of Theorem 2. At the true date, (2.1) and (2.3) imply

$$S_{k_0,t} = \begin{cases} \displaystyle\sum_{s=1}^{t} \varepsilon_s - \frac{t}{k_0}\sum_{s=1}^{k_0} \varepsilon_s, & t < k_0, \\ \displaystyle\sum_{s=k_0+1}^{t} \varepsilon_s - \frac{t-k_0}{n-k_0}\sum_{s=k_0+1}^{n} \varepsilon_s, & t > k_0. \end{cases} \tag{A.10}$$

Note that $S_{k_0,n} = 0$. We also set $S_{k_0,k_0} = 0$ and adopt the convention $S_{k_0,0} = 0$. For $u \in [0,1]$, we have

$$\frac{1}{\sqrt{k_0}} S_{k_0,\lfloor k_0 u\rfloor} = \sqrt{\frac{n}{k_0}}\left(\frac{1}{\sqrt{n}}\sum_{s=1}^{\lfloor k_0 u\rfloor} \varepsilon_s - \frac{\lfloor k_0 u\rfloor}{k_0}\frac{1}{\sqrt{n}}\sum_{s=1}^{k_0}\varepsilon_s\right),$$

and similarly,

$$\frac{1}{\sqrt{n-k_0}} S_{k_0,\,k_0+\lfloor (n-k_0)u\rfloor} = \sqrt{\frac{n}{n-k_0}}\left(\frac{1}{\sqrt{n}}\sum_{s=k_0+1}^{k_0+\lfloor (n-k_0)u\rfloor} \varepsilon_s - \frac{\lfloor (n-k_0)u\rfloor}{n-k_0}\frac{1}{\sqrt{n}}\sum_{s=k_0+1}^{n}\varepsilon_s\right).$$

Since $k_0/n \to \tau_0$ and $(n-k_0)/n \to 1-\tau_0$, Assumption 1 and the continuous mapping theorem yield

$$\left(\left\{\frac{1}{\sqrt{k_0}} S_{k_0,\lfloor k_0 u\rfloor}\right\}_{0\le u\le 1}, \left\{\frac{1}{\sqrt{n-k_0}} S_{k_0,\,k_0+\lfloor (n-k_0)u\rfloor}\right\}_{0\le u\le 1}\right)$$

$$\Rightarrow \left(\left\{ \frac{\mathcal{W}_1(\tau_0 u) - u\mathcal{W}_1(\tau_0)}{\sqrt{\tau_0}} \right\}_{0\le u\le 1}, \left\{ \frac{\mathcal{W}_2\big(\tau_0 + (1-\tau_0)u\big) - \mathcal{W}_2(\tau_0) - u\big\{\mathcal{W}_2(1) - \mathcal{W}_2(\tau_0)\big\}}{\sqrt{1-\tau_0}} \right\}_{0\le u\le 1} \right)$$

as $n \to \infty$. By the definition of an $\mathcal{H}$-valued Brownian motion (e.g., Da Prato and Zabczyk, 2014, Definition 4.2), the process $\left\{ \frac{1}{\sqrt{\tau_0}} \mathcal{W}_1(\tau_0 u) \right\}_{0\le u\le 1}$ is an $\mathcal{H}$-valued Brownian motion on $[0,1]$ with covariance operator $\mathcal{C}_1$. Likewise, $\left\{ \frac{\mathcal{W}_2\big(\tau_0+(1-\tau_0)u\big)-\mathcal{W}_2(\tau_0)}{\sqrt{1-\tau_0}} \right\}_{0\le u\le 1}$ is an $\mathcal{H}$-valued Brownian motion with covariance operator $\mathcal{C}_2$. These two Brownian motions are independent because $\mathcal{W}_1$ and $\mathcal{W}_2$ are independent. Consequently, the limiting pair above has the same distribution as $(\mathcal{B}_1, \mathcal{B}_2)$ as defined in Theorem 2, and hence

$$\left(\left\{ \frac{1}{\sqrt{k_0}} S_{k_0,\lfloor k_0 u\rfloor} \right\}_{0\le u\le 1}, \left\{ \frac{1}{\sqrt{n-k_0}} S_{k_0,\, k_0+\lfloor (n-k_0)u\rfloor} \right\}_{0\le u\le 1} \right) \Rightarrow (\mathcal{B}_1, \mathcal{B}_2). \tag{A.11}$$

The definitions of the two residual partial sums give the exact identities

$$\int_0^1 \left\| \frac{1}{\sqrt{k_0}} S_{k_0,\lfloor k_0 u\rfloor} \right\|_{\mathcal{H}}^2 \mathrm{d}u = \frac{1}{k_0^2} \sum_{t=1}^{k_0-1} \left\| S_{k_0,t} \right\|_{\mathcal{H}}^2,$$
$$\int_0^1 \left\| \frac{1}{\sqrt{n-k_0}} S_{k_0,\, k_0+\lfloor (n-k_0)u\rfloor} \right\|_{\mathcal{H}}^2 \mathrm{d}u = \frac{1}{(n-k_0)^2} \sum_{t=k_0+1}^{n-1} \left\| S_{k_0,t} \right\|_{\mathcal{H}}^2.$$

Then, it follows from (3.4) that

$$U_n(k_0) = \int_0^1 \left\| \frac{1}{\sqrt{k_0}} S_{k_0,\lfloor k_0 u\rfloor} \right\|_{\mathcal{H}}^2 \mathrm{d}u + \int_0^1 \left\| \frac{1}{\sqrt{n-k_0}} S_{k_0,\, k_0+\lfloor (n-k_0)u\rfloor} \right\|_{\mathcal{H}}^2 \mathrm{d}u + o_{\mathbb{P}}(1). \tag{A.12}$$

We shall show that the map

$$(f_1, f_2) \mapsto \int_0^1 \|f_1(u)\|_{\mathcal{H}}^2 \,\mathrm{d}u + \int_0^1 \|f_2(u)\|_{\mathcal{H}}^2 \,\mathrm{d}u \tag{A.13}$$

is continuous at every pair of continuous paths (f_1, f_2). Since $\mathcal{B}_1$ and $\mathcal{B}_2$ have continuous sample paths almost surely, this result, together with an application of the continuous mapping theorem (CMT) to (A.12), immediately yields (4.2).

It remains to verify the required continuity of the map in (A.13). Let $(f_{1,n}, f_{2,n})$ converge to (f_1, f_2) in the product Skorokhod topology, where f_1 and f_2 are continuous. By Ethier and Kurtz (2005, Lemma 10.1), applied in their notation with $E = \mathcal{H}$ and

$r(x, y) = \|x - y\|_{\mathcal{H}}$, Skorokhod convergence to a continuous limit implies uniform convergence. Hence, for $\ell = 1, 2$,

$$\left| \int_0^1 \|f_{\ell,n}(u)\|_{\mathcal{H}}^2 \,\mathrm{d}u - \int_0^1 \|f_\ell(u)\|_{\mathcal{H}}^2 \,\mathrm{d}u \right| \\ \leq \sup_{0 \leq u \leq 1} \|f_{\ell,n}(u) - f_\ell(u)\|_{\mathcal{H}} \left\{ \sup_{0 \leq u \leq 1} \|f_{\ell,n}(u)\|_{\mathcal{H}} + \sup_{0 \leq u \leq 1} \|f_\ell(u)\|_{\mathcal{H}} \right\} \to 0.$$

Therefore, the map in (A.13) is continuous at every pair of continuous paths. This completes the proof. □

Proof of Theorem 3. Similar to (A.12), the test statistic admits the representation:

$$U_n(k) = \int_0^1 \left\| \frac{1}{\sqrt{k}} S_{k,\lfloor ku \rfloor} \right\|_{\mathcal{H}}^2 \mathrm{d}u + \int_0^1 \left\| \frac{1}{\sqrt{n-k}} S_{k,\, k+\lfloor (n-k)u \rfloor} \right\|_{\mathcal{H}}^2 \mathrm{d}u + o_{\mathbb{P}}(1). \tag{A.14}$$

It therefore suffices to derive the joint weak limit of the two normalized residual partial sum processes.

Suppose first that $\tau < \tau_0$. Then $k = \lfloor n\tau \rfloor < \lfloor n\tau_0 \rfloor = k_0$ for all sufficiently large n, and (2.3) gives

$$S_{k,t} = \begin{cases} \sum_{s=1}^{t} \varepsilon_s - \frac{t}{k} \sum_{s=1}^{k} \varepsilon_s, & t < k, \\ \sum_{s=k+1}^{t} \varepsilon_s - \frac{t-k}{n-k} \sum_{s=k+1}^{n} \varepsilon_s + \left[(t-k_0)_+ - \frac{t-k}{n-k}(n-k_0) \right] \delta_n, & t > k, \end{cases} \tag{A.15}$$

where $(x)_+ = \max\{x, 0\}$. First, the left candidate segment (i.e., $t < k$) lies entirely before the true break. Set $t = \lfloor ku \rfloor$. By Assumption 1, its normalized stochastic residual partial sum process converges to $\tau^{-1/2}\{\mathcal{W}_1(\tau u) - u\mathcal{W}_1(\tau)\}$, which has the same distribution as $\mathcal{B}_1^\circ$. Second, consider the right candidate segment (i.e., $t > k$) and set $t = k + \lfloor (n-k)u \rfloor$. The key normalized error partial sum is

$$\frac{1}{\sqrt{n-k}} \sum_{s=k+1}^{t} \varepsilon_s = \begin{cases} \frac{1}{\sqrt{n-k}} \sum_{s=k+1}^{t} \varepsilon_s, & t \leq k_0, \\ \frac{1}{\sqrt{n-k}} \left(\sum_{s=k+1}^{k_0} \varepsilon_s + \sum_{s=k_0+1}^{t} \varepsilon_s \right), & t > k_0. \end{cases} \tag{A.16}$$

By Assumption 1, after rescaling the right candidate segment to $[0, 1]$, this process converges to the process obtained by joining the increments of $\mathcal{W}_1$ over $[\tau, \tau_0]$ with those of $\mathcal{W}_2$ over $[\tau_0, 1]$. By (4.5), the limiting process has the same distribution as $\widetilde{\mathcal{W}}_{\rho_0}$. Subtracting $(t-k)/(n-k)\sum_{s=k+1}^{n}\varepsilon_s$ ties down the process at $u=1$, yielding $\widetilde{\mathcal{B}}_{\rho_0}$. Moreover, $\mathcal{B}_1^{\circ}$ and $\widetilde{\mathcal{B}}_{\rho_0}$ are independent because they depend on disjoint increments of $\mathcal{W}_1$ and on the independent process $\mathcal{W}_2$. Finally, the last term in the second line of (A.15) is the contribution of the local break. Under $t = k + \lfloor (n-k)u \rfloor$ and $\delta_n = d/\sqrt{n}$, and using $(k_0-k)/(n-k) \to \rho_0$, normalization by $(n-k)^{-1/2}$ yields uniform convergence over $u \in [0,1]$ to $\sqrt{1-\tau}\, g_{\rho_0}(u)d$. Consequently,

$$\left(\left\{ \frac{1}{\sqrt{k}} S_{k,\lfloor ku \rfloor} \right\}_{0\le u\le 1}, \left\{ \frac{1}{\sqrt{n-k}} S_{k,\,k+\lfloor (n-k)u \rfloor} \right\}_{0\le u\le 1} \right)$$
$$\Rightarrow \left(\left\{ \mathcal{B}_1^{\circ}(u) \right\}_{0\le u\le 1}, \left\{ \widetilde{\mathcal{B}}_{\rho_0}(u) + \sqrt{1-\tau}\, g_{\rho_0}(u)\, d \right\}_{0\le u\le 1} \right) \quad \text{(A.17)}$$

in the product Skorokhod space.

Suppose next that $\tau > \tau_0$. We then obtain

$$S_{k,t} = \begin{cases} \displaystyle\sum_{s=1}^{t} \varepsilon_s - \frac{t}{k}\sum_{s=1}^{k} \varepsilon_s + \left[(t-k_0)_+ - \frac{t}{k}(k-k_0) \right] \delta_n, & t < k, \\ \displaystyle\sum_{s=k+1}^{t} \varepsilon_s - \frac{t-k}{n-k} \sum_{s=k+1}^{n} \varepsilon_s, & t > k. \end{cases} \quad \text{(A.18)}$$

Similar arguments yield

$$\left(\left\{ \frac{1}{\sqrt{k}} S_{k,\lfloor ku \rfloor} \right\}_{0\le u\le 1}, \left\{ \frac{1}{\sqrt{n-k}} S_{k,\,k+\lfloor (n-k)u \rfloor} \right\}_{0\le u\le 1} \right)$$
$$\Rightarrow \left(\left\{ \widetilde{\mathcal{B}}_{\rho_0}(u) + \sqrt{\tau}\, g_{\rho_0}(u)\, d \right\}_{0\le u\le 1}, \left\{ \mathcal{B}_2^{\circ}(u) \right\}_{0\le u\le 1} \right) \quad \text{(A.19)}$$

in the product Skorokhod space.

All limiting processes have continuous sample paths. As shown below (A.13), the CMT applied to (A.14) therefore yields (4.7) and (4.8). □

B Auxiliary results

Lemma B.1. *Let $\mathcal{W}$ be an $\mathcal{H}$-valued Brownian motion with covariance operator $\mathcal{C}$, and define $\mathcal{B}(u) = \mathcal{W}(u) - u\mathcal{W}(1)$ for $u \in [0,1]$. Let $\{(\lambda_j, \phi_j)\}_{j\geq 1}$ denote the eigenpairs of $\mathcal{C}$. Then*

$$\int_0^1 \|\mathcal{B}(u)\|_{\mathcal{H}}^2 \,\mathrm{d}u \stackrel{d}{=} \sum_{j=1}^{\infty} \sum_{m=1}^{\infty} \frac{\lambda_j}{\pi^2 m^2} Z_{j,m}^2, \tag{B.1}$$

where the $Z_{j,m}$ are independent standard normal random variables. The double sum is almost surely finite.

Proof of Lemma B.1. By Da Prato and Zabczyk (2014, Proposition 4.3(ii)), $\mathcal{W}$ admits the series representation $\mathcal{W}(u) = \sum_{j=1}^{\infty} \lambda_j^{1/2} \beta_j(u) \phi_j$, where the β_j are independent standard Brownian motions. It follows that

$$\mathcal{B}(u) = \sum_{j=1}^{\infty} \sqrt{\lambda_j}\, \big[\beta_j(u) - u\beta_j(1)\big] \phi_j = \sum_{j=1}^{\infty} \sqrt{\lambda_j}\, B_j(u)\, \phi_j,$$

where $B_j(u) = \beta_j(u) - u\beta_j(1)$ are independent standard Brownian bridges. Parseval's identity, together with Tonelli's theorem for nonnegative functions, therefore yields

$$\int_0^1 \|\mathcal{B}(u)\|_{\mathcal{H}}^2 \,\mathrm{d}u = \sum_{j=1}^{\infty} \lambda_j \int_0^1 B_j(u)^2 \,\mathrm{d}u. \tag{B.2}$$

Note that $\int_0^1 B_j(u)^2 \,\mathrm{d}u = \sum_{m=1}^{\infty} Z_{j,m}^2/(\pi^2 m^2)$ almost surely; see Shorack and Wellner (1986, Chapter 5.3, Theorem 1, p. 215). Together with (B.2) above, this implies (B.1). Finally, Tonelli's theorem implies

$$\mathbb{E}\left(\sum_{j=1}^{\infty} \sum_{m=1}^{\infty} \frac{\lambda_j}{\pi^2 m^2} Z_{j,m}^2\right) = \sum_{j=1}^{\infty} \lambda_j \sum_{m=1}^{\infty} \frac{1}{\pi^2 m^2} = \frac{\mathrm{tr}(\mathcal{C})}{6} < \infty.$$

Since the series is nonnegative and has finite expectation, it converges to a finite value almost surely. □

C Additional simulation results

Tables C.3 and C.4 report the results for the independent and FAR error settings with $\tau_0 = 0.5$. The same findings as those discussed in Section 5.3 emerge.

Table C.3: Empirical coverage and length under the *independent* error settings with $\tau_0 = 0.5$. The break magnitude satisfies $n\|\delta_n\|_{\mathcal{H}}^2 = \Delta^2$. The Oracle–fLBI, H_0–fLBI, and H_1–fLBI procedures use, respectively, population covariance operators, $H_0(k)$ residuals, and $H_1(k)$ residuals; ARS denotes the interval of Aue et al. (2018).

			Empirical coverage				Empirical length			
Setting	n	Δ	Oracle–fLBI	H_0–fLBI	H_1–fLBI	ARS	Oracle–fLBI	H_0–fLBI	H_1–fLBI	ARS
IND–COM	200	5	0.943	0.960	0.933	0.849	0.534	0.567	0.388	0.135
		6	0.948	0.961	0.943	0.868	0.370	0.438	0.286	0.093
		7	0.949	0.964	0.947	0.880	0.281	0.354	0.224	0.069
		8	0.953	0.961	0.951	0.891	0.229	0.301	0.187	0.052
	400	5	0.949	0.958	0.936	0.855	0.527	0.539	0.415	0.142
		6	0.949	0.960	0.946	0.882	0.367	0.405	0.304	0.098
		7	0.949	0.956	0.945	0.883	0.278	0.322	0.237	0.072
		8	0.951	0.958	0.948	0.900	0.226	0.269	0.195	0.055
IND–HET	200	5	0.949	0.963	0.939	0.835	0.435	0.487	0.317	0.105
		6	0.953	0.965	0.945	0.871	0.307	0.376	0.237	0.074
		7	0.953	0.965	0.954	0.880	0.238	0.310	0.189	0.056
		8	0.956	0.963	0.953	0.896	0.196	0.268	0.158	0.043
	400	5	0.947	0.953	0.933	0.854	0.429	0.448	0.335	0.110
		6	0.954	0.958	0.943	0.873	0.301	0.342	0.251	0.078
		7	0.952	0.957	0.947	0.885	0.235	0.276	0.198	0.058
		8	0.951	0.959	0.951	0.900	0.194	0.234	0.164	0.045

Table C.4: Empirical coverage and length under the *serially dependent* error settings with $\tau_0 = 0.5$. The break magnitude satisfies $n\|\delta_n\|_{\mathcal{H}}^2 = \Delta^2$. The Oracle–fLBI, H_0–fLBI, and H_1–fLBI procedures use, respectively, population covariance operators, $H_0(k)$ residuals, and $H_1(k)$ residuals; ARS denotes the interval of Aue et al. (2018).

			Empirical coverage				Empirical length			
Setting	n	Δ	Oracle–fLBI	H_0–fLBI	H_1–fLBI	ARS	Oracle–fLBI	H_0–fLBI	H_1–fLBI	ARS
FAR–COM	200	5	0.951	0.960	0.932	0.854	0.564	0.541	0.367	0.132
		6	0.952	0.960	0.942	0.875	0.386	0.422	0.274	0.090
		7	0.952	0.964	0.948	0.895	0.289	0.343	0.217	0.067
		8	0.957	0.961	0.953	0.906	0.235	0.295	0.182	0.051
	400	5	0.952	0.959	0.939	0.859	0.544	0.525	0.404	0.140
		6	0.954	0.958	0.944	0.884	0.374	0.398	0.299	0.097
		7	0.952	0.955	0.944	0.889	0.283	0.318	0.234	0.072
		8	0.953	0.958	0.947	0.906	0.229	0.267	0.193	0.054
FAR–HET	200	5	0.954	0.964	0.941	0.843	0.463	0.465	0.302	0.103
		6	0.959	0.965	0.948	0.876	0.321	0.364	0.229	0.073
		7	0.958	0.962	0.951	0.893	0.247	0.303	0.184	0.055
		8	0.964	0.964	0.954	0.907	0.203	0.263	0.154	0.042
	400	5	0.953	0.952	0.933	0.858	0.444	0.439	0.328	0.109
		6	0.957	0.959	0.945	0.879	0.308	0.338	0.247	0.078
		7	0.953	0.958	0.947	0.893	0.239	0.273	0.196	0.058
		8	0.955	0.960	0.952	0.900	0.197	0.232	0.162	0.045

References in Online Appendix

Aue, A., G. Rice, and O. Sönmez (2018). Detecting and dating structural breaks in functional data without dimension reduction. *Journal of the Royal Statistical Society Series B: Statistical Methodology 80*(3), 509–529.

Da Prato, G. and J. Zabczyk (2002). *Second Order Partial Differential Equations in Hilbert Spaces*, Volume 293. Cambridge University Press.

Da Prato, G. and J. Zabczyk (2014). *Stochastic Equations in Infinite Dimensions* (2 ed.). Encyclopedia of Mathematics and its Applications. Cambridge University Press.

Ethier, S. N. and T. G. Kurtz (2005). *Markov Processes: Characterization and Convergence.* Hoboken, NJ: John Wiley & Sons.

Rudin, W. (1976). *Principles of Mathematical Analysis* (3rd ed.). New York: McGraw-Hill.

Shorack, G. R. and J. A. Wellner (1986). *Empirical Processes with Applications to Statistics.* Wiley Series in Probability and Mathematical Statistics. New York: John Wiley & Sons.

Simon, B. (2005). *Trace Ideals and Their Applications* (2 ed.), Volume 120 of *Mathematical Surveys and Monographs.* Providence, RI: American Mathematical Society.