Contributions to the theory of weak convergences in Hilbert spaces and its applications

SOKENDAI (The Graduate University for Advanced Studies)

Koji Tsukuda

Contents

Chapter 1

Introduction and preliminaries

1.1 Historical backgrounds and goals of this thesis

The weak convergence theory for random elements taking values in metric spaces such as functional spaces endowed with some metrics, has been developed over decades because it is of interest in itself as well as usefulness in statistical applications including goodness of fits tests and change point detection tests. Especially, a well-known goodness of fit test "Kolmogorov-Smirnov test" was validated by the so-called functional central limit theorem proven by M. D. Donsker following the heuristic approach stated by J. L. Doob, although the null distribution was originally derived by a direct calculation of characteristic functions much earlier. On the other hand, weak convergence theories in metric spaces were conclusively established by the landmark paper by Prohorov (1956). In that paper, a separable Hilbert space was one of concrete examples of metric spaces. After this work, the weak convergence theory in Banach spaces including a Skorokhod space D, which is the space of càdlàg functions endowed with the Skorokhod metric, and the space ℓ^∞, which is the space of bounded functions endowed with the supremum norm, has been developed much further and applied to many problems. See the books by Billingsley (1999) and by van der Vaart and Wellner (1996). On the other hand, the weak convergence theory in separable Hilbert spaces, especially L^2 spaces, has not been so developed and applied as compared to D and ℓ^∞ spaces. However, an elegant treatment of the Anderson-Darling statistic written in Example 1.8.6 of the book by van der Vaart and Wellner (1996) is of intense interest enough that it lets us foresee its extensive applicability. A possible direction is to consider dependent random variables. In the most previous works such as Khmaladze (1979) which considered empirical processes and Mason (1984) which considered quantile processes, independent random variables were argued. Though Oliveira and Suquet (1995, 1996, 1998) and Morel and Suquet (2002) treated empirical processes for some dependent cases, associated case and mixing case, they did not consider random fields which have the Anderson-Darling type weight function. Another possible direction is to expand the field of applications. Khmaladze (1979) and LaRiccia and Mason (1986) applied the weak convergence results of empirical processes and quantile processes to goodness of fit tests. Suquet and Viano (1998) applied the results of Oliveira and Suquet (1995) to change point tests. Some other statistical applications are in Oliveira and Suquet (1996). See Oliveira (2012) for some results about functional limit theorems of associated sequences in L^p spaces, including L^2 spaces, and the Skorokhod spaces.

Broadly speaking, this thesis has two goals. The one is to develop the limit theorem of random fields which have the weight function equivalent of the Anderson-Darling statistic used for testing change point hypotheses. We consider the model not only the independent random variables but stochastic processes. Especially, random fields of stochastic integrals taking values in L^2 spaces are considered. This part is based on Tsukuda (2015) and Tsukuda and Nishiyama (2014, 2015).

The other goal is to derive the new functional central limit theorems in L^2 spaces for the number of partitions by the Ewens partition and random mappings, which are important examples of random combinatorial structures, with a weight function. Because of this weight function, ℓ^∞ space may be not suitable as a framework to discuss the asymptotic behavior of random fields considered in this thesis. Because only one functional space discussed before is the Skorokhod space in this field, this thesis gives novel results. This part is based on Tsukuda (2014).

Although these two goals are concerned with different problems, there is a common treatment by the limit theorem in L^2 spaces at the base of the problems. Let us call it "L^2 space approach". We expect that more new limit theorems in an L^2 space can be acquired by this approach and it makes this approach valiant.

1.2 The organization of the thesis

The rest of this thesis is consisted as follows. In chapter 2, the weak convergence theory in Hilbert spaces is summarized and a new theorem with simple tightness criterion, which may be convenient especially for martingale random fields, is presented.

Part I is concerned with change point tests for some stochastic processes. Historical backgrounds and the key idea of our approach are explained in chapter 3. Chapter 4 contains some results on convergences of random elements taking values in L^2 spaces. Change point tests for continuous stochastic processes and discrete ones are argued in chapter 5 and 6, respectively. Chapter 7 includes comparisons of likelihood ratio methods and Z process methods and other problems for some independent cases.

Part II is devoted to L^2 functional limit theorems for two famous random logarithmic combinatorial structures: the Ewens sampling formula and random mappings. This part consists of three chapters. The goal of this part is formulated and historical backgrounds are explained in chapter 8. In chapter 9, a new functional CLT for the Ewens sampling formula is presented. The proof contains verifying a Poisson process approximation in $L^2([0,1], du)$ and establishing a functional CLT for a homogeneous Poisson process in an L^2 space. In chapter 10, a new functional CLT for random mappings, which is the other problem of this part, is presented. The argument for a Poisson process approximation which is different from chapter 8 is contained.

1.3 Preliminaries

Let us make some conventions. In this thesis, convergences as $T \to \infty$ and $n \to \infty$ are considered. The notations $\to^p$ and $\to^d$ denote convergence in probability and convergence in distribution, respectively. The notation l.i.m. means the limit in the second mean, where this "mean" is meant the expectation. "The sum" $\sum_{k=1}^{0} a_k$ is equal to 0 for any $\{a_\cdot\}$. The notation $1\{\cdot\}$ denotes the indicator function. The binary relations $a \wedge b$ and $a \vee b$ for $a, b \in \mathbb{R}$ mean $\min(a, b)$ and $\max(a, b)$, respectively. Let us denote the transpose of vector or matrix by superscript $\top$. The i-th element of vector x is denoted by $(x)_{(i)}$ and the finite dimensional Euclidean norm of vector x is denoted by $\|x\| = \sqrt{x^\top x}$. The (i, j) element of matrix A is denoted by

$(A)_{(i,j)}$ and the operator norm of matrix A is denoted by $\|A\|_{OP}$, that is,

$$\|A\|_{OP} = \sup_{x \in \mathbb{R}^d, \|x\|=1} \|Ax\| = \sup_{x \in \mathbb{R}^d, \|x\|>0} \frac{\|Ax\|}{\|x\|}.$$

Moreover, the Frobenius norm of matrix A which is defined by

$$\|A\| = \sqrt{\mathrm{tr}(A^\top A)} = \sqrt{\sum_i \sum_j |(A)_{(i,j)}|^2}$$

is denoted by $\|A\|$. Note that it holds that

$$\begin{aligned} \|A\|_{OP} &= \max_\sigma \sigma(A) \\ &\leq \sqrt{\sum_\sigma (\sigma(A))^2} = \|A\| \\ &\leq \sum_i \sum_j |(A)_{(i,j)}|, \end{aligned}$$

where $\sigma(A)$ denotes the singular value of the matrix A. The expectation and the variance, or the covariance matrix, of X which has density function $f(x, \theta)$ is denoted by $\mathbb{E}_\theta[X]$, $Var_\theta[X]$, respectively. In particular, for a random vector X, $\mathbb{E}_\theta[X]$ denotes the expectation of each element of X and $Var_\theta[X]$ denotes the covariance matrix of X. For a random matrix X, $\mathbb{E}_\theta[X]$ also denotes the expectation of each element of X. The covariance of random variables X_1 and X_2 is denoted by $Cov_\theta(X_1, X_2)$. For these three notations, when there is no risk of confusion, we omit subscript θ.

Let us introduce a functional space $L^2(S, \mathbb{R}^d, ds)$, or abbreviation form $L^2(S, ds)$, $L^2(ds)$ or $L^2(S)$, where S is a bounded subset of the Euclid space. Consider the inner product

$$\langle z_1, z_2 \rangle_{L^2(S)} = \int_S z_1(s)^\top z_2(s) ds,$$

where z_1 and z_2 are d dimensional vector-valued functions on S and ds is the Lebesgue measure. The functional space $L^2(S, \mathbb{R}^d, ds)$ is equivalence classes of square integrable real vector functions on a bounded set S, that is, the set of all measurable functions $z : S \to \mathbb{R}^d$ which satisfy $\|z\|^2_{L^2(S)} = \langle z, z \rangle_{L^2(S)} < \infty$. This space is a separable Hilbert space with respect to L^2

distance $d_2(z_1, z_2) = \|z_1 - z_2\|_{L^2(S)}$. Note that $\|\cdot\|_{L^2(S)}$ is different from $\|\cdot\|$, which is the Euclidean norm. In many cases, the inner product $\langle\cdot,\cdot\rangle_{L^2([0,1],du)}$ and the norm $\|\cdot\|_{L^2([0,1],du)}$ are denoted by $\langle\cdot,\cdot\rangle_{L^2}$ and $\|\cdot\|_{L^2}$ for simplicity. Let us denote a complete orthonormal system for $L^2([0,1], \mathbb{R}^d, du)$ space by $e_1, e_2, \ldots$. Note that, for example, it can be constructed as follows: If $\{e'_j;\ j \in J\}$ is a complete orthonormal system for $L^2([0,1], \mathbb{R}, du)$,

$$\left\{(e'_j, 0, \ldots, 0)^\top, (0, e'_j, 0, \ldots, 0)^\top, \ldots, (0, \ldots, 0, e'_j)^\top;\ j \in J\right\}$$

becomes a complete orthonormal system for $L^2([0,1], \mathbb{R}^d, du)$. The predictable quadratic variation process of martingales $t \rightsquigarrow M_t$ is denoted by $t \rightsquigarrow \langle M\rangle_t$ and note that it is different from the inner product $\langle\cdot,\cdot\rangle_{L^2(S)}$.

In the proofs, we sometimes omit integral region for simplicity if there is no risk of confusion.

Chapter 2

The weak convergence theory in separable Hilbert spaces

2.1 Known results

In, this section, let us introduce the result written in van der Vaart and Wellner (1996) with the assumption of the measurability. The conditions for the weak convergence of random elements taking values in separable Hilbert space was given by Prohorov (1956). Let $\mathbb{H}$ be a real separable Hilbert space with inner product $\langle \cdot, \cdot \rangle_{\mathbb{H}}$ and a complete orthonormal system $\{e_i\}$. A $\mathbb{H}$-valued random sequence X_n is said to be *asymptotically finite dimensional* if for any $\delta, \varepsilon > 0$, there exists a finite subset $\{e_i : i \in I\}$ of the complete orthonormal system such that

$$\limsup_{n\to\infty} \mathbb{P}\left(\sum_{j\notin I} \langle X_n, e_j \rangle_{\mathbb{H}}^2 > \delta\right) < \varepsilon.$$

This tightness criterion is established by Prohorov (1956) and the phrase "asymptotically finite dimensional" is apparently firstly used in van der Vaart and Wellner (1996).

The weak convergence in separable Hilbert spaces is characterized by the following theorem which clearly generalizes the Cramér-Wold theorem which characterizes weak convergences in finite dimensional Euclid spaces, see, for example, Billingsley (2012).

Theorem 2.1.1. *A sequence of random variables $X_n : \Omega_n \to \mathbb{H}$ converges in distribution to a tight random variable X if and only if it is asymptotically finite dimensional and the sequence $\langle X_n, h\rangle_{\mathbb{H}}$ converges in distribution to $\langle X, h\rangle_{\mathbb{H}}$ for every $h \in \mathbb{H}$.*

See section 1.8 of van der Vaart and Wellner (1996) for more details and the proof. It should be noted that the measurability of $\{X_\cdot\}$ is not assumed in their book. In this case, we have to argue the asymptotic measurability instead. By this theorem, we can prove the following CLT in a Hilbert space.

Example: the central limit theorem in a Hilbert space Consider the sequence $\{X_\cdot\}$ of i.i.d. random elements which take their values in a separable Hilbert space $\mathbb{H}$. Assume that $\mathbb{E}[\langle X_1, h\rangle_{\mathbb{H}}] = 0$ for any $h \in \mathbb{H}$ and $\mathbb{E}[\|X_1\|^2_{\mathbb{H}}] < \infty$. It holds that $n^{-1/2}\sum X_k$ converges to a Gaussian field G which satisfies $\langle G, h\rangle \sim N(0, \mathbb{E}[\langle X_1, h\rangle^2_{\mathbb{H}}])$ for any $h \in \mathbb{H}$.

Actually, it follows from the CLT that

$$\left\langle \frac{1}{\sqrt{n}}\sum_{k=1}^{n} X_k, h \right\rangle_{\mathbb{H}} = \frac{1}{\sqrt{n}}\sum_{k=1}^{n}\langle X_k, h\rangle_{\mathbb{H}} \to^d N(0, \mathbb{E}[\langle G, h\rangle^2_{\mathbb{H}}])$$

for any h. Moreover, letting $\{e_j : j \in J\}$ be a complete orthonormal system for $\mathbb{H}$, it holds that

$$\begin{aligned}
\mathbb{E}\left[\sum_{j>J_0}\left\langle \frac{1}{\sqrt{n}}\sum_{k=1}^{n} X_k, e_j \right\rangle^2_{\mathbb{H}}\right] &= \frac{1}{n}\sum_{j>J_0}\mathbb{E}\left[\left(\sum_{k=1}^{n}\langle X_k, e_j\rangle_{\mathbb{H}}\right)^2\right] \\
&= \frac{1}{n}\sum_{j>J_0}\mathbb{E}\left[\sum_{k=1}^{n}\langle X_k, e_j\rangle^2_{\mathbb{H}}\right] \\
&= \sum_{j>J_0}\mathbb{E}\left[\langle X_1, e_j\rangle^2_{\mathbb{H}}\right] \\
&= \mathbb{E}\left[\sum_{j>J_0}\langle X_1, e_j\rangle^2_{\mathbb{H}}\right]
\end{aligned}$$

converges to 0 as $J_0 \to \infty$. That is because, by the Bessel inequality, the integrand (of the expectation) is bounded above by $\|X_1\|^2_{\mathbb{H}}$ which is integrable.

By this CLT and the continuous mapping theorem, we can derive the asymptotic distribution of the Anderson-Darling test statistic.

Example: the Anderson-Darling test statistic for the goodness of fit test Consider the sequence $\{X_\cdot\}$ of real valued i.i.d. random variables with continuous distribution function F.

Let F_0 be a given continuous cumulative distribution function. We wish to test:

$\mathcal{H}_0$: $F = F_0$
$\mathcal{H}_1$: $F \neq F_0$

Consider the measure dF on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ Let us define the random field

$$t \rightsquigarrow Z_k(t) = \frac{1\{X_k \leq t\} - F(t)}{\sqrt{F(t)(1 - F(t))}}$$

which takes its value on $L^2(\mathbb{R}, dF)$. It holds that $\mathbb{E}\left[\langle Z_1, h\rangle_{L^2(\mathbb{R})}\right] = 0$ and that $\mathbb{E}\left[\|Z_1\|^2_{L^2(\mathbb{R})}\right] = 1$ by the Fubini theorem. Due to the central limit theorem in a separable Hilbert space,

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} Z_k(\cdot) = \frac{\sqrt{n}(\mathbb{F}_n(\cdot) - F(\cdot))}{\sqrt{F(\cdot)((1 - F(\cdot))}} \to^d G(\cdot)$$

in $L^2(\mathbb{R}, dF)$, where the empirical distribution function of $X_1, \ldots, X_n$ is denoted by $\mathbb{F}_n$ and $t \rightsquigarrow G(t)$ is a Gaussian field. It follows from the Fubini theorem that

$$
\begin{aligned}
& \mathbb{E}[\langle Z_1, h\rangle^2_{L^2(\mathbb{R})}] \\
= & \int_{\mathbb{R}}\int_{\mathbb{R}} \mathbb{E}[Z_1(s)Z_1(t)]h(s)h(t)dF(s)dF(t) \\
= & \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{F(s\wedge t)-F(s)F(t)}{\sqrt{F(s)((1-F(s))F(t)((1-F(t))}}h(s)h(t)dF(s)dF(t) \\
= & \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{F(s)\wedge F(t)-F(s)F(t)}{\sqrt{F(s)((1-F(s))F(t)((1-F(t))}}h(s)h(t)dF(s)dF(t) \\
= & \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{\mathbb{E}[B^\circ(F(s))B^\circ(F(t))]}{\sqrt{F(s)((1-F(s))F(t)((1-F(t))}}h(s)h(t)dF(s)dF(t) \\
= & \mathbb{E}\left[\left\langle \frac{B^\circ(F(\cdot))}{\sqrt{F(\cdot)(1-F(\cdot))}}, h\right\rangle^2_{L^2(\mathbb{R})}\right],
\end{aligned}
$$

where $u \rightsquigarrow B^\circ(u)$ denotes the (1 dimensional) standard Brownian bridge which is defined by $B^\circ(u) := B(u)-uB(1)$ for any u, where $u \rightsquigarrow B(u)$ denotes the standard Brownian motion. The first and the second moments of the standard Brownian bridge are given by $\mathbb{E}[B^\circ(u)] = 0$ and $\mathbb{E}[B^\circ(u)B^\circ(v)] = u\wedge v - uv$ for any u, v. Therefore, in this case,

$$
G(\cdot) = \frac{B^\circ(F(\cdot))}{\sqrt{F(\cdot)(1-F(\cdot))}}.
$$

By the convergence above and the continuous mapping theorem, it holds that

$$
\begin{aligned}
\int_{\mathbb{R}} \frac{n(\mathbb{F}_n(t)-F(t))^2}{F(t)(1-F(t))}dF(t) \quad &\to^d \quad \int_{\mathbb{R}} \left(\frac{B^\circ(F(t))}{\sqrt{F(t)(1-F(t))}}\right)^2 dF(t) \\
&= \quad \int_0^1 \left(\frac{B^\circ(u)}{\sqrt{u(1-u)}}\right)^2 du.
\end{aligned}
$$

Now, let us discuss the goodness of fit test. We can use

$$
\int_{\mathbb{R}} \frac{n(\mathbb{F}_n(t)-F_0(t))^2}{F_0(t)(1-F_0(t))}dF_0(t)
$$

as the test statistic. This statistic is called the Anderson-Darling (AD) statistic. We can construct the approximate rejection region by the asymptotic distributions under $\mathcal{H}_0$ derived above.

2.2 A new criterion

Sufficient conditions are given in the following proposition in order to check that a given sequence of random elements taking values in $\mathbb{H}$ is asymptotically finite dimensional.

Proposition 2.2.1. *A sequence of random variables $X_n : \Omega \to \mathbb{H}$ is asymptotically finite dimensional if there exists the random variable X such that*

$$\mathbb{E}[\|X_n\|_{\mathbb{H}}^2] \to \mathbb{E}[\|X\|_{\mathbb{H}}^2] < \infty \tag{2.2.1}$$

and

$$\mathbb{E}[\langle X_n, e_j\rangle_{\mathbb{H}}^2] \to \mathbb{E}[\langle X, e_j\rangle_{\mathbb{H}}^2], \quad \forall j \in J, \tag{2.2.2}$$

as $n \to \infty$, where $\{e_j : j \in J\}$ is a complete orthonormal system of $\mathbb{H}$.

Proof of the Proposition 2.2.1. It is enough to show that $\forall \epsilon > 0$, there exists a finite subset $\{e_i : i \in I\}$ of the complete orthonormal system such that

$$\limsup_{n\to\infty} \mathbb{E}\left[\sum_{j\notin I} \langle X_n, e_j\rangle_{\mathbb{H}}^2\right] < \epsilon$$

by the Markov inequality. The Parseval identity yields that

$$\|X\|_{\mathbb{H}}^2 = \sum_{j\in I} \langle X, e_j\rangle_{\mathbb{H}}^2 + \sum_{j\notin I} \langle X, e_j\rangle_{\mathbb{H}}^2,$$

so, it holds that for any $\epsilon > 0$ there exists a finite subset $I \subset J$ such that

$$\sum_{j\in I} \mathbb{E}\left[\langle X, e_j\rangle_{\mathbb{H}}^2\right] > \mathbb{E}\left[\|X\|_{\mathbb{H}}^2\right] - \epsilon.$$

Thence, we have from the assumptions that

$$\begin{aligned}\mathbb{E}\left[\sum_{j\notin I} \langle X_n, e_j\rangle_{\mathbb{H}}^2\right] &= \mathbb{E}\left[\|X_n\|_{\mathbb{H}}^2\right] - \mathbb{E}\left[\sum_{j\in I} \langle X_n, e_j\rangle_{\mathbb{H}}^2\right] \\ &\to \mathbb{E}\left[\|X\|_{\mathbb{H}}^2\right] - \mathbb{E}\left[\sum_{j\in I} \langle X, e_j\rangle_{\mathbb{H}}^2\right] < \epsilon\end{aligned}$$

for enough large finite set I. This completes the proof. □

Corollary 2.2.1. *If $\{X_n\}$ satisfies (2.2.1), (2.2.2) and the sequence $\langle X_n, h\rangle_{\mathbb{H}}$ converges in distribution to $\langle X, h\rangle_{\mathbb{H}}$ for every $h \in \mathbb{H}$, then $X_n \to^d X$ in $\mathbb{H}$.*

It will appear easier to check this set of sufficient conditions especially for some martingales and we shall use this corollary frequently in the thesis.

Part I

Applications to change point tests

Chapter 3

Introduction and the key idea

3.1 Historical backgrounds and rough explanations

As it is written in Chapter 1, the weak convergences of random elements taking values in Hilbert spaces was firstly established by Prohorov (1956). After Prohorov's paper, the weak convergence theory in Hilbert space progresses owing to some works; for example, see Parthasarathy (1967), Jakubowski (1980), Dedecker and Merlevède (2003), Merlevède (2003) and their references. On the other hand, the applications to statistical problems seem to be much less numerous than ones of other functional spaces like ℓ^∞ space, which is a space of bounded functions equipped with the supremum norm.

A successful applications of the weak convergence theory in L^2 space is to derive the asymptotic distribution of Anderson-Darling (AD) test statistic for goodness of fit test. While the asymptotic distribution of the test statistic was derived by a direct calculation of characteristic functions in the original paper by Anderson and Darling (1952), the book for modern theories of empirical processes by van der Vaart and Wellner (1996) contains an elegant proof based on the L^2 limit theory. See Section 1.8 of their book, and see also Khmaladze (1979) and LaRiccia and Mason (1986). Moreover, there is an application to change point problems for independent observations with the Anderson-Darling type weight function which needs more delicate arguments: see Tsukuda and Nishiyama (2014). The important point here is that we have to treat the weight function of the form $(u(1-u))^{-1/2}$ for $u \in (0,1)$ and it preclude the use of the weak convergence theory in ℓ^∞

space. In order to treat this weight function, most of the previous works of change point problems adopted the theory of Hungarian construction; we refer to Csörgő et al. (1986), Csörgő et al. (1993) and Csörgő and Horváth (1997) for the details. Indeed the approach by this celebrated theory gives us powerful tools in many situations, but it is not clear that we can apply it to the problems in this part, which are stated below , and we think it is natural to consider L^2 space as frameworks of weak convergences in order to treat this weight function. There are numerous studies which treat change point problems; see Csörgő and Horváth (1997), Brodsky and Darkhovsky (2000) and Chen and Gupta (2012) for reviews, and see Horváth and Rice (2014) for current progresses. Especially, there are several papers which treat change detections in stochastic process models; for example, diffusion processes with continuous observations: Lee et al. (2006), Mihalache (2012), Negri and Nishiyama (2012) and Dehling et al. (2014), diffusion processes with discrete observations: DeGregorio and Iacus (2008) and Song and Lee (2009), counting processes: Matthews et al. (1985) and Liang et al. (1990), AR(p) processes: Gombay (2008). Negri and Nishiyama (2014) adopts general framework called Z-process methods. However, none of them seem to have the same view as us. It is noted that although Suquet and Viano (1998) applied L^2 limit theorems to change point problems for some dependent cases, mixing and associated cases, they proved weak convergences of some statistics which do not have a weight function $(u(1-u))^{-1/2}$.

Let us make a rough explanation of our results. Consider a parametric model $\{P_\theta\}$ indexed by parameter θ. Let $t \rightsquigarrow X_t,\ t \in [0,\infty)$ be a semimartingale whose compensator is $\int a_s(\theta)ds$ and $\{\xi_k\}_{k=1,2,\ldots}$ be a martingale difference sequence under a probability measure P_θ for every θ. We shall propose a general approach based on the theory of weak convergence of random elements taking values in L^2 spaces: for continuous time stochastic processes, consider

$$(u,\theta) \rightsquigarrow \mathbb{Z}_T(u,\theta) = \frac{1}{\sqrt{T}} \int_0^T w_s^T(u) H_s(\theta)(dX_s - a_s(\theta)ds), \tag{3.1.1}$$

where H is a predictable process and

$$w_s^T : (0,1) \ni u \mapsto w_s^T(u) = \frac{1\{s \le Tu\} - u}{\sqrt{u(1-u)}}, \quad s \in [0,T];$$

and for discrete time stochastic processes, consider

$$(u,\theta) \rightsquigarrow \mathbb{Z}_n(u,\theta) = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} w_k^n(u) H_{k-1}(\theta) \xi_k(\theta), \tag{3.1.2}$$

where H_{k-1} is a measurable-$\mathcal{F}_{k-1}$ random element and

$$w_k^n : (0,1) \ni u \mapsto \begin{cases} 0, & u \in \left(0, \frac{1}{n}\right), \\ \frac{1\{k \le nu\} - [nu]/n}{\sqrt{[nu]/n(1-[nu]/n)}}, & u \in \left[\frac{1}{n}, 1\right), \ k = 1, \ldots, n. \end{cases}$$

Let us call (3.1.1) and (3.1.2) *pinned Z-process*, because if we assign the solution, or approximate solution, of estimating equations

$$\frac{1}{T} \int_0^T H_s(\theta)(dX_s - a_s(\theta)ds) = 0$$

and

$$\frac{1}{n} \sum_{k=1}^{n} H_{k-1}(\theta) \xi_k(\theta) = 0$$

for θ, (3.1.1) and (3.1.2) equal to the partial sums of estimating equations if we attach the suitable weight functions and rate constants. Pinned Z-processes converge to centered Gaussian fields as T, or n, tends to infinity. This idea basically comes from the the work of Horváth and Parzen (1994). They studied the asymptotic behavior of a Fisher score change process, it is the partial sum of the likelihood equation, for general independent observation cases under the null hypothesis. See also Negri and Nishiyama (2012), they refined the idea and applied it to a change detection of drift parameters in an ergodic diffusion process model. The proof for the limit theorem of Negri and Nishiyama (2012), especially the proof for the asymptotic tightness, is based on the tightness criterion for martingales taking values in ℓ^∞ spaces developed by Nishiyama (1999). As for the tightness criterion for martingales taking values in ℓ^∞ spaces, see also Nishiyama (2000) and references therein. Moreover, this approach is generalized to wide class of stochastic processes by Negri and Nishiyama (2014). However, as it is stated above, we cannot apply this kinds of weak convergence theorem to the current problems because the random fields $\mathbb{Z}_T$ and $\mathbb{Z}_n$ are not bounded in $(0,1)$ with respect to u. Hence, we regard random fields (3.1.1) and (3.1.2) as the elements in some space $L^2([0,1], du)$ and prove the limit theorems in this space.

3.2 A change point test for an independent random variables - the key idea -

Let us describe the problem for independent observations. Let $(\mathcal{X}, \mathcal{A}, \mu)$ be a measure space. Let $X_k,\ k = 1, \ldots, n$ be independent random variables taking values in $\mathcal{X}$, whose probability density functions with respect to the measure μ are $f(x; \theta_{(1)}), \ldots, f(x; \theta_{(n)})$, where $\theta \in \Theta \subset \mathbb{R}^d$ and Θ is a bounded open convex set. For this model, what we wish to test is that

$\mathcal{H}_0$: $\exists \theta_0 \in \Theta$ *such that* $\theta_{(k)} = \theta_0,\ \forall k = 1, \ldots, n$
$\mathcal{H}_1$: $\exists \theta_0, \theta_1 \in \Theta,\ \exists u_* \in (0,1)$ *such that* $\theta_{(k)} = \theta_0,\ \forall k = 1, \ldots, [nu_*]$
and that $\theta_{(k)} = \theta_1 \neq \theta_0,\ \forall k = [nu_*] + 1, \ldots, n$

The likelihood is given by

$$\prod_{k=1}^{n} f(X_k, \theta_{(k)}),$$

and the log likelihood is

$$\sum_{k=1}^{n} \log f(X_k, \theta_{(k)}) = \sum_{k=1}^{n} l_{\theta_{(k)}}(X_k),$$

where $l_\theta(x) = \log f(x, \theta)$. Consider the likelihood equation

$$\frac{1}{n} \sum_{k=1}^{n} \dot{l}_\theta(X_k) = 0.$$

The notation $\hat{\theta}_n$ denotes the solution, or an approximate solution in the sense that

$$\frac{1}{n} \sum_{k=1}^{n} \dot{l}_{\hat{\theta}_n}(X_k) = o_P(n^{-1/2}),$$

of the above equation.

Suppose the following conditions. Let us assume that l_θ is second order differentiable with respect to θ. For any $\theta_0 \in \Theta$, there exists a nonnegative measurable function K which satisfy

$$\int_{\mathcal{X}} K(x) f(x, \theta_0) \mu(dx) < \infty,$$

and

$$|\partial_i l_{\theta_1}(x) - \partial_i l_{\theta_2}(x)| \le K(x) \|\theta_1 - \theta_2\|, \ \forall \theta_1, \theta_2 \in N \tag{3.2.1}$$

$$|\partial_{ij} l_{\theta_1}(x) - \partial_{ij} l_{\theta_2}(x)| \le K(x) \|\theta_1 - \theta_2\|, \ \forall \theta_1, \theta_2 \in N \tag{3.2.2}$$

for all $i, j = 1, \ldots, d$, where N is neighborhood of θ_0. The matrix

$$I_\theta = \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}_{\theta_{(k)}}[\dot{l}_\theta(X_k) \dot{l}_\theta(X_k)^\top],$$

is assumed to be a positive definite matrix for all θ. Assume that for all $i, j = 1, \ldots, d$, for all $k = 1, \ldots, n$ and all $\theta \in \Theta$,

$$\mathbb{E}_{\theta_{(k)}} \left[(\partial_i \partial_j l_\theta(X_k))^2 \right] < \infty,$$
$$\mathbb{E}_{\theta_{(k)}} \left[(K(X_k))^2 \right] < \infty,$$

and there exist a $\delta > 0$ such that for all $i = 1, \ldots, d$,

$$\mathbb{E}_{\theta_{(k)}} \left[|\partial_i l_\theta(X_k)|^{2+\delta} \right] < \infty. \tag{3.2.3}$$

Let us assume

$$\inf_{\theta : \|\theta - \theta_0\| > \varepsilon} \left\| \int_{\mathcal{X}} \dot{l}_\theta(x) f(x, \theta_0) \mu(dx) \right\| > 0, \quad \forall \theta_0 \in \Theta, \ \forall \varepsilon > 0. \tag{3.2.4}$$

As for the estimator $\hat{\theta}_n$, it holds that the following properties: Under $\mathcal{H}_0$, it holds that $\sqrt{n}(\hat{\theta}_n - \theta_0) \to^d N(0, I_{\theta_0}^{-1})$. Under $\mathcal{H}_1$, it holds that $\hat{\theta}_n \to^p \theta_*$, where θ_* is a point in Θ such that

$$\theta_* \ne \theta_0, \theta_* \ne \theta_1, u_* \mathbb{E}_{\theta_0}[\dot{l}_{\theta_*}(X_1)] + (1 - u_*) \mathbb{E}_{\theta_1}[\dot{l}_{\theta_*}(X_1)] = 0.$$

Here, let us explain *Z-process methods*. Solutions to estimating equations

$$\Psi_n(\theta) = \frac{1}{n}\sum_{k=1}^{n}\psi(X_k,\theta) = 0$$

are sometimes called Z-estimators (see Chapter 3.3 of van der Vaart and Wellner (1996) and Chapter 5 of van der Vaart (1998)), where $\psi(x,\theta)$ is a finite dimensional vector. See $\psi(X_k,\theta)$ as $\dot{l}_\theta(X_k)$ in this section. We shall consider change point tests based on estimating equations. For the purpose of it, let us introduce *Z-process* $\Psi_n(u,\theta)$ and *pinned Z-process* $\Psi_n^\circ(u,\theta)$ as follows:

$$\begin{aligned}\Psi_n(u,\theta) &= \frac{1}{n}\sum_{k=1}^{[nu]}\psi(X_k,\theta),\\ \Psi_n^\circ(u,\theta) &= \frac{1}{n}\sum_{k=1}^{n}(1\{k\le nu\}-s_n(u))\psi(X_k,\theta),\end{aligned}$$

where

$$s_n(u) = \frac{[nu]}{n}, \quad u\in(0,1)$$

and $\sum_{k=1}^{0}$ is always zero hereafter. We call $\Psi_n(u,\theta_0)$ *Z-motion*, and $\Psi_n^\circ(u,\theta_0)$ *Z-bridge* because it holds that

$$\sqrt{n}\hat{I}_n^{-\frac{1}{2}}\Psi_n(\cdot,\theta_0) \to^d B_d(\cdot)$$

and

$$\sqrt{n}\hat{I}_n^{-\frac{1}{2}}\Psi_n^\circ(\cdot,\theta_0) \to^d B_d^\circ(\cdot),$$

where $B_d(\cdot)$ and $B_d^\circ(\cdot)$ are the d dimensional standard Brownian motion and the standard Brownian bridge, respectively and $\hat{I}_n$ is a consistent estimator of $Var[\psi(X_1,\theta_0)]$ under $\mathcal{H}_0$. For general θ, $\Psi_n(\cdot,\theta)$ is not equal to $\Psi_n^\circ(\cdot,\theta)$. However, the important point is that

$$\Psi_n(\cdot,\hat{\theta}_n) = \Psi_n^\circ(\cdot,\hat{\theta}_n) \approx \Psi_n^\circ(\cdot,\theta_0), \tag{3.2.5}$$

in some sense. Because of this relationship, we can use functions of $\Psi_n(\cdot,\hat{\theta}_n)$ as test statistics.

Now, let us propose the test statistic

$$AD_n = \sum_{j=1}^{n-1} \frac{n^2 \Phi_{n,j}(\hat{\theta}_n)^\top \hat{I}_n^{-1} \Phi_{n,j}(\hat{\theta}_n)}{j(n-j)},$$

where

$$\hat{I}_n = \frac{1}{n} \sum_{k=1}^{n} \dot{l}_{\hat{\theta}_n}(X_k) \dot{l}_{\hat{\theta}_n}(X_k)^\top, \tag{3.2.6}$$

and $(\Phi_{n,j}(\theta))_{j=1,\dots,n-1}$ is calculated by

$$\Phi_{n,j}(\theta) = \frac{1}{n^2} \left[(n-j) \sum_{k=1}^{j} \dot{l}_\theta(X_k) - j \sum_{k=j+1}^{n} \dot{l}_\theta(X_k) \right].$$

Remark 3.2.1. *In this case, under $\mathcal{H}_0$, the pinned Z-process is*

$$\begin{aligned} \Psi_n^\circ(u,\theta) &= \frac{1}{n} \left[(1-s_n(u)) \sum_{k=1}^{ns_n(u)} \dot{l}_\theta(X_k) - s_n(u) \sum_{k=ns_n(u)+1}^{n} \dot{l}_\theta(X_k) \right] \\ &= \frac{1}{n} \sum_{k=1}^{n} \left(1\{k \le nu\} - s_n(u) \right) \dot{l}_\theta(X_k) \end{aligned}$$

for $u \in (0,1)$ and a direct computation shows that

$$\begin{aligned} AD_n &= \int_0^1 \frac{n \Psi_n^\circ(u,\hat{\theta}_n)^\top \hat{I}_n^{-1} \Psi_n^\circ(u,\hat{\theta}_n)}{s_n(u)(1-s_n(u))} du \\ &= \left\| \frac{\sqrt{n} \hat{I}_n^{-1/2} \Psi_n^\circ(\cdot,\hat{\theta}_n)}{\sqrt{s_n(\cdot)(1-s_n(\cdot))}} \right\|_{L^2(S)}^2, \end{aligned} \tag{3.2.7}$$

where let us make a convention that if $u < 1/n$ then

$$\frac{1\{k \le nu\} - s_n(u)}{\sqrt{s_n(u)(1-s_n(u))}} = 0.$$

In the proof of the following theorem, we use this expression (3.2.7) for AD_n. Under $\mathcal{H}_1$, the corresponding term is $M_n(u)/\sqrt{n}$ where

$$M_n(u) = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \left(1\{k \le nu\} - s_n(u) \right) \left(\dot{l}_{\theta_*}(X_k) - \mathbb{E}_{\theta_{(k)}}[\dot{l}_{\theta_*}(X_k)] \right)$$

for $u \in (0,1)$.

Remark 3.2.2. *For $\hat{I}_n$, any consistent estimator for Fisher information matrix under $\mathcal{H}_0$ can be used. We can always construct it by (3.2.6), but in some cases we can construct more sensible estimators like*

$$\hat{I}_n = \frac{-1}{n} \sum_{k=1}^{n} \ddot{l}_{\hat{\theta}_n}(X_k).$$

For example, it becomes a constant for normal observations with known variance. We will use this $\hat{I}_n$ in Section 7.

As for this test, the following Theorem holds.

Theorem 3.2.1. *(i) Under $\mathcal{H}_0$, the asymptotic distribution of AD_n is*

$$\int_0^1 \left\| \frac{B_d^\circ(u)}{\sqrt{u(1-u)}} \right\|^2 du.$$

(ii) Under $\mathcal{H}_1$, the test is consistent.

In order to prove this theorem, let us prepare following lemmas.

Lemma 3.2.1. *(i) Under $\mathcal{H}_0$, it holds that $\hat{I}_n \to^p I_{\theta_0}$. (ii) Under $\mathcal{H}_1$, it holds that $\hat{I}_n \to^p I_{\theta_*}$.*

Lemma 3.2.2. *(i) Under $\mathcal{H}_0$, it holds that*

$$\mathbb{E}\left[\frac{n\Psi_n^\circ(u,\theta_0)^\top I_{\theta_0}^{-1} \Psi_n^\circ(u,\theta_0)}{s_n(u)(1-s_n(u))} \right] = d$$

for all $u \in (0,1)$ and for all $n \in \mathbb{N}$.
(ii) Under $\mathcal{H}_1$, it holds that

$$\begin{aligned} &\mathbb{E}_{\theta_{true}} \left[\frac{M_n(u)^\top I_{\theta_*}^{-1} M_n(u)}{s_n(u)(1-s_n(u))} \right] \\ \leq\ &\mathbb{E}_{\theta_0}\left[\dot{l}_{\theta_*}(X_1)^\top I_{\theta_*}^{-1} \dot{l}_{\theta_*}(X_1) \right] + \mathbb{E}_{\theta_1}\left[\dot{l}_{\theta_*}(X_1)^\top I_{\theta_*}^{-1} \dot{l}_{\theta_*}(X_1) \right], \end{aligned}$$

for all $u \in (0,1)$ and for all $n \in \mathbb{N}$, where $\mathbb{E}_{\theta_{true}}$ denotes integration with the true probability measure under $\mathcal{H}_1$.

Remark 3.2.3. *Lemma 3.2.2 implies that random elements*

$$\left\| \frac{\sqrt{n} I_{\theta_0}^{-1/2} \Psi_n^\circ(\cdot, \theta_0)}{\sqrt{s_n(\cdot)(1 - s_n(\cdot))}} \right\|_{L^2}^2$$

and

$$\left\| \frac{I_{\theta_*}^{-1/2} M_n(\cdot)}{\sqrt{s_n(\cdot)(1 - s_n(\cdot))}} \right\|_{L^2}^2,$$

are asymptotically tight in $\mathbb{R}$ *because, by the Fubini theorem, their expectations do not depend on* n *and they are finite under* $\mathcal{H}_0$ *and* $\mathcal{H}_1$*, respectively. Moreover, it holds that*

$$\left\| \frac{\sqrt{n} I_{\theta_0}^{-1/2} \Psi_n^\circ(\cdot, \theta_0)}{\sqrt{s_n(\cdot)(1 - s_n(\cdot))}} \right\|_{L^2}^2 < \infty \text{ and } \left\| \frac{I_{\theta_*}^{-1/2} M_n(\cdot)}{\sqrt{s_n(\cdot)(1 - s_n(\cdot))}} \right\|_{L^2}^2 < \infty,$$

almost surely under $\mathcal{H}_0$ *and* $\mathcal{H}_1$*, respectively, for all* n*.*

Lemma 3.2.3. *(i) Under* $\mathcal{H}_0$*, it holds that*

$$n \int_0^1 \left\| \frac{\Psi_n^\circ(u, \hat{\theta}_n) - \Psi_n^\circ(u, \theta_0)}{\sqrt{(s_n(u))(1 - s_n(u))}} \right\|^2 du \to^p 0.$$

(ii) Under $\mathcal{H}_1$*, it holds that*

$$\int_0^1 \left\| \frac{\Psi_n^\circ(u, \hat{\theta}_n) - \Psi_n^\circ(u, \theta_*)}{\sqrt{(s_n(u))(1 - s_n(u))}} \right\|^2 du \to^p 0.$$

Lemma 3.2.4. *Under* $\mathcal{H}_0$*, the sequence of random vector*

$$\left\langle \frac{\sqrt{n} I_{\theta_0}^{-1/2} \Psi_n^\circ(\cdot, \theta_0)}{\sqrt{s_n(\cdot)(1 - s_n(\cdot))}}, h \right\rangle$$

converges to $\langle G, h \rangle$ *in distribution for every* $h \in L^2([0,1], \mathbb{R}^d, du)$*, where*

$$G(u) = \frac{B_d^\circ(u)}{\sqrt{u(1-u)}}, \; u \in (0,1).$$

The following lemma is concerned with confirming Prohorov's criterion for tightness in L^2 space.

Lemma 3.2.5. *Under $\mathcal{H}_0$, the sequence of random maps*

$$\frac{\sqrt{n} I_{\theta_0}^{-1/2} \Psi_n^\circ(\cdot, \theta_0)}{\sqrt{(s_n(\cdot))(1 - s_n(\cdot))}}$$

is asymptotically finite dimensional.

Now let us start to prove Theorem 3.2.1 by using above lemmas.

PROOF OF THE LEMMA 3.2.1(I). We shall derive the asymptotic distribution of AD_n. Due to Lemma 3.2.1(i) and Lemma 3.2.3(i), it holds that

$$AD_n = \left\| \frac{\sqrt{n} I_{\theta_0}^{-1/2} \Psi_n^\circ(\cdot, \theta_0)}{\sqrt{(s_n(\cdot))(1 - s_n(\cdot))}} \right\|_{L^2}^2 + o_P(1).$$

Lemma 3.2.2-3.2.4 (i) leads that

$$\frac{\sqrt{n} I_{\theta_0}^{-1/2} \Psi_n^\circ(\cdot, \theta_0)}{\sqrt{s_n(\cdot)(1 - s_n(\cdot))}} \to^d G(\cdot) \quad \text{in } L^2([0,1], \mathbb{R}^d, du).$$

Hence, the continuous mapping theorem yields the conclusion. □

PROOF OF THE THEOREM 3.2.1(II). Due to Lemma 3.2.1(ii), Lemma 3.2.3(ii) and the continuous mapping theorem, it holds that

$$AD_n = n \times \left(\int_0^1 \frac{\Psi_n^\circ(u, \theta_*)^\top \hat{I}_n^{-1} \Psi_n^\circ(u, \theta_*)}{s_n(u)(1 - s_n(u))} du + o_P(1) \right).$$

Recall that, generally, when M is a non negative definite matrix, it holds that

$$\begin{aligned} 2(v^\top M^{-1} v + w^\top M^{-1} w) &= (v + w)^\top M^{-1}(v + w) + (v - w)^\top M^{-1}(v - w) \\ &\geq (v - w)^\top M^{-1}(v - w) \end{aligned}$$

for every $v, w \in \mathbb{R}^d$. Since I_{θ_*} is a positive definite matrix, denoting

$$A_n(u) = \frac{1}{n}\sum_{k=1}^{n}(1\{k \le nu\} - s_n(u))\,\mathbb{E}_{\theta_{(k)}}[\dot{l}_{\theta_*}(X_k)],$$

this inequality yields that

$$\begin{aligned} & 2\int_0^1 \frac{\Psi_n^\circ(u,\theta_*)^\top \hat{I}_n^{-1}\Psi_n^\circ(u,\theta_*)}{s_n(u)(1-s_n(u))}du \\ \ge\ & \int_0^1 \frac{A_n(u)^\top \hat{I}_n^{-1} A_n(u)}{s_n(u)(1-s_n(u))}du - 2\int_0^1 \frac{M_n(u)^\top \hat{I}_n^{-1} M_n(u)}{ns_n(u)(1-s_n(u))}du. \end{aligned}$$

The first term is asymptotically tight because it holds that

$$\begin{aligned} & \int_0^1 \frac{A_n(u)^\top \hat{I}_n^{-1} A_n(u)}{s_n(u)(1-s_n(u))}du \\ =\ & \int_0^1 \frac{\sum_{k=1}^n (1\{k \le nu\} - s_n(u))^2 \mathbb{E}_{\theta_{(k)}}[\dot{l}_{\theta_*}(X_k)]^\top \hat{I}_n^{-1}\mathbb{E}_{\theta_{(k)}}[\dot{l}_{\theta_*}(X_k)]}{ns_n(u)(1-s_n(u))}du \\ \le\ & \int_0^1 \frac{\sum_{k=1}^n (1\{k \le nu\} - s_n(u))^2}{ns_n(u)(1-s_n(u))}\Big(\mathbb{E}_{\theta_0}[\dot{l}_{\theta_*}(X_1)]^\top \hat{I}_n^{-1}\mathbb{E}_{\theta_0}[\dot{l}_{\theta_*}(X_1)] \\ & +\mathbb{E}_{\theta_1}[\dot{l}_{\theta_*}(X_1)]^\top \hat{I}_n^{-1}\mathbb{E}_{\theta_1}[\dot{l}_{\theta_*}(X_1)]\Big)\,du \\ =\ & \mathbb{E}_{\theta_0}[\dot{l}_{\theta_*}(X_1)]^\top \hat{I}_n^{-1}\mathbb{E}_{\theta_0}[\dot{l}_{\theta_*}(X_1)] + \mathbb{E}_{\theta_1}[\dot{l}_{\theta_*}(X_1)]^\top \hat{I}_n^{-1}\mathbb{E}_{\theta_1}[\dot{l}_{\theta_*}(X_1)], \end{aligned}$$

so it holds that

$$\int_0^1 \frac{A_n(u)^\top \hat{I}_n^{-1} A_n(u)}{s_n(u)(1-s_n(u))}du = \int_0^1 \frac{A_n(u)^\top I_{\theta_*}^{-1} A_n(u)}{s_n(u)(1-s_n(u))}du + o_P(1).$$

By the Remark 3.2.3 and the Slutsky theorem, it holds that

$$n \times \int_0^1 \frac{M_n(u)^\top \hat{I}_n^{-1} M_n(u)}{ns_n(u)(1-s_n(u))}du = \int_0^1 \frac{M_n(u)^\top I_{\theta_*}^{-1} M_n(u)}{s_n(u)(1-s_n(u))}du + o_P(1)$$

is asymptotically tight in $\mathbb{R}$. Moreover, we have

$$\int_0^1 \frac{A_n(u)^\top I_{\theta_*}^{-1} A_n(u)}{s_n(u)(1-s_n(u))}du = \int_0^{u_*} \frac{A_n(u)^\top I_{\theta_*}^{-1} A_n(u)}{s_n(u)(1-s_n(u))}du + \int_{u_*}^1 \frac{A_n(u)^\top I_{\theta_*}^{-1} A_n(u)}{s_n(u)(1-s_n(u))}du.$$

For $u < [nu_*]/n$, it holds that

$$\begin{aligned}A_n(u) &= [(1-s_n(u))s_n(u) - s_n(u)(s_n(u_*) - s_n(u))]\mathbb{E}_{\theta_0}[\dot{l}_{\theta_*}(X_1)] \\ &\quad - s_n(u)(1-s_n(u_*))\mathbb{E}_{\theta_1}[\dot{l}_{\theta_*}(X_1)] \\ &= s_n(u)(1-s_n(u_*))(\mathbb{E}_{\theta_0}[\dot{l}_{\theta_*}(X_1)] - \mathbb{E}_{\theta_1}[\dot{l}_{\theta_*}(X_1)]) \\ &\to u(1-u_*)(\mathbb{E}_{\theta_0}[\dot{l}_{\theta_*}(X_1)] - \mathbb{E}_{\theta_1}[\dot{l}_{\theta_*}(X_1)]), \end{aligned} \tag{3.2.8}$$

while for $u \geq ([nu_*]+1)/n$ it holds that

$$\begin{aligned}A_n(u) &= (1-s_n(u))s_n(u_*)\mathbb{E}_{\theta_0}[\dot{l}_{\theta_*}(X_1)] \\ &\quad + [(1-s_n(u))(s_n(u) - s_n(u_*)) - s_n(u)(1-s_n(u))]\mathbb{E}_{\theta_1}[\dot{l}_{\theta_*}(X_1)] \\ &\to (1-u)u_*(\mathbb{E}_{\theta_0}[\dot{l}_{\theta_*}(X_1)] - \mathbb{E}_{\theta_1}[\dot{l}_{\theta_*}(X_1)]), \end{aligned} \tag{3.2.9}$$

uniformly for $u \in (0,1)$. Each of the right-hand sides of (3.2.8) and (3.2.9) cannot be 0 because it holds that $\mathbb{E}_{\theta_0}[\dot{l}_{\theta_*}(X_1)] \neq \mathbf{0}$, $\mathbb{E}_{\theta_1}[\dot{l}_{\theta_*}(X_1)] \neq \mathbf{0}$ and $\mathbb{E}_{\theta_0}[\dot{l}_{\theta_*}(X_1)] \neq \mathbb{E}_{\theta_1}[\dot{l}_{\theta_*}(X_1)]$. Denoting $\Delta = \mathbb{E}_{\theta_0}[\dot{l}_{\theta_*}(X_1)] - \mathbb{E}_{\theta_1}[\dot{l}_{\theta_*}(X_1)] \neq 0$, it implies that, since I_{θ_*} is a positive definite matrix,

$$\begin{aligned} &\liminf_{n\to\infty} \int_0^{u_*} \frac{A_n(u)^\top I_{\theta_*}^{-1} A_n(u)}{s_n(u)(1-s_n(u))} du \\ \geq\ & \liminf_{n\to\infty} \int_0^{u_*} 1\left\{u < \frac{[nu_*]}{n}\right\} \frac{A_n(u)^\top I_{\theta_*}^{-1} A_n(u)}{s_n(u)(1-s_n(u))} du \\ \geq\ & \int_0^{u_*} \liminf_{n\to\infty} 1\left\{u < \frac{[nu_*]}{n}\right\} \frac{A_n(u)^\top I_{\theta_*}^{-1} A_n(u)}{s_n(u)(1-s_n(u))} du \\ =\ & \int_0^{u_*} \lim_{n\to\infty} 1\left\{u < \frac{[nu_*]}{n}\right\} \frac{A_n(u)^\top I_{\theta_*}^{-1} A_n(u)}{s_n(u)(1-s_n(u))} du \\ =\ & \int_0^{u_*} \frac{(1-u_*)\Delta^\top I_{\theta_*}^{-1}\Delta}{1-u} du = (1-u_*)\Delta^\top I_{\theta_*}^{-1}\Delta \cdot \log\frac{1}{1-u_*} \\ \geq\ & u_*(1-u_*)\Delta^\top I_{\theta_*}^{-1}\Delta > 0 \end{aligned}$$

and that

$$\begin{aligned} &\liminf_{n\to\infty} \int_{u_*}^{1} \frac{A_n(u)^\top I_{\theta_*}^{-1} A_n(u)}{s_n(u)(1-s_n(u))} du \\ \geq\ & \int_{u_*}^{1} \frac{u_*\Delta^\top I_{\theta_*}^{-1}\Delta}{u} du = u_*\Delta^\top I_{\theta_*}^{-1}\Delta \cdot \log\frac{1}{u_*} > 0. \end{aligned}$$

Therefore, we can conclude that the test is consistent. □

Now, let us prove the lemmas. As for the proofs of (ii) of the Lemmas are similar to (i) except the Lemma 3.2.3 (ii), which is rather easier , so we omit them. In the proofs,

$$\frac{1\{k \le nu\} - s_n(u)}{\sqrt{s_n(u)(1-s_n(u))}}$$

is denoted by $w_k^n(u)$ for simplicity.

Proof of the Lemma 3.2.1(i). It holds that

$$\begin{aligned}
\hat{I}_n &= \frac{1}{n}\sum_{k=1}^n \dot{l}_{\hat{\theta}_n}(X_k)\dot{l}_{\hat{\theta}_n}(X_k)^\top \\
&= \frac{1}{n}\sum_{k=1}^n \dot{l}_{\theta_0}(X_k)\dot{l}_{\theta_0}(X_k)^\top \\
&\quad + \frac{1}{n}\sum_{k=1}^n \Big(\dot{l}_{\theta_0}(X_k)(\dot{l}_{\hat{\theta}_n}(X_k) - \dot{l}_{\theta_0}(X_k))^\top + (\dot{l}_{\hat{\theta}_n}(X_k) - \dot{l}_{\theta_0}(X_k))\dot{l}_{\theta_0}(X_k)^\top\Big) \\
&\quad + \frac{1}{n}\sum_{k=1}^n (\dot{l}_{\hat{\theta}_n}(X_k) - \dot{l}_{\theta_0}(X_k))(\dot{l}_{\hat{\theta}_n}(X_k) - \dot{l}_{\theta_0}(X_k))^\top
\end{aligned}$$

The second and third term is $o_P(1)$, because the assumption (3.2.1) and the Schwartz inequality yield that

$$\begin{aligned}
&\frac{1}{n}\sum_{k=1}^n \partial_i l_{\theta_0}(X_k)(\partial_j l_{\hat{\theta}_n}(X_k) - \partial_j l_{\theta_0}(X_k)) \\
\le\ & \sqrt{\frac{1}{n}\sum_{k=1}^n (\partial_i l_{\theta_0}(X_k))^2 \frac{1}{n}\sum_{k=1}^n (\partial_j l_{\hat{\theta}_n}(X_k) - \partial_j l_{\theta_0}(X_k))^2} \\
\le\ & \sqrt{\frac{1}{n}\sum_{k=1}^n (\partial_i l_{\theta_0}(X_k))^2 \frac{1}{n}\sum_{k=1}^n (K(X_k))^2 \|\hat{\theta}_n - \theta_0\|^2} \\
\to^p\ & 0
\end{aligned}$$

by the law of large numbers, and other term is also converge to 0 in probability by the same reason. Hence, the law of large numbers yields that

$$\hat{I}_n \quad \to^p \quad \mathbb{E}[\dot{l}_{\theta_0}(X_1)\dot{l}_{\theta_0}(X_1)^\top] = I_{\theta_0}.$$

This completes the proof. □

Proof of the Lemma 3.2.2(i). Firstly, $Var[\Psi_n^\circ(u,\theta_0)]$ is calculated as follows:

$$Var[\Psi_n^\circ(u,\theta_0)] = \frac{s_n(u)(1-s_n(u))}{n} Var[\dot{l}_{\theta_0}(X_1)] = \frac{s_n(u)(1-s_n(u))}{n} I_{\theta_0}.$$

Thus, it holds that

$$\mathbb{E}\left[\frac{n\Psi_n^\circ(u,\theta_0)^\top I_{\theta_0}^{-1}\Psi_n^\circ(u,\theta_0)}{s_n(u)(1-s_n(u))}\right] = \frac{n\mathrm{tr}\left(Var\left[I_{\theta_0}^{-\frac{1}{2}}\Psi_n^\circ(u,\theta_0)\right]\right)}{s_n(u)(1-s_n(u))} = \mathrm{tr}\left(I_d\right) = d,$$

where tr(A) denotes the trace of matrix A and the notation I_d denotes d dimensional identity matrix. This completes the proof. □

Proof of the Lemma 3.2.2(ii). For the simplicity, let us denote $\dot{l}_{\theta_*}(X_k) - \mathbb{E}_{\theta_{(k)}}[\dot{l}_{\theta_*}(X_k)]$ by $\dot{l}^C_{\theta_*}(X_k)$ The left-hand side of the claim is equal to

$$\begin{aligned}
&\mathbb{E}_{\theta_{true}}\left[\frac{1}{n}\sum_{k=1}^{n}\frac{(1\{k\le nu\}-s_n(u))^2\,\dot{l}^C_{\theta_*}(X_k)^\top I_{\theta_*}^{-1}\dot{l}^C_{\theta_*}(X_k)}{s_n(u)(1-s_n(u))}\right]\\
=\ & \frac{1}{n}\sum_{k=1}^{[nu_*]}\frac{(1-s_n(u))^2\,\mathbb{E}_{\theta_0}\left[\dot{l}^C_{\theta_*}(X_1)^\top I_{\theta_*}^{-1}\dot{l}^C_{\theta_*}(X_1)\right]}{s_n(u)(1-s_n(u))}\\
&+\frac{1}{n}\sum_{k=[nu_*]+1}^{n}\frac{(s_n(u))^2\,\mathbb{E}_{\theta_1}\left[\dot{l}^C_{\theta_*}(X_1)^\top I_{\theta_*}^{-1}\dot{l}^C_{\theta_*}(X_1)\right]}{s_n(u)(1-s_n(u))}\\
=\ & (1-s_n(u))\,\mathbb{E}_{\theta_0}\left[\dot{l}^C_{\theta_*}(X_1)^\top I_{\theta_*}^{-1}\dot{l}^C_{\theta_*}(X_1)\right] + (s_n(u))\,\mathbb{E}_{\theta_1}\left[\dot{l}^C_{\theta_*}(X_1)^\top I_{\theta_*}^{-1}\dot{l}^C_{\theta_*}(X_1)\right]\\
\le\ & \mathbb{E}_{\theta_0}\left[\dot{l}^C_{\theta_*}(X_1)^\top I_{\theta_*}^{-1}\dot{l}^C_{\theta_*}(X_1)\right] + \mathbb{E}_{\theta_1}\left[\dot{l}^C_{\theta_*}(X_1)^\top I_{\theta_*}^{-1}\dot{l}^C_{\theta_*}(X_1)\right].
\end{aligned}$$

This completes the proof. □

Proof of the Lemma 3.2.3(i). It follows from the Taylor expansion that

$$\frac{\sqrt{n}\left(\Psi_n^\circ(u,\hat{\theta}_n)-\Psi_n^\circ(u,\theta_0)\right)}{\sqrt{s_n(u)(1-s_n(u))}} = \frac{1}{n}\sum_{k=1}^{n} w_k^n(u)\ddot{l}_{\tilde{\theta}_n}(X_k)\sqrt{n}(\hat{\theta}_n-\theta_0),$$

where $\tilde{\theta}_n$ is between $\hat{\theta}_n$ and θ_0. Since $\sqrt{n}(\hat{\theta}_n - \theta_0) = O_P(1)$, let us show that

$$\int_0^1 \left(\frac{1}{n} \sum_{k=1}^n w_k^n(u) (\ddot{l}_{\tilde{\theta}_n}(X_k) - \ddot{l}_{\theta_0}(X_k)) \right)^2 du \to^p 0.$$

By the Schwartz inequality, it holds that

$$\begin{aligned}
& \int_0^1 \left(\frac{1}{n} \sum_{k=1}^n w_k^n(u) (\ddot{l}_{\tilde{\theta}_n}(X_k) - \ddot{l}_{\theta_0}(X_k)) \right)^2 du \\
\leq & \int_0^1 \frac{1}{n^2} \sum_{k=1}^n w_k^n(u) \sum_{k=1}^n \left(\ddot{l}_{\tilde{\theta}_n}(X_k) - \ddot{l}_{\theta_0}(X_k) \right)^2 du \\
= & \frac{1}{n} \sum_{k=1}^n \left(\ddot{l}_{\tilde{\theta}_n}(X_k) - \ddot{l}_{\theta_0}(X_k) \right)^2. \qquad (3.2.10)
\end{aligned}$$

Because of the assumption (3.2.2), the (i,j)-th element of (3.2.10) is bounded by $\sum_{k=1}^n (K(X_k))^2 \|\tilde{\theta}_n - \theta_0\|^2/n$, which converges to 0 in probability. Next let us show that

$$\int_0^1 \left\| \frac{1}{n} \sum_{k=1}^n w_k^n(u) \ddot{l}_{\theta_0}(X_k) \sqrt{n}(\hat{\theta}_n - \theta_0) \right\|^2 du \to^p 0.$$

Since it holds that $\sqrt{n}(\hat{\theta}_n - \theta_0) = O_P(1)$, it is sufficient to show the following:

$$\begin{aligned}
& \mathbb{E}\left[\int_0^1 \left(\frac{1}{n} \sum_{k=1}^n w_k^n(u) \ddot{l}_{\theta_0}(X_k) \right)^2 du \right] \\
= & \int_0^1 \frac{1}{n^2} \mathbb{E}\left[\left(\sum_{k=1}^n w_k^n(u) \ddot{l}_{\theta_0}(X_k) \right)^2 \right] du \\
= & \int_0^1 \frac{1}{n^2} \mathbb{E}\left[\left(\sum_{k=1}^n w_k^n(u) (\ddot{l}_{\theta_0}(X_k) - \mathbb{E}[\ddot{l}_{\theta_0}(X_1)]) \right)^2 \right] du \\
= & \int_0^1 \frac{1}{n^2} \sum_{k=1}^n w_k^n(u)^2 \mathbb{E}\left[\left(\ddot{l}_{\theta_0}(X_k) - \mathbb{E}[\ddot{l}_{\theta_0}(X_1)] \right)^2 \right] du \\
= & \frac{1}{n} \left(\mathbb{E}\left[\left(\ddot{l}_{\theta_0}(X_1) \right)^2 \right] - \left(\mathbb{E}\left[\ddot{l}_{\theta_0}(X_1) \right] \right)^2 \right) \to 0.
\end{aligned}$$

This completes the proof. □

Proof of the Lemma 3.2.4. Firstly, it holds that

$$\left\langle \frac{\sqrt{n} I_{\theta_0}^{-1/2} \Psi_n^{\circ}(\cdot, \theta_0)}{\sqrt{s_n(\cdot)(1 - s_n(\cdot))}}, h \right\rangle_{L^2} = \sum_{k=1}^{n} \frac{I_{\theta_0}^{-1/2}}{\sqrt{n}} \int_0^1 w_k^n(u) \dot{l}_{\theta_0}(X_k)^\top h(u) du$$

because of the definition of the inner product. Its expectation is that

$$\sum_{k=1}^{n} \frac{I_{\theta_0}^{-1/2}}{\sqrt{n}} \int_0^1 w_k^n(u) \mathbb{E}[\dot{l}_{\theta_0}(X_k)^\top] h(u) du = 0,$$

and its variance is that

$$\begin{aligned}
& \sum_{k=1}^{n} Var\left[\frac{1}{\sqrt{n}} I_{\theta_0}^{-1/2} \int_0^1 w_k^n(u) \dot{l}_{\theta_0}(X_k)^\top h(u) du\right] \\
= & \int_0^1 \int_0^1 \frac{1}{n} \sum_{k=1}^{n} w_k^n(u) w_k^n(v) h(u) h(v)^\top du dv \\
= & \int_0^1 \int_0^1 \frac{s_n(u \wedge v) - s_n(u) s_n(v)}{\sqrt{s_n(u)(1 - s_n(u)) s_n(v)(1 - s_n(v))}} h(u) h(v)^\top du dv \\
\to & \int_0^1 \int_0^1 \frac{u \wedge v - uv}{\sqrt{uv(1-u)(1-v)}} h(u) h(v)^\top du dv.
\end{aligned}$$

as $n \to \infty$. Further, we shall show that the Lyapunov Condition:

$$\sum_{k=1}^{n} \mathbb{E}\left[\left|\int_0^1 \frac{1}{\sqrt{n}} I_{\theta_0}^{-\frac{1}{2}} w_k^n(u) \dot{l}_{\theta_0}(X_k)^\top h(u) du\right|^{2+\delta}\right] \to 0,$$

for some $\delta > 0$. The Schwartz inequality yields that

$$\begin{aligned}
& \sum_{k=1}^{n} \mathbb{E}\left[\left|\int_0^1 \frac{1}{\sqrt{n}} I_{\theta_0}^{-\frac{1}{2}} w_k^n(u) \dot{l}_{\theta_0}(X_k)^\top h(u) du\right|^{2+\delta}\right] \\
\leq & \left(\int_0^1 \|h(u)\|^2 du\right)^{\frac{2+\delta}{2}} \sum_{k=1}^{n} \mathbb{E}\left[\left|\int_0^1 \left\|\frac{1}{\sqrt{n}} I_{\theta_0}^{-\frac{1}{2}} w_k^n(u) \dot{l}_{\theta_0}(X_k)\right\|^2 du\right|^{\frac{2+\delta}{2}}\right].
\end{aligned}$$

Since h is square integrable, it is sufficient to show that

$$\sum_{k=1}^{n} \mathbb{E}\left[\left|\int_0^1 \left|\frac{1}{\sqrt{n}} w_k^n(u) \left(I_{\theta_0}^{-\frac{1}{2}} \dot{l}_{\theta_0}(X_k)\right)_{(i)}\right|^2 du\right|^{\frac{2+\delta}{2}}\right] \tag{3.2.11}$$

converges to 0 for some δ and for all i. Because of the assumption (3.2.3) for all i, it holds that

$$\begin{aligned}\mathbb{E}\left[\left|(I_{\theta_0}^{-\frac{1}{2}}\dot{l}_{\theta_0}(X_1))_{(i)}\right|^{2+\delta}\right] &\le \mathbb{E}\left[\left|(I_{\theta_0}^{-\frac{1}{2}}\mathbf{1})_{(i)}\max_{i=1,\dots,d}|\partial_i l_\theta(X_1)|\right|^{2+\delta}\right]\\ &= (I_{\theta_0}^{-\frac{1}{2}}\mathbf{1})_{(i)}\max_{i=1,\dots,d}\mathbb{E}\left[|\partial_i l_\theta(X_1)|^{2+\delta}\right]<\infty,\end{aligned}$$

where the notation $\mathbf{1}$ denotes the d dimensional vector whose all elements are 1. Therefore, since $0<\delta<2$, (3.2.11) is bounded above by

$$\begin{aligned}&\sum_{k=1}^{n}\mathbb{E}\left[\int_0^1\left|\frac{w_k^n(u)}{\sqrt{n}}\left(I_{\theta_0}^{-\frac{1}{2}}\dot{l}_{\theta_0}(X_k)\right)_{(i)}\right|^{2+\delta}du\right]\\ =\;&\int_0^1\sum_{k=1}^{n}\left|\frac{w_k^n(u)}{\sqrt{n}}\right|^{2+\delta}\mathbb{E}\left[\left|\left(I_{\theta_0}^{-\frac{1}{2}}\dot{l}_{\theta_0}(X_k)\right)_{(i)}\right|^{2+\delta}\right]du\\ =\;&\frac{1}{n^{\delta/2}}\mathbb{E}\left[\left|(I_{\theta_0}^{-\frac{1}{2}}\dot{l}_{\theta_0}(X_1))_{(i)}\right|^{2+\delta}\right]\int_0^1\frac{s_n(u)^{1+\delta}+(1-s_n(u))^{1+\delta}}{(s_n(u)(1-s_n(u)))^{\delta/2}}du\\ \le\;&\frac{1}{n^{\delta/2}}\mathbb{E}\left[\left|(I_{\theta_0}^{-\frac{1}{2}}\dot{l}_{\theta_0}(X_1))_{(i)}\right|^{2+\delta}\right]\int_0^1\frac{2}{(s_n(u)(1-s_n(u)))^{\delta/2}}du\to 0.\end{aligned}$$

In consequence, the Lyapunov condition holds. Hence, the multivariate central limit theorem yields that

$$\left\langle\frac{\sqrt{n}I_{\theta_0}^{-1/2}\Psi_n^{\circ}(\cdot,\theta_0)}{\sqrt{s_n(\cdot)(1-s_n(\cdot))}},h\right\rangle_{L^2}\to^d\langle G,h\rangle_{L^2}$$

for every $h\in L^2([0,1],du)$, where G is a Gaussian variable such that

$$\mathbb{E}[\langle G,h\rangle_{L^2}]=\mathbf{0},$$

and

$$\mathbb{E}[\langle G,h\rangle_{L^2}^2]=\int_0^1\int_0^1\frac{u\wedge v-uv}{\sqrt{uv(1-u)(1-v)}}h(u)h(v)^\top dudv,$$

for every $h\in L^2([0,1],du)$. Such a G is $B_d^{\circ}(u)/\sqrt{u(1-u)},\ u\in(0,1)$. This completes the proof. □

PROOF OF THE LEMMA 3.2.5. Let us use the Proposition 2.2.1. As for (2.2.1), it holds that

$$\mathbb{E}\left[\left\|\frac{\sqrt{n}I_{\theta_0}^{-1/2}\Psi_n^\circ(\cdot,\theta_0)}{\sqrt{(s_n(\cdot))(1-s_n(\cdot))}}\right\|_{L^2}^2\right] = d = \mathbb{E}\left[\left\|\frac{B^\circ(\cdot)}{\sqrt{\cdot(1-\cdot)}}\right\|_{L^2}^2\right].$$

Moreover, (2.2.2) holds if we take e_j as h for the convergence of

$$Var\left[\left\langle\frac{\sqrt{n}I_{\theta_0}^{-1/2}\Psi_n^\circ(\cdot,\theta_0)}{\sqrt{s_n(\cdot)(1-s_n(\cdot))}}, h\right\rangle_{L^2}\right]$$

in the proof of the Lemma 3.2.4. □

Moreover, let us provide another proof of the Lemma 3.2.5 which corrects mistake in Tsukuda and Nishiyama (2014).

Another proof of the Lemma 3.2.5 By the Parseval identity, it is sufficient to prove that

$$\lim_{J\to\infty}\limsup_{n\to\infty} E\left[\sum_{j>J}\left|\left\langle\sqrt{\frac{n}{s_n(\cdot)(1-s_n(\cdot))}}I_{\theta_0}^{-\frac{1}{2}}\Psi_n^\circ(\cdot,\theta_0), e_j\right\rangle_{L^2}\right|^2\right] = 0$$

$$\iff \lim_{J\to\infty}\limsup_{n\to\infty} E\left[\sum_{j>J}\left|\left\langle\frac{1}{\sqrt{n}}\sum_{k=1}^n w_k^n(\cdot)I_{\theta_0}^{-\frac{1}{2}}\dot{l}_{\theta_0}(X_k), e_j\right\rangle_{L^2}\right|^2\right] = 0$$

$$\iff \lim_{J\to\infty}\limsup_{n\to\infty} E\left[\left\|\frac{1}{\sqrt{n}}\sum_{k=1}^n w_k^n(\cdot)I_{\theta_0}^{-\frac{1}{2}}\dot{l}_{\theta_0}(X_k)\right\|_{L^2}^2 - \sum_{j\le J}\left|\left\langle\frac{1}{\sqrt{n}}\sum_{k=1}^n w_k^n(\cdot)I_{\theta_0}^{-\frac{1}{2}}\dot{l}_{\theta_0}(X_k), e_j\right\rangle_{L^2}\right|^2\right] = 0.$$

It follows from the Lemma 3.2.2 and the Fubini theorem that

$$E\left[\left\|1/\sqrt{n}\sum_{k=1}^n w_k^n(\cdot)I_{\theta_0}^{-\frac{1}{2}}\dot{l}_{\theta_0}(X_k)\right\|_{L^2}^2\right] = d.$$

It holds that

$$\begin{aligned}
&\limsup_{n\to\infty} E\left[-\sum_{j\le J}\left|\left\langle \frac{1}{\sqrt{n}}\sum_{k=1}^n w_k^n(\cdot) I_{\theta_0}^{-\frac{1}{2}}\dot{l}_{\theta_0}(X_k), e_j\right\rangle_{L^2}\right|^2\right] \\
= \; & -\liminf_{n\to\infty}\sum_{j\le J} E\left[\left|\frac{1}{\sqrt{n}}\sum_{k=1}^n \left\langle w_k^n(\cdot) I_{\theta_0}^{-\frac{1}{2}}\dot{l}_{\theta_0}(X_k), e_j\right\rangle_{L^2}\right|^2\right] \\
= \; & -\liminf_{n\to\infty}\sum_{j\le J} E\left[\frac{1}{n}\sum_{k=1}^n \left|\left\langle w_k^n(\cdot) I_{\theta_0}^{-\frac{1}{2}}\dot{l}_{\theta_0}(X_1), e_j\right\rangle_{L^2}\right|^2\right] \\
= \; & -\liminf_{n\to\infty}\int_0^1\int_0^1 \frac{1}{n}\sum_{k=1}^n w_k^n(u) w_k^n(v) \\
& \sum_{j=1}^J E\left[I_{\theta_0}^{-\frac{1}{2}}\dot{l}_{\theta_0}(X_1)^\top e_j(u) I_{\theta_0}^{-\frac{1}{2}}\dot{l}_{\theta_0}(X_1)^\top e_j(v)\right] dudv \\
= \; & -\liminf_{n\to\infty}\int_0^1\int_0^1 \frac{1}{n}\sum_{k=1}^n w_k^n(u) w_k^n(v) d\sum_{j=1}^J e_j(u)^\top e_j(v) dudv.
\end{aligned}$$

Since $\sum_{k=1}^n w_k^n(u)w_k^n(v) \le n$ by the Schwartz inequality and $|\sum_{j=1}^J e_j(u)^\top e_j(v)|$ is integrable with respect to $dudv$, the Fatou lemma gives the upper bound

$$\begin{aligned}
& -\int_0^1\int_0^1 \liminf_{n\to\infty}\frac{1}{n}\sum_{k=1}^n w_k^n(u) w_k^n(v) d\sum_{j=1}^J e_j(u)^\top e_j(v) dudv \\
= \; & -\int_0^1\int_0^1 \frac{u\wedge v - uv}{\sqrt{uv(1-u)(1-v)}} d\sum_{j=1}^J e_j(u)^\top e_j(v) dudv \\
= \; & -\int_0^1\int_0^1 \frac{E[B^\circ(u)B^\circ(v)]}{\sqrt{uv(1-u)(1-v)}} d\sum_{j=1}^J e'_{[\frac{j-1}{d}]+1}(u) e'_{[\frac{j-1}{d}]+1}(v) dudv \\
= \; & -dE\left[\sum_{j=1}^J\left|\left\langle \frac{B^\circ(\cdot)}{\sqrt{\cdot(1-\cdot)}}, e'_{[\frac{j-1}{d}]+1}\right\rangle_{L^2}\right|^2\right]
\end{aligned}$$

since it holds that

$$
\begin{aligned}
& \liminf_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} w_k^n(u) w_k^n(v) \\
= & \liminf_{n\to\infty} \frac{s_n(u\wedge v) - s_n(u)s_n(v)}{\sqrt{s_n(u)s_n(v)(1-s_n(u)(1-s_n(v)))}} \\
= & \frac{u\wedge v - uv}{\sqrt{uv(1-u)(1-v)}}
\end{aligned}
$$

and that

$$
\frac{s_n(u\wedge v) - s_n(u)s_n(v)}{\sqrt{s_n(u)s_n(v)(1-s_n(u)(1-s_n(v)))}} \le 1.
$$

On the other hand, the Bessel inequality implies that

$$
\sum_{j=1}^{J} \left| \left\langle \frac{B^\circ(\cdot)}{\sqrt{\cdot(1-\cdot)}}, e'_{[\frac{j-1}{d}]+1} \right\rangle_{L^2} \right|^2 \le \left\| \frac{B^\circ(\cdot)}{\sqrt{\cdot(1-\cdot)}} \right\|_{L^2}^2 .
$$

Thence, the dominated convergence theorem yields that

$$
\begin{aligned}
& -d \lim_{J\to\infty} E\left[\sum_{j=1}^{J} \left| \left\langle \frac{B^\circ(\cdot)}{\sqrt{\cdot(1-\cdot)}}, e'_{[\frac{j-1}{d}]+1} \right\rangle_{L^2} \right|^2 \right] \\
= & -dE\left[\left\| \frac{B^\circ(\cdot)}{\sqrt{\cdot(1-\cdot)}} \right\|_{L^2}^2 \right] = -d.
\end{aligned}
$$

This completes the proof.

Remark 3.2.4. *This section is mainly from Tsukuda and Nishiyama (2014). In the paper, there are some mistakes. Assumption (3.2.4) is added. The claims of their Lemma 3(ii), 4 (ii), 5(ii) and 6 (ii) contain mistakes. In this thesis, the claims of the Lemma 3(ii) and 4(ii) are corrected in the Lemma 3.2.2 and 3.2.3, respectively. The claims of the Lemmas 5(ii) and 6(ii) are deleted in the Lemma 3.2.4 and 3.2.5. The proof of their Lemma 5 (i) is wrong and the corrected one is contained in the proof of the Lemma 3.2.5 (i) of this thesis. Due to the mistake in Lemma 4 (ii), the proof of the Theorem 1 (ii) also contains some mistakes, in addition to other mistakes appeared in the equations (7), (8) and the later. These mistakes are also corrected in this thesis. The table 3 is wrong and the corrected version is the Table 7.3 in this thesis.*

3.3 Bridging a gap between the independent case and stochastic processes

The goal of this part is generalizing this independent setup to some semimartingales and establishing similar methodologies to prove weak convergences. In order to bridge this gap, let us rewrite the lemmas in the preceding section by the notation which we shall use in the Chapter 5 and Chapter 6. Define the random field

$$(u,\theta) \rightsquigarrow \mathbb{Z}_n(u,\theta) = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} w_k^n(u) \dot{l}_\theta(X_k),$$

where

$$w_k^n : (0,1) \ni u \mapsto \begin{cases} 0, & u \in \left(0, \frac{1}{n}\right), \\ \frac{1\{k \le nu\} - [nu]/n}{\sqrt{[nu]/n(1-[nu]/n)}}, & u \in \left[\frac{1}{n}, 1\right), \end{cases} \quad k = 1, \dots, n.$$

Its "predictable projection" to the true model is

$$(u,\theta) \rightsquigarrow \mathbb{Z}_n^p(u,\theta) = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} w_k^n(u) \mathbb{E}_{\theta_{(k)}}[\dot{l}_\theta(X_k)].$$

The difference between $\mathbb{Z}$ and $\mathbb{Z}^p$ is denoted by $\mathbb{M}$:

$$(u,\theta) \rightsquigarrow \mathbb{M}_n(u,\theta) = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} w_k^n(u) (\dot{l}_\theta(X_k) - \mathbb{E}_{\theta_{(k)}}[\dot{l}_\theta(X_k)]).$$

Lemma 3.3.1. *(i) Under $\mathcal{H}_0$, it holds that*

$$\mathbb{E}\left[\mathbb{Z}_n(u,\theta_0)^\top I_{\theta_0}^{-1} \mathbb{Z}_n(u,\theta_0)\right] = d$$

for all $u \in (0,1)$ and for all $n \in \mathbb{N}$.
(ii) Under $\mathcal{H}_1$, it holds that

$$\mathbb{E}_{\theta_{true}}\left[\mathbb{M}_n(u,\theta_*)^\top I_{\theta_*}^{-1} \mathbb{M}_n(u,\theta_*)\right] = d,$$

for all $u \in (0,1)$ and for all $n \in \mathbb{N}$, where $\mathbb{E}_{\theta_{true}}$ denotes integration with the true probability measure under $\mathcal{H}_1$.

Lemma 3.3.2. *(i) Under $\mathcal{H}_0$, it holds that*

$$\left\|\mathbb{Z}_n(\cdot,\hat{\theta}_n)-\mathbb{Z}_n(\cdot,\theta_0)\right\|_{L^2}^2 \to^p 0.$$

(ii) Under $\mathcal{H}_1$, it holds that

$$\frac{1}{n}\left\|\mathbb{Z}_n(\cdot,\hat{\theta}_n)-\mathbb{Z}_n(\cdot,\theta_*)\right\|_{L^2}^2 \to^p 0.$$

Lemma 3.3.3. *Under $\mathcal{H}_0$, the sequence of random vector $\langle I_{\theta_0}^{-1/2}\mathbb{Z}_n(\cdot,\theta_0),h\rangle_{L^2}$ converges to $\langle G,h\rangle_{L^2}$ in distribution for every $h\in L^2([0,1],du)$, where*

$$u \rightsquigarrow G(u)=\frac{B_d^\circ(u)}{\sqrt{u(1-u)}}.$$

Lemma 3.3.4. *Under $\mathcal{H}_0$, the sequence of random maps $I_{\theta_0}^{-1/2}\mathbb{Z}_n(\cdot,\theta_0)$ is asymptotically finite dimensional.*

Chapter 4

On convergences of some random fields

4.1 Limit theorems for stochastic integrals taking values in L^2 spaces

First of all, set a measurable space and introduce a filtration. Let us consider a locally square integrable martingale M whose predictable quadratic variation process is

$$\langle M \rangle_\cdot = \int_0^\cdot \lambda_s ds,$$

where $\{\lambda_\cdot\}$ is a non negative adapted process which satisfies

$$\sup_{s \in [0,\infty)} \mathbb{E}[\lambda_s] < \infty.$$

It leads that M is a martingale. Define the following random field:

$$(u, \theta) \rightsquigarrow \mathcal{M}_T(u, \theta) = \frac{1}{\sqrt{T}} \int_0^T w_s^T(u) H_s(\theta) dM_s,$$

where

$$w_s^T(u) = \frac{1\{s \le Tu\} - u}{\sqrt{u(1-u)}}, \quad \forall u \in (0, 1),$$

θ is an element of open bounded subset Θ of $\mathbb{R}^d$ and $H(\theta)$ is a d dimensional predictable process such that

$$\int_0^T \|H_s(\theta)\|^2 d\langle M \rangle_s < \infty, \quad a.s. \quad \forall \theta \in \Theta.$$

Proposition 4.1.1. *(i) If there exists a positive $\delta \in \mathbb{R}$ such that*

$$\int_\Theta \sup_{s\in[0,\infty)} \mathbb{E}[\|H_s(\theta)\|^{2+\delta}\lambda_s]d\theta < \infty \tag{4.1.1}$$

holds, then it holds that

$$\int_\Theta \sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right] d\theta < \infty. \tag{4.1.2}$$

(ii) If (4.1.2) holds, then it holds that

$$\mathbb{E}\left[\|\mathcal{M}_T\|^2_{L^2([0,1]\times\Theta)}\right] < \infty.$$

In particular, $\mathcal{M}_T$ takes its values in $L^2([0,1]\times\Theta)$ a.s..
(iii) If (4.1.2) holds and there exists the following limit

$$C(\theta,\eta) = \underset{T\to\infty}{\mathrm{l.i.m.}} \frac{1}{T}\int_0^T H_s(\theta)H_s(\eta)^\top\lambda_s ds, \tag{4.1.3}$$

then it holds that

$$\int_\Theta \mathrm{tr}C(\theta,\theta)d\theta < \infty.$$

Proof of the Proposition 4.1.1. (i) The Jensen inequality yields that

$$\begin{aligned}
\mathbb{E}[\|H_s(\theta)\|^2\lambda_s] &= \mathbb{E}[\|H_s(\theta)\|^2\lambda_s(1\{\|H_s(\theta)\|\geq 1\} + 1\{\|H_s(\theta)\| < 1\})] \\
&\leq \mathbb{E}[\|H_s(\theta)\|^{2+\delta}\lambda_s] + \mathbb{E}[\lambda_s] \\
&\leq \sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta)\|^{2+\delta}\lambda_s\right] + \sup_{s\in[0,\infty)} \mathbb{E}\left[\lambda_s\right].
\end{aligned}$$

The right-hand side does not depend on s, and by integrating both sides with respect to θ, we obtain the conclusion because of the assumption (4.1.1).

(ii) The left-hand side of the claim is equal to

$$
\begin{aligned}
& \mathbb{E}\left[\left\|\frac{1}{\sqrt{T}}\int_0^T w_s^T H_s dM_s\right\|^2_{L^2([0,1]\times\Theta)}\right] \\
= \ & \mathbb{E}\left[\int\int\left\|\frac{1}{\sqrt{T}}\int_0^T w_s^T(u)H_s(\theta)dM_s\right\|^2 d\theta du\right] \\
= \ & \mathbb{E}\left[\int\int\frac{1}{T}\int_0^T \left(w_s^T(u)\right)^2 \|H_s(\theta)\|^2\lambda_s ds d\theta du\right] \\
= \ & \int\int\left(\frac{1}{T}\int_0^T \left(w_s^T(u)\right)^2 \mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right] ds\right) d\theta du \\
\le \ & \int \sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right] d\theta < \infty
\end{aligned}
$$

by a martingale property and the Fubini theorem. This completes the proof.

(iii) It holds that

$$
\lim_{T\to\infty}\mathbb{E}\left[\frac{1}{T}\int_0^T H_s(\theta)H_s(\eta)^\top\lambda_s ds\right] = C(\theta,\eta)
$$

by the Jensen inequality. Hence, the conclusion is obvious. □

In order to establish the limit behavior for $\mathcal{M}_T(\cdot,\cdot)$, recalling the Theorem 2.1, what we have to prove is that $\mathcal{M}_T(\cdot,\cdot)$ is asymptotically finite dimensional and the weak convergence of the martingale $\langle\mathcal{M}_T, h\rangle_{L^2([0,1]\times\Theta)}$ taking values in $\mathbb{R}$, for every $h\in L^2([0,1]\times\Theta, du\times d\theta)$. Through proving them, we can get the following limit theorem.

Theorem 4.1.1. *Suppose that there exists a $\delta>0$ which satisfies (4.1.1) and that*

$$
\int_\Theta\int_\Theta \sup_{s\in[0,\infty)}\left\|\mathbb{E}[H_s(\theta)H_s(\eta)^\top\lambda_s]\right\|^2_{OP} d\theta d\eta < \infty \tag{4.1.4}
$$

and (4.1.3) hold. The random field $\mathcal{M}_T(\cdot,\cdot)$ converges to $\Gamma(\cdot,\cdot)$ weakly in $L^2([0,1]\times\Theta, du\times d\theta)$ as $T\to\infty$, where Γ is a Gaussian field satisfying

$$
\mathbb{E}[\langle\Gamma, h\rangle_{L^2([0,1]\times\Theta)}] = 0
$$

and

$$
\begin{aligned}
&\mathbb{E}[\langle \Gamma, h\rangle^2_{L^2([0,1]\times\Theta)}] \\
= &\int_0^1\int_0^1\int_\Theta\int_\Theta \frac{u\wedge v - uv}{\sqrt{u(1-u)v(1-v)}} h^\top(u,\theta)C(\theta,\eta)h(v,\eta)d\theta d\eta du dv,
\end{aligned}
$$

for every $h \in L^2([0,1]\times\Theta, du\times d\theta)$.

Proof of the theorem 4.1.1. We use Corollary 2.2.1. Firstly we check the criterion (2.2.1) as follows

$$
\begin{aligned}
&\mathbb{E}\left[\left\| \frac{1}{\sqrt{T}}\int_0^T w_s^T H_s dM_s\right\|^2_{L^2([0,1]\times\Theta)}\right] \\
= &\int\int\left(\frac{1}{T}\int_0^T \left(w_s^T(u)\right)^2 \mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right]ds\right)d\theta du \\
\to &\int \mathrm{tr}C(\theta,\theta)d\theta < \infty.
\end{aligned}
$$

The limit operation above is due to the dominated convergence theorem because the pointwise convergence holds by the assumtion (4.1.3).

$$
\begin{aligned}
&\frac{1}{T}\int_0^T \left(w_s^T(u)\right)^2 \mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right]ds \\
= &\left(\frac{1-u}{Tu}\int_0^{Tu} + \frac{u}{T(1-u)}\int_{Tu}^T\right)\mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right]ds \\
\to &(1-u)\mathrm{tr}C(\theta,\theta) + u\mathrm{tr}C(\theta,\theta) = \mathrm{tr}C(\theta,\theta)
\end{aligned}
$$

and it holds that

$$
\frac{1}{T}\int_0^T \left(w_s^T(u)\right)^2 \mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right]ds \le \sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right]
$$

and the right-hand side is integrable with respect to $du\times d\theta$. Next we argue the convergence of the inner product

$$
\left\langle \frac{1}{\sqrt{T}}\int_0^T w_s^T H_s dM_s, h\right\rangle_{L^2([0,1]\times\Theta)}
$$

for $h \in L^2([0,1] \times \Theta, du \times d\theta)$. The preceding display is equal to

$$\frac{1}{\sqrt{T}} \int_0^T \left\langle w_s^T H_s, h \right\rangle_{L^2([0,1]\times\Theta)} dM_s$$

by the Fubini theorem for stochastic integrals (see 5.5 of Liptser and Shiryaev (2001)) which can be applicable by the Proposition 4.1.1(ii). We shall apply the martingale CLT. As for the variance, it holds that

$$\begin{aligned} V_T &:= \mathbb{E}\left[\left(\frac{1}{\sqrt{T}} \int_0^T \left\langle w_s^T H_s, h \right\rangle_{L^2([0,1]\times\Theta)} dM_s\right)^2\right] \\ &= \mathbb{E}\left[\frac{1}{T} \int_0^T \left\langle w_s^T H_s, h \right\rangle^2_{L^2([0,1]\times\Theta)} \lambda_s ds\right] \end{aligned}$$

and the right-hand side is equal to

$$\begin{aligned} &\mathbb{E}\left[\frac{1}{T} \int_0^T \int\int\int\int w_s^T(u) w_s^T(v) h(u,\theta)^\top H_s(\theta) H_s(\eta)^\top h(v,\eta) d\theta d\eta du dv \lambda_s ds\right] \\ = &\int\int\int\int \frac{1}{T} \int_0^T w_s^T(u) w_s^T(v) h(u,\theta)^\top \mathbb{E}\left[H_s(\theta) H_s(\eta)^\top \lambda_s\right] h(v,\eta) ds d\theta d\eta du dv \end{aligned}$$

We shall check that the dominated convergence theorem which yields that

$$V_T \to \int\int\int\int \frac{u \wedge v - uv}{\sqrt{u(1-u)v(1-v)}} h(u,\theta)^\top C(\theta,\eta) h(v,\eta) d\theta d\eta du dv$$

can be applied. The pointwise convergence holds by the assumption (4.1.3).

By the Schwartz inequality, it holds that

$$
\begin{aligned}
&\frac{1}{T}\int_0^T w_s^T(u)w_s^T(v)h(u,\theta)^\top \mathbb{E}\left[H_s(\theta)H_s(\eta)^\top \lambda_s\right] h(v,\eta)ds \\
\leq\ & \sqrt{\frac{1}{T^2}\int_0^T \left(w_s^T(u)h(u,\theta)^\top \mathbb{E}\left[H_s(\theta)H_s(\eta)^\top\lambda_s\right] h(v,\eta)\right)^2 ds \int_0^T (w_s^T(v))^2 ds} \\
\leq\ & \sqrt{\frac{1}{T}\int_0^T (w_s^T(u))^2 \sup_{s\in[0,\infty)} \left(h(u,\theta)^\top \mathbb{E}\left[H_s(\theta)H_s(\eta)^\top\lambda_s\right] h(v,\eta)\right)^2 ds} \\
=\ & \sup_{s\in[0,\infty)} \left|h(u,\theta)^\top \mathbb{E}\left[H_s(\theta)H_s(\eta)^\top\lambda_s\right] h(v,\eta)\right| \\
\leq\ & \|h(u,\theta)\| \sup_{s\in[0,\infty)} \left\|\mathbb{E}\left[H_s(\theta)H_s(\eta)^\top\lambda_s\right] h(v,\eta)\right\| \\
\leq\ & \|h(u,\theta)\|\|h(v,\eta)\| \sup_{s\in[0,\infty)} \left\|\mathbb{E}\left[H_s(\theta)H_s(\eta)^\top\lambda_s\right]\right\|_{OP},
\end{aligned}
$$

The right-hand side is integrable with respect to $du \times dv \times d\theta \times d\eta$ by the Schwartz inequality and by the assumption (4.1.4). Therefore, we can apply the dominated convergence theorem. It also leads that (2.2.2) holds. Finally, let us check the Lyapunov condition:

$$
\mathbb{E}\left[\frac{1}{T^{(2+\delta_0)/2}}\int_0^T \left\langle w_s^T H_s, h\right\rangle_{L^2([0,1]\times\Theta)}^{2+\delta_0} \lambda_s ds\right] \to 0
$$

for some $\delta_0 > 0$. The Schwartz inequality gives the upper bound of the left-hand side

$$
\frac{1}{T^{(2+\delta_0)/2}}\mathbb{E}\left[\int_0^T \left\|w_s^T H_s\right\|_{L^2([0,1]\times\Theta)}^{2+\delta_0} \lambda_s ds\right] \|h\|_{L^2([0,1]\times\Theta)}^{2+\delta_0}.
$$

Next, it follows from the Jensen inequality that

$$
\begin{aligned}
&\frac{1}{T^{(2+\delta_0)/2}}\mathbb{E}\left[\int_0^T \left\|w_s^T H_s\right\|_{L^2([0,1]\times\Theta)}^{2+\delta_0} \lambda_s ds\right] \\
\leq\ & \frac{1}{T^{(2+\delta_0)/2}}\mathbb{E}\left[\int_0^T \int\int \|w_s^T(u)H_s(\theta)\|^{2+\delta_0} dud\theta\lambda_s ds\right] \\
=\ & \frac{1}{T^{(2+\delta_0)/2}}\int\int\int_0^T |w_s^T(u)|^{2+\delta_0}\mathbb{E}\left[\|H_s(\theta)\|^{2+\delta_0}\lambda_s\right] dsdud\theta \\
\leq\ & \frac{1}{T^{\delta_0/2}}\int \frac{u^{1+\delta_0}+(1-u)^{1+\delta_0}}{(u(1-u))^{\delta_0/2}}du \int \sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta)\|^{2+\delta_0}\lambda_s\right] d\theta.
\end{aligned}
$$

If we choose $\delta_0 < 2$, which is possible by the assumption (4.1.1), the right-hand side converges to 0, since

$$\begin{aligned} & \int \frac{u^{1+\delta_0} + (1-u)^{1+\delta_0}}{(u(1-u))^{\delta_0/2}} du \\ \leq \; & \int \left(\frac{1}{(1-u)^{\delta_0/2}} + \frac{1}{u^{\delta_0/2}} \right) du = 2 \int \frac{1}{u^{\delta_0/2}} du = 2. \end{aligned}$$

Hence, the martingale central limit theorem yields the conclusion. □

The following proposition and theorem are corresponding one to the proposition 4.1.1 (i)(ii) and the theorem 4.1.1, when fixing θ.

Proposition 4.1.2. *Fix a $\theta \in \Theta$. (i) If there exists a positive $\delta \in \mathbb{R}$ such that*

$$\sup_{s \in [0,\infty)} \mathbb{E}[\|H_s(\theta)\|^{2+\delta} \lambda_s] < \infty \tag{4.1.5}$$

holds, then it holds that

$$\sup_{s \in [0,\infty)} \mathbb{E}\left[\|H_s(\theta)\|^2 \lambda_s\right] < \infty. \tag{4.1.6}$$

(ii) If (4.1.6) holds, then it holds that

$$\mathbb{E}\left[\|\mathcal{M}_T(\cdot,\theta)\|^2_{L^2([0,1])}\right] < \infty.$$

In particular, $\mathcal{M}_T(\cdot,\theta)$ almost surely takes its values in $L^2([0,1], du)$.

Proof of the Proposition 4.1.2. (i) The Jensen inequality yields that

$$\begin{aligned} \mathbb{E}[\|H_s(\theta)\|^2 \lambda_s] &= \mathbb{E}[\|H_s(\theta)\|^2 \lambda_s (1\{\|H_s(\theta)\| \geq 1\} + 1\{\|H_s(\theta)\| < 1\})] \\ &\leq \mathbb{E}[\|H_s(\theta)\|^{2+\delta} \lambda_s] + \mathbb{E}[\lambda_s] \\ &\leq \sup_{s \in [0,\infty)} \mathbb{E}\left[\|H_s(\theta)\|^{2+\delta} \lambda_s\right] + \sup_{s \in [0,\infty)} \mathbb{E}\left[\lambda_s\right]. \end{aligned}$$

This completes the proof.

(ii) The left-hand side of the claim is equal to

$$\begin{aligned}
& \mathbb{E}\left[\int_0^1 \left\| \frac{1}{\sqrt{T}} \int_0^T w_s^T(u) H_s(\theta) dM_s \right\|^2 du\right] \\
= & \int_0^1 \mathbb{E}\left[\frac{1}{T}\int_0^T \left(w_s^T(u)\right)^2 \|H_s(\theta)\|^2 \lambda_s ds\right] du \\
= & \int_0^1 \left(\frac{1}{T}\int_0^T \left(w_s^T(u)\right)^2 \mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right] ds\right) du \\
\leq & \sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right] < \infty
\end{aligned}$$

by a martingale property and the Fubini theorem. This completes the proof.

Theorem 4.1.2. *Fix a $\theta \in \Theta$. Suppose that there exists the following limit*

$$C(\theta, \eta) = \underset{T\to\infty}{\text{l.i.m.}} \frac{1}{T}\int_0^T H_s(\theta) H_s(\eta)^\top \lambda_s ds. \tag{4.1.7}$$

If there exists a positive $\delta \in \mathbb{R}$ which satisfies

$$\sup_{s\in[0,\infty)} \mathbb{E}[\|H_s(\theta)\|^{2+\delta}\lambda_s] < \infty \tag{4.1.8}$$

and it holds that

$$\sup_{s\in[0,\infty)} \left\|\mathbb{E}[H_s(\theta)H_s(\theta)^\top\lambda_s]\right\|_{OP} < \infty, \tag{4.1.9}$$

then the random field $\mathcal{M}_T(\cdot,\theta)$ converges to

$$\Gamma(\cdot,\theta) = \frac{C(\theta,\theta)^{1/2} B_d^\circ(\cdot)}{w(\cdot)}$$

weakly in $L^2([0,1], du)$ as $T \to \infty$, where B_d° denotes the d dimensional standard Brownian bridge and $w(u) = \sqrt{u(1-u)}$ for $u \in [0.1]$.

Proof of the theorem 4.1.2. We use Corollary 2.2.1. Firstly we check the criterion (2.2.1) as follows

$$\begin{aligned}
& \mathbb{E}\left[\left\| \frac{1}{\sqrt{T}} \int_0^T w_s^T H_s(\theta) dM_s \right\|^2_{L^2([0,1])}\right] \\
= & \int \left(\frac{1}{T}\int_0^T \left(w_s^T(u)\right)^2 \mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right] ds\right) du \\
\to & \operatorname{tr} C(\theta,\theta) < \infty.
\end{aligned}$$

The limit operation above is due to the bounded convergence theorem because the pointwise convergence holds by the assumption (4.1.7)

$$\begin{aligned}
& \frac{1}{T}\int_0^T \left(w_s^T(u)\right)^2 \mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right] ds \\
= & \left(\frac{1-u}{Tu}\int_0^{Tu} + \frac{u}{T(1-u)}\int_{Tu}^{T}\right) \mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right] ds \\
\to & (1-u)\mathrm{tr}C(\theta,\theta) + u\mathrm{tr}C(\theta,\theta) = \mathrm{tr}C(\theta,\theta)
\end{aligned}$$

for all $u \in (0,1)$ and it holds that

$$\frac{1}{T}\int_0^T \left(w_s^T(u)\right)^2 \mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right] ds \le \sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right].$$

Next we argue the convergence of the inner product

$$\left\langle \frac{1}{\sqrt{T}}\int_0^T w_s^T H_s(\theta) dM_s, h\right\rangle_{L^2([0,1])}$$

for $h \in L^2([0,1], du)$. The preceding display is equal to

$$\frac{1}{\sqrt{T}}\int_0^T \left\langle w_s^T H_s(\theta), h\right\rangle_{L^2([0,1])} dM_s$$

by the Fubini theorem for stochastic integrals. We shall apply the martingale CLT. As for the variance, it holds that

$$\begin{aligned}
V_T &:= \mathbb{E}\left[\left(\frac{1}{\sqrt{T}}\int_0^T \left\langle w_s^T H_s(\theta), h\right\rangle_{L^2([0,1])} dM_s\right)^2\right] \\
&= \mathbb{E}\left[\frac{1}{T}\int_0^T \left\langle w_s^T H_s(\theta), h\right\rangle_{L^2([0,1])}^2 \lambda_s ds\right]
\end{aligned}$$

and the right-hand side is equal to

$$\begin{aligned}
& \mathbb{E}\left[\frac{1}{T}\int_0^T \int\int w_s^T(u) w_s^T(v) h(u)^\top H_s(\theta) H_s(\theta)^\top h(v) du dv \lambda_s ds\right] \\
= & \int\int \frac{1}{T}\int_0^T w_s^T(u) w_s^T(v) h(u)^\top \mathbb{E}\left[H_s(\theta) H_s(\theta)^\top \lambda_s\right] h(v) ds du dv
\end{aligned}$$

We shall check that the dominated convergence theorem which yields that

$$V_T \to \int\int \frac{u \wedge v - uv}{\sqrt{u(1-u)v(1-v)}} h(u)^\top C(\theta, \theta) h(v) du dv$$

can be applied. The pointwise convergence holds by the assumption (4.1.7). Because of the Schwartz inequality, it holds that

$$\begin{aligned}
& \frac{1}{T} \int_0^T w_s^T(u) w_s^T(v) h(u)^\top \mathbb{E}\left[H_s(\theta) H_s(\theta)^\top \lambda_s\right] h(v) ds \\
\leq & \sqrt{\frac{1}{T^2} \int_0^T \left(w_s^T(u) h(u)^\top \mathbb{E}\left[H_s(\theta) H_s(\theta)^\top \lambda_s\right] h(v)\right)^2 ds \int_0^T (w_s^T(v))^2 ds} \\
\leq & \sqrt{\frac{1}{T} \int_0^T (w_s^T(u))^2 \sup_{s \in [0,\infty)} \left(h(u)^\top \mathbb{E}\left[H_s(\theta) H_s(\theta)^\top \lambda_s\right] h(v)\right)^2 ds} \\
= & \sup_{s \in [0,\infty)} \left| h(u)^\top \mathbb{E}\left[H_s(\theta) H_s(\theta)^\top \lambda_s\right] h(v) \right|.
\end{aligned}$$

The right-hand side is integrable by the Schwartz inequality for the Euclid inner product, which gives the upper bound of the right-hand side

$$\|h(u)\| \|h(v)\| \sup_{s \in [0,\infty)} \left\| \mathbb{E}\left[H_s(\theta) H_s(\theta)^\top \lambda_s\right] \right\|_{OP},$$

and by the assumption (4.1.9). Therefore, we can apply the dominated convergence theorem. It also leads that (2.2.2) holds. Finally, let us check the Lyapunov condition:

$$\mathbb{E}\left[\frac{1}{T^{(2+\delta_0)/2}} \int_0^T \left\langle w_s^T H_s(\theta), h \right\rangle_{L^2([0,1])}^{2+\delta_0} \lambda_s ds\right] \to 0$$

for some $\delta_0 > 0$. The Schwartz inequality and the Jensen inequality give the upper bound of the left-hand side

$$\begin{aligned}
& \frac{1}{T^{(2+\delta_0)/2}} \mathbb{E}\left[\int_0^T \left\| w_s^T H_s(\theta) \right\|_{L^2([0,1])}^{2+\delta_0} \lambda_s ds\right] \|h\|_{L^2([0,1])}^{2+\delta_0} \\
\leq & \frac{1}{T^{(2+\delta_0)/2}} \mathbb{E}\left[\int_0^T \int \|w_s^T(u) H_s(\theta)\|^{2+\delta_0} du \lambda_s ds\right] \|h\|_{L^2([0,1])}^{2+\delta_0} \\
= & \frac{1}{T^{(2+\delta_0)/2}} \int \int_0^T |w_s^T(u)|^{2+\delta_0} \mathbb{E}\left[\|H_s(\theta)\|^{2+\delta_0} \lambda_s\right] ds du \|h\|_{L^2([0,1])}^{2+\delta_0} \\
\leq & \frac{1}{T^{\delta_0/2}} \int \frac{u^{1+\delta_0} + (1-u)^{1+\delta_0}}{(u(1-u))^{\delta_0/2}} du \sup_{s \in [0,\infty)} \mathbb{E}\left[\|H_s(\theta)\|^{2+\delta_0} \lambda_s\right] \|h\|_{L^2([0,1])}^{2+\delta_0}.
\end{aligned}$$

If we choose $\delta_0 < 2$, which is possible by the assumption (4.1.8), the right-hand side converges to 0. Hence, the martingale central limit theorem yields the conclusion. $\square$

4.2 Limit theorems for discrete time martingales taking values in L^2 spaces

First of all, set a measurable space and introduce a filfration. Let us consider a martingale difference sequence $\{\xi_\cdot\}$ which satisfies

$$\sup_{k=1,2,\dots} \mathbb{E}[\xi_k^2] < \infty.$$

Define the following random field:

$$(u,\theta) \rightsquigarrow \mathcal{M}_n(u,\theta) = \frac{1}{\sqrt{n}} \sum_{k=1}^n w_k^n(u) H_{k-1}(\theta)\xi_k,$$

where

$$w_k^n(u) = \begin{cases} 0, & u \in \left(0, \frac{1}{n}\right), \\ \frac{1\{k \le nu\} - [nu]/n}{\sqrt{[nu]/n(1-[nu]/n)}}, & u \in \left[\frac{1}{n}, 1\right), \end{cases} \quad k = 1, \dots, n,$$

θ is an element of open bounded subset Θ of $\mathbb{R}^d$ and $H(\theta)$ satisfies that

$$\sum_{k=1}^n \|H_{k-1}(\theta)\|^2 \mathbb{E}[\xi_k^2 | \mathcal{F}_{k-1}] < \infty, \quad a.s..$$

Proposition 4.2.1. *(i) If there exists a $\delta > 0$ such that*

$$\int_\Theta \sup_{k=1,2,\dots} \mathbb{E}[\|H_{k-1}(\theta)\|^{2+\delta}\xi_k^2] d\theta < \infty. \tag{4.2.1}$$

holds, then it holds that

$$\int_\Theta \sup_{k=1,2,\dots} \mathbb{E}\left[\|H_{k-1}(\theta)\|^2 \xi_k^2\right] d\theta < \infty. \tag{4.2.2}$$

(ii) If (4.2.2) holds, then it holds that

$$\mathbb{E}\left[\|\mathcal{M}_n\|^2_{L^2([0,1]\times\Theta)}\right] < \infty.$$

In particular, $\mathcal{M}_n$ takes values in $L^2([0,1]\times\Theta)$ a.s..
(iii) If (4.2.2) holds and there exists the following limit

$$C(\theta,\eta) = \operatorname*{l.i.m.}_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} H_{k-1}(\theta)H_{k-1}(\eta)^\top \mathbb{E}[\xi_k^2|\mathcal{F}_{k-1}], \tag{4.2.3}$$

then it holds that

$$\int_\Theta \mathrm{tr}C(\theta,\theta)d\theta < \infty.$$

Proof of the Proposition 4.2.1. (i) The Jensen inequality yields that

$$\begin{aligned}
& \mathbb{E}[\|H_{k-1}(\theta)\|^2\xi_k^2] \\
= \ & \mathbb{E}[\|H_{k-1}(\theta)\|^2\xi_k^2(1\{\|H_{k-1}(\theta)\|\geq 1\} + 1\{\|H_{k-1}(\theta)\| < 1\})] \\
\leq \ & \mathbb{E}[\|H_{k-1}(\theta)\|^{2+\delta}\xi_k^2] + \mathbb{E}[\xi_k^2] \\
\leq \ & \sup_{k=1,2,\ldots} \mathbb{E}\left[\|H_{k-1}(\theta)\|^{2+\delta}\xi_k^2\right] + \sup_{k=1,2,\ldots} \mathbb{E}\left[\xi_k^2\right].
\end{aligned}$$

The right-hand side does not depend on s, and by integrating both sides with respect to θ, we obtain the conclusion because of the assumption (4.2.1).

(ii) The left-hand side of the claim is equal to

$$\begin{aligned}
& \mathbb{E}\left[\left\|\frac{1}{\sqrt{n}}\sum_{k=1}^n w_k^n H_{k-1}\xi_k\right\|^2_{L^2([0,1]\times\Theta)}\right] \\
= \ & \mathbb{E}\left[\int\int\left\|\frac{1}{\sqrt{n}}\sum_{k=1}^n w_k^n(u) H_{k-1}(\theta)\xi_k\right\|^2 d\theta du\right] \\
= \ & \mathbb{E}\left[\int\int\frac{1}{n}\sum_{k=1}^n (w_k^n(u))^2\|H_{k-1}(\theta)\|^2\xi_k^2 d\theta du\right] \\
= \ & \int\int\left(\frac{1}{n}\sum_{k=1}^n (w_k^n(u))^2\,\mathbb{E}\left[\|H_{k-1}(\theta)\|^2\xi_k^2\right]\right) d\theta du \\
\leq \ & \int \sup_{k=1,2,\ldots}\mathbb{E}\left[\|H_{k-1}(\theta)\|^2\xi_k^2\right] d\theta < \infty
\end{aligned}$$

by martingale property and the Fubini theorem. This completes the proof.

(iii) It holds that

$$\lim_{n\to\infty} \mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n} H_{k-1}(\theta)H_{k-1}(\eta)^\top \xi_k^2\right] = C(\theta,\eta)$$

by the Jensen inequality. Hence, the conclusion is obvious. □

Theorem 4.2.1. *Suppose that there exists a* $\delta > 0$ *which satisfies (4.2.1) holds and that*

$$\int_\Theta \int_\Theta \sup_{k=1,2,\ldots} \left\| \mathbb{E}[H_{k-1}(\theta)H_{k-1}(\theta)^\top \xi_k^2] \right\|_{OP}^2 d\theta d\eta < \infty \tag{4.2.4}$$

and (4.2.3) hold. The random field $\mathcal{M}_n(\cdot,\cdot)$ *converges to* $\Gamma(\cdot,\cdot)$ *weakly in* $L^2([0,1]\times\Theta, du\times d\theta)$ *as* $n\to\infty$*, where* Γ *is a Gaussian field satisfying*

$$\mathbb{E}[\langle \Gamma, h\rangle_{L^2([0,1]\times\Theta)}] = 0$$

and

$$\begin{aligned}&\mathbb{E}[\langle \Gamma, h\rangle^2_{L^2([0,1]\times\Theta)}] \\ =\ & \int_0^1\int_0^1\int_\Theta\int_\Theta \frac{u\wedge v - uv}{\sqrt{u(1-u)v(1-v)}} h^\top(u,\theta)C(\theta,\eta)h(v,\eta)d\theta d\eta du dv,\end{aligned}$$

for every $h \in L^2([0,1]\times\Theta, du\times d\theta)$*.*

PROOF OF THE THEOREM 4.2.1. We use Corollary 2.2.1. Let us check the criterion (2.2.1) as follows

$$\begin{aligned}&\mathbb{E}\left[\int\int \left\| \frac{1}{\sqrt{n}}\sum_{k=1}^{n} w_k^n(u)H_{k-1}(\theta)\xi_k\right\|^2 d\theta du\right] \\ =\ & \int\int \mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n}(w_k^n(u))^2\|H_{k-1}(\theta)\|^2\xi_k^2\right] d\theta du \\ \to\ & \int \mathrm{tr} C(\theta,\theta)d\theta < \infty.\end{aligned}$$

The limit operation above holds by the dominated convergence theorem because (4.2.3) yields the pointwise convergence

$$\begin{aligned} & \mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n}(w_k^n(u))^2\|H_{k-1}(\theta)\|^2\xi_k^2\right] \\ = & \mathbb{E}\left[\left((1-u)\frac{1}{[nu]}\sum_{k=1}^{[nu]} + u\frac{1}{n-[nu]}\sum_{k=[nu]+1}^{n}\right)\|H_{k-1}(\theta)\|^2\xi_k^2\right] \\ \to & (1-u)\mathrm{tr}C(\theta,\theta) + u\mathrm{tr}C(\theta,\theta) = \mathrm{tr}C(\theta,\theta) \end{aligned}$$

and because the domination holds by the assumption

$$\begin{aligned} & \mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n}(w_k^n(u))^2\|H_{k-1}(\theta)\|^2\xi_k^2\right] \\ \leq & \frac{1}{n}\sum_{k=1}^{n}(w_k^n(u))^2 \sup_{k=1,2,\ldots}\mathbb{E}\left[\|H_{k-1}(\theta)\|^2\xi_k^2\right] \\ = & \sup_{k=1,2,\ldots}\mathbb{E}\left[\|H_{k-1}(\theta)\|^2\xi_k^2\right]. \end{aligned}$$

The convergence of the inner product

$$\left\langle \frac{1}{\sqrt{n}}\sum_{k=1}^{n} w_k^n H_{k-1}\xi_k, h\right\rangle_{L^2([0,1]\times\Theta)} = \frac{1}{\sqrt{n}}\sum_{k=1}^{n}\langle w_k^n H_{k-1}, h\rangle_{L^2([0,1]\times\Theta)}\,\xi_k$$

shall be given by the martingale central limit theorem. As for the variance, the tower property yields that

$$\begin{aligned} V_n := & \mathbb{E}\left[\left(\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\langle w_k^n H_{k-1}, h\rangle_{L^2([0,1]\times\Theta)}\,\xi_k\right)^2\right] \\ = & \mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n}\langle w_k^n H_{k-1}, h\rangle_{L^2([0,1]\times\Theta)}^2\,\xi_k^2\right] \end{aligned}$$

and the right-hand side is equal to

$$\begin{aligned} & \mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n}\int\int\int\int w_k^n(u)w_k^n(v)h(u,\theta)^\top H_{k-1}(\theta)H_{k-1}(\eta)^\top h(v,\eta)d\theta d\eta du dv\xi_k^2\right] \\ = & \int\int\int\int\frac{1}{n}\sum_{k=1}^{n} w_k^n(u)w_k^n(v)h(u,\theta)^\top\mathbb{E}\left[H_{k-1}(\theta)H_{k-1}(\eta)^\top\xi_k^2\right]h(v,\eta)d\theta d\eta du dv \end{aligned}$$

We shall check that the dominated convergence theorem which yields that

$$V_n \to \int\int\int\int \frac{u \wedge v - uv}{\sqrt{u(1-u)v(1-v)}} h(u,\theta)^\top C(\theta,\eta) h(v,\eta) d\theta d\eta du dv,$$

can be applied. The pointwise convergence holds by the assumption (4.2.3). Because of the Schwartz inequality, it holds that

$$\begin{aligned}
& \frac{1}{n}\sum_{k=1}^{n} w_k^n(u) w_k^n(v) h(u,\theta)^\top \mathbb{E}\left[H_{k-1}(\theta)H_{k-1}(\eta)^\top \xi_k^2\right] h(v,\eta) \\
\leq & \sqrt{\frac{1}{n^2}\sum_{k=1}^{n}(w_k^n(u) h(u,\theta)^\top \mathbb{E}\left[H_{k-1}(\theta)H_{k-1}(\eta)^\top \xi_k^2\right] h(v,\eta))^2 \sum_{k=1}^{n}(w_k^n(v))^2} \\
\leq & \sqrt{\frac{1}{n}\sum_{k=1}^{n}(w_k^n(u))^2 \sup_{k=1,2,\ldots} (h(u,\theta)^\top \mathbb{E}\left[H_{k-1}(\theta)H_{k-1}(\eta)^\top \xi_k^2\right] h(v,\eta))^2} \\
= & \sup_{k=1,2,\ldots} \left| h(u,\theta)^\top \mathbb{E}\left[H_{k-1}(\theta)H_{k-1}(\eta)^\top \xi_k^2\right] h(v,\eta)\right| \\
\leq & \|h(u,\theta)\| \sup_{k=1,2,\ldots} \left\| \mathbb{E}\left[H_{k-1}(\theta)H_{k-1}(\eta)^\top \xi_k^2\right] h(v,\eta)\right\| \\
\leq & \|h(u,\theta)\|\|h(v,\eta)\| \sup_{k=1,2,\ldots} \left\| \mathbb{E}\left[H_{k-1}(\theta)H_{k-1}(\eta)^\top \xi_k^2\right]\right\|_{OP}.
\end{aligned}$$

The right-hand side is integrable by the Schwartz inequality and by the assumption (4.2.4). Therefore, we can apply the dominated convergence theorem. It also leads that (2.2.2) holds. Finally, let us check the Lyapunov condition:

$$\mathbb{E}\left[\frac{1}{n^{(2+\delta_0)/2}}\sum_{k=1}^{n} \langle w_k^n H_{k-1}, h\rangle_{L^2([0,1]\times\Theta)}^{2+\delta_0} \xi_k^2\right] \to 0$$

for some $\delta_0 > 0$. The Schwartz inequality gives the upper bound of the right-hand side

$$\frac{1}{n^{(2+\delta_0)/2}}\mathbb{E}\left[\sum_{k=1}^{n} \|w_k^n H_{k-1}\|_{L^2([0,1]\times\Theta)}^{2+\delta_0} \xi_k^2\right] \|h\|_{L^2([0,1]\times\Theta)}^{2+\delta_0}.$$

Then, the Jensen inequality yields that

$$\begin{aligned}
& \frac{1}{n^{(2+\delta_0)/2}}\mathbb{E}\left[\sum_{k=1}^{n} \|w_k^n H_s\|_{L^2([0,1]\times\Theta)}^{2+\delta_0} \xi_k^2\right] \\
\leq \quad & \frac{1}{n^{(2+\delta_0)/2}}\mathbb{E}\left[\sum_{k=1}^{n} \int\int \|w_k^n(u) H_{k-1}(\theta)\|^{2+\delta_0} dud\theta \xi_k^2\right] \\
= \quad & \frac{1}{n^{(2+\delta_0)/2}} \int\int \sum_{k=1}^{n} |w_k^n(u)|^{2+\delta_0} \mathbb{E}\left[\|H_{k-1}(\theta)\|^{2+\delta_0}\xi_k^2\right] dud\theta \\
\leq \quad & \frac{1}{n^{\delta_0/2}} \int_{\frac{1}{n}}^{1} W^n(u)du \int \sup_{k=1,2,\ldots} \mathbb{E}\left[\|H_{k-1}(\theta)\|^{2+\delta_0}\xi_k^2\right] d\theta,
\end{aligned}$$

where

$$W^n(u) = \frac{[nu]^{1+\delta_0} + (n-[nu])^{1+\delta_0}}{([nu](n-[nu]))^{\delta_0/2}}.$$

If we choose $\delta_0 < 2$, which is possible by the assumption (4.1.1), the right-hand side converges to 0, since

$$\begin{aligned}
\int_{\frac{1}{n}}^{1} W^n(u)du \quad & \leq \quad \int_{\frac{1}{n}}^{1} \left(\frac{1}{(1-[nu]/n)^{\delta_0/2}} + \frac{1}{([nu]/n)^{\delta_0/2}} \right) du \\
& < \quad \int_{\frac{1}{n}}^{1} \left(\frac{1}{(1-u)^{\delta_0/2}} + \frac{1}{(u-1/n)^{\delta_0/2}} \right) du \\
& = \quad 2\int_{0}^{1-\frac{1}{n}} \frac{1}{u^{\delta_0/2}} du = 2\left(1-\frac{1}{n}\right)^{1-\delta_0/2} < 2.
\end{aligned}$$

Hence, the martingale central limit theorem yields the conclusion. □

The following proposition and theorem are corresponding one to the proposition 4.2.1 (i)(ii) and the theorem 4.2.1, when fixing θ. The proof of the proposition 4.2.2 is omitted because it is very similar to the proof of the proposition 4.2.1.

Proposition 4.2.2. *Fix a $\theta \in \Theta$. (i) If there exists a positive $\delta \in \mathbb{R}$ such that*

$$\sup_{k=1,2,\ldots} \mathbb{E}[\|H_{k-1}(\theta)\|^{2+\delta}\xi_k^2] < \infty \tag{4.2.5}$$

holds, then it holds that

$$\sup_{k=1,2,\ldots} \mathbb{E}\left[\|H_{k-1}(\theta)\|^{2}\xi_k^2\right] < \infty. \tag{4.2.6}$$

(ii) If (4.2.6) holds, then it holds that

$$\mathbb{E}\left[\|\mathcal{M}_n(\cdot,\theta)\|^2_{L^2([0,1])}\right] < \infty.$$

In particular, $\mathcal{M}_T(\cdot,\theta)$ almost surely takes its values in $L^2([0,1],du)$.

Theorem 4.2.2. *Fix a $\theta \in \Theta$. Suppose that there exists the following limit*

$$C(\theta,\eta) = \underset{n\to\infty}{\text{l.i.m.}}\frac{1}{n}\sum_{k=1}^{n} H_{k-1}(\theta)H_{k-1}(\eta)^\top \mathbb{E}[\xi_k^2|\mathcal{F}_{k-1}]. \tag{4.2.7}$$

If there exists a positive $\delta \in \mathbb{R}$ which satisfies

$$\sup_{k=1,2,\ldots} \mathbb{E}[\|H_{k-1}(\theta)\|^{2+\delta}\xi_k^2] < \infty \tag{4.2.8}$$

and it holds that

$$\sup_{k=1,2,\ldots} \left\|\mathbb{E}[H_{k-1}(\theta)H_{k-1}(\theta)^\top \xi_k^2]\right\|_{OP} < \infty, \tag{4.2.9}$$

then the random field $\mathcal{M}_n(\cdot,\theta)$ converges to

$$\Gamma(\cdot,\theta) = \frac{C(\theta,\theta)^{1/2}B_d^\circ(\cdot)}{w(\cdot)}$$

weakly in $L^2([0,1],du)$ as $n\to\infty$, where B_d° denotes the d dimensional Brownian bridge and $w(u)=\sqrt{u(1-u)}$ for $u \in [0.1]$.

Proof of the theorem 4.2.2. We use Corollary 2.2.1. Let us check the criterion (2.2.1) as follows

$$\begin{aligned}
&\mathbb{E}\left[\int\int\left\|\frac{1}{\sqrt{n}}\sum_{k=1}^{n} w_k^n(u)H_{k-1}(\theta)\xi_k\right\|^2 d\theta du\right] \\
= &\int \mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n}(w_k^n(u))^2 H_{k-1}(\theta)^\top H_{k-1}(\theta)\xi_k^2\right] du \\
\to &\ \mathrm{tr}C(\theta,\theta) < \infty.
\end{aligned}$$

The last inequality is led by the definition of $C(\theta, \eta)$ and the assumption (4.2.8). The limit operation above is due to the bounded convergence theorem because the assumption (4.2.7) yields the pointwise convergence

$$\begin{aligned} & \mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n}(w_k^n(u))^2 H_{k-1}(\theta)^\top H_{k-1}(\theta)\xi_k^2\right] \\ = & \mathbb{E}\left[\left((1-u)\frac{1}{[nu]}\sum_{k=1}^{[nu]} + u\frac{1}{n-[nu]}\sum_{k=[nu]+1}^{n}\right) H_{k-1}(\theta)^\top H_{k-1}(\theta)\xi_k^2\right] \\ \to & (1-u)\mathrm{tr}C(\theta,\theta) + u\mathrm{tr}C(\theta,\theta) = \mathrm{tr}C(\theta,\theta) \end{aligned}$$

for all $u \in (0, 1)$ and it holds that

$$\mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n}(w_k^n(u))^2 H_{k-1}(\theta)^\top H_{k-1}(\theta)\xi_k^2\right] \leq \sup_{k=1,2,\ldots} \mathbb{E}\left[H_{k-1}(\theta)^\top H_{k-1}(\theta)\xi_k^2\right].$$

Next we argue the convergence of the inner product

$$\left\langle \frac{1}{\sqrt{n}}\sum_{k=1}^{n} w_k^n H_{k-1}(\theta)\xi_k, h\right\rangle_{L^2([0,1])} = \frac{1}{\sqrt{n}}\sum_{k=1}^{n}\langle w_k^n H_{k-1}(\theta), h\rangle_{L^2([0,1])}\xi_k.$$

As for the variance, the tower property yields that

$$\begin{aligned} V_n &:= \mathbb{E}\left[\left(\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\langle w_k^n H_{k-1}(\theta), h\rangle_{L^2([0,1])}\xi_k\right)^2\right] \\ &= \mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n}\langle w_k^n H_{k-1}(\theta), h\rangle_{L^2([0,1])}^2 \xi_k^2\right] \end{aligned}$$

and the right-hand side is equal to

$$\begin{aligned} & \mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n}\int\int w_k^n(u)w_k^n(v)h(u)^\top H_{k-1}(\theta)H_{k-1}(\theta)^\top h(v)dudv\xi_k^2\right] \\ = & \int\int\frac{1}{n}\sum_{k=1}^{n} w_k^n(u)w_k^n(v)h(u)^\top\mathbb{E}\left[H_{k-1}(\theta)H_{k-1}(\theta)^\top\xi_k^2\right]h(v)dudv \end{aligned}$$

We shall check that the dominated convergence theorem which yields that

$$V_n \to \int\int \frac{u \wedge v - uv}{\sqrt{u(1-u)v(1-v)}} h(u)^\top C(\theta,\theta) h(v) du dv$$

can be applied. The pointwise convergence holds by the assumption (4.2.7). By the Schwartz inequality, it holds that

$$\begin{aligned}
& \frac{1}{n}\sum_{k=1}^n w_k^n(u) w_k^n(v) h(u)^\top \mathbb{E}\left[H_{k-1}(\theta)H_{k-1}(\theta)^\top \xi_k^2\right] h(v) \\
\leq & \sqrt{\frac{1}{n^2}\sum_{k=1}^n (w_k^n(u)h(u)^\top \mathbb{E}\left[H_{k-1}(\theta)H_{k-1}(\theta)^\top \xi_k^2\right] h(v))^2 \sum_{k=1}^n (w_k^n(v))^2} \\
\leq & \sqrt{\frac{1}{n}\sum_{k=1}^n (w_k^n(u))^2 \sup_{k=1,2,\dots} (h(u)^\top \mathbb{E}\left[H_{k-1}(\theta)H_{k-1}(\theta)^\top \xi_k^2\right] h(v))^2} \\
= & \sup_{k=1,2,\dots} \left| h(u)^\top \mathbb{E}\left[H_{k-1}(\theta)H_{k-1}(\theta)^\top \xi_k^2\right] h(v) \right|.
\end{aligned}$$

The right-hand side is integrable by the Schwartz inequality for the Euclid inner product, which gives the following upper bound

$$\begin{aligned}
& \|h(u,\theta)\| \sup_{k=1,2,\dots} \left\| \mathbb{E}\left[H_{k-1}(\theta)H_{k-1}(\eta)^\top \xi_k^2\right] h(v,\eta) \right\| \\
\leq & \|h(u,\theta)\| \|h(v,\eta)\| \sup_{k=1,2,\dots} \left\| \mathbb{E}\left[H_{k-1}(\theta)H_{k-1}(\eta)^\top \xi_k^2\right] \right\|_{OP}
\end{aligned}$$

and by the assumption (4.2.9). Therefore, we can apply the dominated convergence theorem. It also leads that (2.2.2) holds. Finally, let us check the Lyapunov condition:

$$\mathbb{E}\left[\frac{1}{n^{(2+\delta_0)/2}} \sum_{k=1}^n \langle w_k^n H_s(\theta), h \rangle_{L^2([0,1])}^{2+\delta_0} \xi_k^2 \right] \to 0$$

for some $\delta_0 > 0$. The Schwartz inequality and the Jensen inequality give the upper bound of the left-hand side

$$
\begin{aligned}
& \frac{1}{n^{(2+\delta_0)/2}} \mathbb{E}\left[\sum_{k=1}^{n} \|w_k^n H_s(\theta)\|_{L^2([0,1])}^{2+\delta_0} \xi_k^2\right] \|h\|_{L^2([0,1])}^{2+\delta_0} \\
\leq \ & \frac{1}{n^{(2+\delta_0)/2}} \mathbb{E}\left[\sum_{k=1}^{n} \int \|w_k^n(u) H_{k-1}(\theta)\|^{2+\delta_0} du \xi_k^2\right] \|h\|_{L^2([0,1])}^{2+\delta_0} \\
= \ & \frac{1}{n^{(2+\delta_0)/2}} \int \sum_{k=1}^{n} |w_k^n(u)|^{2+\delta_0} \mathbb{E}\left[|H_{k-1}(\theta)^\top H_{k-1}(\theta)|^{1+\delta_0/2} \xi_k^2\right] dsdu \|h\|_{L^2([0,1])}^{2+\delta_0} \\
\leq \ & \frac{1}{n^{\delta_0/2}} \int_{\frac{1}{n}}^{1} W^n(u) du \sup_{k=1,2,\ldots} \mathbb{E}\left[|H_{k-1}(\theta)^\top H_{k-1}(\theta)|^{1+\delta_0/2} \xi_k^2\right] \|h\|_{L^2([0,1])}^{2+\delta_0}.
\end{aligned}
$$

If we choose $\delta_0 < 2$, which is possible by the assumption (4.1.1), the right-hand side converges to 0. Hence, the martingale central limit theorem yields the conclusion. □

Chapter 5

Continuous time stochastic processes

5.1 The limit distribution of Z-process

Let us describe the following problem. Firstly, set a measurable space and introduce a filtration. For a one dimensional continuous time adapted process $t \rightsquigarrow X_t,\ t \in [0, \infty)$ whose state space is $\mathbb{R}$, consider a parametric model $\{P_\theta\}$ indexed by $\theta \in \Theta$, where Θ is a bounded open subset of $\mathbb{R}^d$, such that under P_θ, X has the representation

$$X_t = A_t(\theta) + M_t(\theta), \quad \forall t \in [0, \infty)$$

where

$$t \rightsquigarrow A_t(\theta) = \int_0^t a_s(\theta) ds$$

is a predictable finite variation stochastic process and $M(\theta)$ is a martingale starting at zero whose predictable quadratic variation is

$$t \rightsquigarrow \langle M(\theta) \rangle_t = \int_0^t b_s(\theta) ds,$$

where suppose that

$$\sup_{s \in [0,\infty)} \mathbb{E}[b_s(\theta)] < \infty, \quad \forall \theta \in \Theta.$$

The true value of θ for X_t is denoted by $\theta_{(t)}$.

For the model above, we wish to test the following hypotheses:

$\mathcal{H}_0$: $\exists\theta_0 \in \Theta$ *such that* $\theta_{(t)} = \theta_0, \ \forall t \in [0, T]$
$\mathcal{H}_1$: $\exists\theta_0, \theta_1 \in \Theta, \ \exists u_* \in (0,1)$ *such that* $\theta_{(t)} = \theta_0, \ \forall t \in [0, Tu_*)$
and that $\theta_{(t)} = \theta_1 \neq \theta_0, \ \forall t \in [Tu_*, T]$

In order to estimate parameter θ, the following estimating equation is considered:

$$\Psi_T(\theta) = \frac{1}{T}\int_0^T H_s(\theta)(dX_s - a_s(\theta)ds) = 0,$$

where H is a d dimensional predictable process such that

$$\frac{1}{T}\int_0^T \|H_s(\theta)\|^2 b_s(\theta_{(s)})ds < \infty, \quad a.s., \quad \forall\theta \in \Theta.$$

The solution, or an approximate solution, of the estimating equation is denoted by $\hat{\theta}_T$, i.e. it holds that $\Psi_T(\hat{\theta}_T) = o_P(1)$, and used as an estimator for θ.

Assumptions I (C1) There exists the matrix $C_\kappa(\theta, \eta)$ given by

$$C_\kappa(\theta, \eta) = \underset{T\to\infty}{\text{l.i.m.}} \frac{1}{T}\int_0^T H_s(\theta)H_s(\eta)^\top b_s(\kappa)ds$$

and $C_{\theta_0}(\theta, \theta)$ is a positive definite for every $\theta \in \Theta$.
(C2) There exist the limits

$$\underset{T\to\infty}{\text{l.i.m.}} \frac{1}{T}\int_0^T H_s(\theta)(a_s(\theta) - a_s(\theta_0))ds,$$
$$\underset{T\to\infty}{\text{l.i.m.}} \frac{1}{T}\int_0^T H_s(\theta)\dot{a}_s(\theta_0)ds,$$

for any $\theta_0 \in \Theta$. Moreover, it holds that for any $\theta_0 \in \Theta$ and any $\varepsilon > 0$,

$$\inf_{\theta:\|\theta-\theta_0\|>\varepsilon} \left\| \underset{T\to\infty}{\text{l.i.m.}} \frac{1}{T}\int_0^T H_s(\theta)(a_s(\theta) - a_s(\theta_0))ds \right\| > 0. \tag{5.1.1}$$

(C3) Under $\mathcal{H}_0$, $\sqrt{T}(\hat{\theta}_T - \theta_0) \to^d N(0, C_{\theta_0}(\theta_0, \theta_0)^{-1})$
(C4) Under $\mathcal{H}_1$, $\hat{\theta}_T \to^p \theta_*$ which satisfies

$$\begin{aligned} & u_* \left(\underset{T\to\infty}{\text{l.i.m.}} \frac{1}{Tu_*} \int_0^{Tu_*} H_s(\theta_*)(a_s(\theta_0) - a_s(\theta_*))ds \right) \\ & +(1-u_*) \left(\underset{T\to\infty}{\text{l.i.m.}} \frac{1}{T(1-u_*)} \int_{Tu_*}^{T} H_s(\theta_*)(a_s(\theta_1) - a_s(\theta_*))ds \right) \\ = \; & u_* D^\infty(\theta_0) + (1-u_*) D^\infty(\theta_1) = 0, \end{aligned}$$

where

$$D^\infty(\theta) = \underset{T\to\infty}{\text{l.i.m.}} \frac{1}{T} \int_0^T H_s(\theta_*)(a_s(\theta) - a_s(\theta_*))ds.$$

(C5) $H_s(\theta)$ is continuously differentiable with respect to θ.
(C6) $a_s(\theta)$ is second order continuously differentiable with respect to θ.
(C7) Under $\mathcal{H}_1$, it holds that

$$\sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta_*)\|^2 a_s^2(\theta)\right] < \infty$$

for $\theta \in \{\theta_0, \theta_1, \theta_*\}$.
(C8) Under $\mathcal{H}_0$, it holds that

$$\sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta_0)\|^2 (\partial_i a_s(\theta_0))^2\right] < \infty$$

for $i = 1, \ldots, d$. Under $\mathcal{H}_1$, it holds that

$$\sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta_*)\|^2 (\partial_i a_s(\theta_*))^2\right] < \infty$$

for $i = 1, \ldots, d$.
(C9) Under $\mathcal{H}_0$, it holds that

$$\sup_{s\in[0,\infty)} \mathbb{E}[\|H_s(\theta_0)\|^{2+\delta} b_s(\theta_0)] < \infty$$

for some $\delta > 0$. Under $\mathcal{H}_1$, it holds that

$$\sup_{s\in[0,\infty)} \mathbb{E}[\|H_s(\theta_*)\|^{2+\delta} b_s(\theta_{(s)})] < \infty$$

for some $\delta > 0$.
(C10) Under $\mathcal{H}_0$, it holds that

$$\sup_{s\in[0,\infty)} \left\|\mathbb{E}[H_s(\theta_0)H_s(\theta_0)^\top b_s(\theta_0)]\right\|_{OP} < \infty.$$

Under $\mathcal{H}_1$, it holds that

$$\sup_{s\in[0,\infty)} \left\|\mathbb{E}[H_s(\theta_*)H_s(\theta_*)^\top b_s(\theta_{(s)})]\right\|_{OP} < \infty.$$

(C11) It holds that

$$\left\| \frac{1}{\sqrt{T}} \int_0^T w_s^T H_s(\theta) dM_s|_{\theta=\theta_n} - \frac{1}{\sqrt{T}} \int_0^T w_s^T H_s(\theta_0) dM_s \right\|_{L^2([0,1])} \to^p 0$$

for any $\theta_0 \in \Theta$, where θ_n is a random sequence which satisfies $\theta_n - \theta_0 = O_P(1/\sqrt{T})$.

Remark 5.1.1. *If $\{M_\cdot\}$ is finite variation processes, (C11) follows from the Lipschitz conditions for $\theta \rightsquigarrow H_s(\theta)$. In the case of ergodic diffusion processes, (C11) also becomes the Lipschitz conditions for $\theta \rightsquigarrow H_s(\theta)$ by the use of The Itô formula under some mild conditions under $\mathcal{H}_0$. See the next section.*

Assumptions I-S Assume (C1)-(C8) and (C11).
(C12) Under $\mathcal{H}_0$, it holds that

$$\sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta_0)\|^4 (b_s(\theta_0))^2\right] < \infty.$$

Under $\mathcal{H}_1$, it holds that

$$\sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta_*)\|^4 (b_s(\theta_{(s)}))^2\right] < \infty.$$

(C13) Under $\mathcal{H}_0$, it holds that

$$\sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta_0)\|^4\right] < \infty.$$

Under $\mathcal{H}_1$, it holds that

$$\sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta_*)\|^4\right] < \infty.$$

Proposition 5.1.1. *Assumptions Ia-S is a sufficient condition for Assumptions Ia. Especially,*
(C12) implies (C10). (C12) and (C13) imply (C9).

Proof of the Proposition 5.1.1 As for the first assertion, it follows from the property of the operator norm and the Jensen inequality that

$$
\begin{aligned}
& \sup_{s\in[0,\infty)} \left\| \mathbb{E}\left[H_s(\theta) H_s(\theta)^\top b_s(\theta_{(s)}) \right] \right\|_{OP}^2 \\
\leq \; & \sup_{s\in[0,\infty)} \sum_{i=1}^{d} \sum_{j=1}^{d} \left| \mathbb{E}\left[\left(H_s(\theta)\right)_{(i)} \left(H_s(\theta)\right)_{(j)} b_s(\theta_{(s)}) \right] \right|^2 \\
\leq \; & \sup_{s\in[0,\infty)} \sum_{i=1}^{d} \sum_{j=1}^{d} \mathbb{E}\left[\left(H_s(\theta)\right)_{(i)}^2 \left(H_s(\theta)\right)_{(j)}^2 \left(b_s(\theta_{(s)})\right)^2 \right] \\
= \; & \sup_{s\in[0,\infty)} \mathbb{E}\left[\sum_{i=1}^{d} \left(H_s(\theta)\right)_{(i)}^2 \sum_{j=1}^{d} \left(H_s(\theta)\right)_{(j)}^2 \left(b_s(\theta_{(s)})\right)^2 \right] \\
\leq \; & \sup_{s\in[0,\infty)} \mathbb{E}\left[\left(\|H_s(\theta)\|^4 + \|H_s(\theta)\|^4 \right) \left(b_s(\theta_{(s)})\right)^2 \right] \\
= \; & 2 \sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta)\|^4 \left(b_s(\theta_{(s)})\right)^2 \right] < \infty.
\end{aligned}
$$

As for the second assertion, it follows from the same reason as the Proposition 4.1.1. This completes the proof. □

Introduce the random field $\{\mathbb{Z}_T(u,\theta); (u,\theta) \in (0,1) \times \Theta\}$ given by

$$
\mathbb{Z}_T(u,\theta) = \frac{1}{\sqrt{T}} \int_0^T w_s^T(u) H_s(\theta)(dX_s - a_s(\theta)ds),
$$

where

$$
w_s^T(u) = \frac{1\{s \leq Tu\} - u}{\sqrt{u(1-u)}}, \quad u \in (0,1).
$$

Its "predictable projection" to the true model is

$$
\mathbb{Z}_T^p(u,\theta) = \frac{1}{\sqrt{T}} \int_0^T w_s^T(u) H_s(\theta)(a_s(\theta_{(s)}) - a_s(\theta))ds.
$$

The difference between $\mathbb{Z}$ and $\mathbb{Z}^p$, which is a martingale random field, is denoted by $\{\mathbb{M}_T(u,\theta); (u,\theta) \in (0,1) \times \Theta\}$, say, it is given by

$$\mathbb{M}_T(u,\theta) = \mathbb{Z}_T(u,\theta) - \mathbb{Z}_T^p(u,\theta) = \frac{1}{\sqrt{T}} \int_0^T w_s^T(u) H_s(\theta) dM_s^{\theta(s)}.$$

Under $\mathcal{H}_0$, it holds that

$$\mathbb{Z}_T^p(\cdot,\theta_0) = 0,$$

so

$$\mathbb{M}_T(\cdot,\theta_0) = \mathbb{Z}_T(\cdot,\theta_0).$$

This relationship gives us the idea to use functions of $\mathbb{Z}_T$ as a test statistic. However, since we cannot know the true value θ_0, it is crucial to hold that under $\mathcal{H}_0$,

$$\mathbb{Z}_T(\cdot,\hat{\theta}_T) - \mathbb{Z}_T(\cdot,\theta_0) \to^p 0$$

which enables us to apply the limit theorem in the preceding chapter. Moreover, in order to ensure the power of the test, it is crucial to hold that under $\mathcal{H}_1$,

$$\frac{1}{\sqrt{T}} \mathbb{Z}_T^p(\cdot,\hat{\theta}_T) \not\to^p 0.$$

Lemma 5.1.1. *Under $\mathcal{H}_0$, it holds that*

$$\left\| \mathbb{Z}_T^p(\cdot,\hat{\theta}_T) \right\|_{L^2}^2 \to^p 0$$

as $T \to \infty$.

Proof of the Lemma 5.1.1 The Taylor expansion yields that

$$\begin{aligned} \mathbb{Z}_T^p(u,\hat{\theta}_T) &= \frac{1}{\sqrt{T}} \int_0^T w_s^T(u) H_s(\hat{\theta}_T)(a_s(\theta_0) - a_s(\hat{\theta}_T)) ds \\ &= \frac{1}{T} \int_0^T w_s^T(u) H_s(\hat{\theta}_T) \dot{a}_s(\tilde{\theta}_T)^\top ds \sqrt{T}(\hat{\theta}_T - \theta_0), \end{aligned}$$

where $\tilde{\theta}_T$ is a value between θ_0 and $\hat{\theta}_T$. Because of the assumption $\sqrt{T}(\hat{\theta}_T - \theta_0) = O_P(1)$, we argue the convergence to 0 in probability in $L^2([0,1],du)$ of the all elements in the following matrix:

$$
\begin{aligned}
& \frac{1}{T}\int_0^T w_s^T(\cdot)H_s(\hat{\theta}_T)\dot{a}_s(\tilde{\theta}_T)^\top ds \\
= & \frac{1}{T}\int_0^T w_s^T(\cdot)H_s(\theta_0)\dot{a}_s(\theta_0)^\top ds \\
& +\frac{1}{T}\int_0^T w_s^T(\cdot)H_s(\theta_0)(\dot{a}_s(\tilde{\theta}_T)-\dot{a}_s(\theta_0))^\top ds \\
& +\frac{1}{T}\int_0^T w_s^T(\cdot)(H_s(\hat{\theta}_T)-H_s(\theta_0))\dot{a}_s(\theta_0)^\top ds.
\end{aligned}
$$

For our purpose, it is sufficient to prove the first term in the right-hand side converges to 0 in $L^2([0,1],du)$ because of the continuous differentiability of $\dot{a}_s(\theta)$ and $H_s(\theta)$ with respect to θ, $\hat{\theta}_T-\theta_0=o_P(1)$ and the Schwartz inequality. The terms in the right-hand side converge to 0 in the second mean for all u because of the assumption (C2). Moreover, by the Schwartz inequality and assumption (C8), it holds that

$$
\begin{aligned}
& \mathbb{E}\left[\left\|\frac{1}{T}\int_0^T w_s^T(u)H_s(\theta_0)\partial_i a_s(\theta_0)ds\right\|^2\right] \\
\leq & \mathbb{E}\left[\frac{1}{T}\int_0^T (w_s^T(u))^2 ds \frac{1}{T}\int_0^T H_s(\theta_0)^\top H_s(\theta_0)\left(\partial_i a_s(\theta_0)\right)^2 ds\right] \\
\leq & \sup_{s\in[0,\infty)} \mathbb{E}\left[H_s(\theta_0)^\top H_s(\theta_0)\left(\partial_i a_s(\theta_0)\right)^2\right] < \infty,
\end{aligned}
$$

for all $i=1,\ldots,d$ and $u\in(0,1)$. Hence, the Fubini theorem and the bounded convergence theorem yield that

$$
\begin{aligned}
& \lim_{T\to\infty}\mathbb{E}\left[\int_0^1\left\|\frac{1}{T}\int_0^T w_s^T(u)H_s(\theta_0)\partial_i a_s(\theta_0)ds\right\|^2 du\right] \\
= & \lim_{T\to\infty}\int_0^1 \mathbb{E}\left[\left\|\frac{1}{T}\int_0^T w_s^T(u)H_s(\theta_0)\partial_i a_s(\theta_0)ds\right\|^2\right] du \\
= & \int_0^1 \lim_{T\to\infty}\mathbb{E}\left[\left\|\frac{1}{T}\int_0^T w_s^T(u)H_s(\theta_0)\partial_i a_s(\theta_0)ds\right\|^2\right] du = 0
\end{aligned}
$$

for all $i=1,\ldots,d$. This completes the proof. □

Next, we discuss a limit theorem for $\mathbb{M}_T(\cdot,\theta_0)$ which is taking values in $L^2([0,1],du)$ spaces, which is a consequence of Theorem 4.1.2.

Lemma 5.1.2. *Under $\mathcal{H}_0$, it holds that the random field $u \rightsquigarrow \mathbb{M}_T(u,\theta_0)$ converges weakly to $u \rightsquigarrow C_{\theta_0}(\theta_0,\theta_0)^{1/2}B_d^\circ(u)/\sqrt{u(1-u)}$ in $L^2([0,1],du)$, where B_d° is the d dimensional standard Brownian bridge.*

Now, let us propose a test statistic AD_T defined by

$$\begin{aligned} AD_T &= \int_0^1 \mathbb{Z}_T(u,\hat{\theta}_T)^\top \hat{C}_T^{-1}\mathbb{Z}_T(u,\hat{\theta}_T)du \\ &= \left\| \hat{C}_T^{-1/2}\mathbb{Z}_T(\cdot,\hat{\theta}_T)\right\|_{L^2}^2, \end{aligned}$$

where $\hat{C}_T$ is a consistent estimator for $C_{\theta_0}(\theta_0,\theta_0)$ under $\mathcal{H}_0$. By the condition (C11), the preceding lemma, the Slutsky theorem and the continuous mapping theorem yield the former assertion of the following theorem.

Theorem 5.1.1. *(i) Under $\mathcal{H}_0$, it holds that*

$$AD_T \to^d \|G\|_{L^2}^2$$

as $T\to\infty$, where $u \rightsquigarrow G(u) = B_d^\circ(u)/\sqrt{u(1-u)}$.
(ii) Under $\mathcal{H}_1$, it holds that

$$AD_T \geq T\left(\frac{2}{3}\Delta^\top C_*^{-1}\Delta - o_P(1)\right) + O_P(1),$$

where $\Delta = u_*(1-u_*)(D^\infty(\theta_0) - D^\infty(\theta_1))$ and the test is consistent.

Proof of the Theorem 5.1.1.(ii). Since $\hat{C}_T^{-1}$ is non-negative definite matrix, it holds that

$$2v_1^\top \hat{C}_T^{-1}v_1 + 2v_2^\top \hat{C}_T^{-1}v_2 \geq (v_1-v_2)^\top \hat{C}_T^{-1}(v_1-v_2)$$

for arbitrary d-dimensional vector v_1, v_2. This property and the inequality $\sqrt{u(1-u)} \leq 1/2$ yield that

$$\begin{aligned} AD_T &\geq \frac{1}{2}\int_0^1 (Z_T^p)^\top(u,\hat{\theta}_T)\hat{C}_T^{-1}Z_T^p(u,\hat{\theta}_T)du - \int_0^1 \mathbb{M}_T(u,\hat{\theta}_T)^\top \hat{C}_T^{-1}\mathbb{M}_T(u,\hat{\theta}_T)du \\ &\geq 2T\int_0^1 A_T^\top(u)\hat{C}_T^{-1}A_T(u)du - \int_0^1 \mathbb{M}_T(u,\hat{\theta}_T)^\top \hat{C}_T^{-1}\mathbb{M}_T(u,\hat{\theta}_T)du, \end{aligned}$$

where

$$\begin{aligned}
A_T(u) &= \frac{\sqrt{u(1-u)}}{\sqrt{T}}\mathbb{Z}_T^p(u,\hat{\theta}_T) \\
&= \frac{1}{T}\int_0^T (1\{s \le Tu\} - u)H_s(\hat{\theta}_T)(a_s(\theta_0)1\{s \le Tu_*\} \\
&\quad +a_s(\theta_1)1\{s \ge Tu_*\} - a_s(\hat{\theta}_T))ds \\
&= \frac{1}{T}\int_0^T (1\{s \le Tu\} - u)H_s(\theta_*)(a_s(\theta_0)1\{s \le Tu_*\} \\
&\quad +a_s(\theta_1)1\{s \ge Tu_*\} - a_s(\theta_*))ds + o_P(1) \\
&= \tilde{A}_T(u) + o_P(1), \quad (say).
\end{aligned}$$

The third equality is obtained by the same reason as Lemma 5.1.1. It holds that, for $u \le u_*$,

$$\begin{aligned}
\tilde{A}_T(u) &= \frac{1-u}{T}\int_0^{Tu} H_s(\theta_*)(a_s(\theta_0) - a_s(\theta_*))ds \\
&\quad -\frac{u}{T}\int_{Tu}^{Tu_*} H_s(\theta_*)(a_s(\theta_0) - a_s(\theta_*))ds \\
&\quad -\frac{u}{T}\int_{Tu_*}^{T} H_s(\theta_*)(a_s(\theta_1) - a_s(\theta_*))ds
\end{aligned}$$

so,

$$\begin{aligned}
\underset{T\to\infty}{\text{l.i.m.}}\tilde{A}_T(u) &= (u(1-u) - u(u_* - u))D^\infty(\theta_0) - u(1-u_*)D^\infty(\theta_1) \\
&= u(1-u_*)(D^\infty(\theta_0) - D^\infty(\theta_1)),
\end{aligned}$$

and for $u > u_*$,

$$\underset{T\to\infty}{\text{l.i.m.}}\tilde{A}_T(u) = u_*(1-u)(D^\infty(\theta_0) - D^\infty(\theta_1)).$$

Let us denote $\underset{T\to\infty}{\text{l.i.m.}}\tilde{A}_T(u)$ by $A_\infty(u)$ for all $u \in (0,1)$. Next, we shall prove

$$\mathbb{E}\left[\|\tilde{A}_T - A_\infty\|_{L^2}^2\right] \to 0. \tag{5.1.2}$$

It holds that for all u,

$$\mathbb{E}\left[\left(\tilde{A}_T(u) - A_\infty(u)\right)^2\right] \le 2\mathbb{E}\left[(\tilde{A}_T(u))^2\right] + 2(A_\infty(u))^2$$

and the first term in the right-hand side is bounded above by

$$\begin{aligned}
& 2\mathbb{E}\left[\frac{1}{T}\int_0^T (1\{s \le Tu\} - u)^2 H_s(\theta_*)^\top H_s(\theta_*)(a_s(\theta_0)1\{s \le Tu_*\}\right. \\
& \quad \left. + a_s(\theta_1)1\{s \ge Tu_*\} - a_s(\theta_*))^2 ds\right] \\
\le\ & \frac{2}{T}\int_0^T (1\{s \le Tu\} - u)^2 ds \sup_{s\in[0,\infty)} \mathbb{E}\left[H_s(\theta_*)^\top H_s(\theta_*)(a_s(\theta_0)1\{s \le Tu_*\}\right. \\
& \quad \left. + a_s(\theta_1)1\{s \ge Tu_*\} - a_s(\theta_*))^2\right] \\
\le\ & 2 \sup_{s\in[0,\infty)} \mathbb{E}\left[H_s(\theta_*)^\top H_s(\theta_*)(a_s(\theta_0)^2 + a_s(\theta_1)^2 + a_s(\theta_*)^2)\right] < \infty.
\end{aligned}$$

Since the left-hand side of (5.1.2) is equal to

$$\int_0^1 \mathbb{E}\left[\left(\tilde{A}_T(u) - A_\infty(u)\right)^2\right] du$$

and $(A_\infty(u))^2$ is integrable with respect to u, the dominated convergence theorem yields (5.1.2), and (5.1.2) yields that $\tilde{A}_T \to^p A_\infty$ in $L^2([0,1], du)$. This result, the Slutsky theorem and the continuous mapping theorem yields that

$$\int_0^1 A_T^\top(u)\hat{C}_T^{-1}A_T(u)du \to^p \int_0^1 A_\infty^\top(u)C_*^{-1}A_\infty(u)du,$$

where $C_* := u_*C_{\theta_0}(\theta_*, \theta_*) + (1-u_*)C_{\theta_1}(\theta_*, \theta_*)$. By simple calculations, the right-hand side is equal to

$$\frac{u_*^2(1-u_*)^2}{3}(D^\infty(\theta_0) - D^\infty(\theta_1))^\top C_*^{-1}(D^\infty(\theta_0) - D^\infty(\theta_1)).$$

Finally, $\mathbb{M}_T(\cdot, \hat{\theta}_T)$ is asymptotically tight in $L^2([0,1], du)$ by the assumption (C11) and the Theorem 4.1.2. The last assertion is followed since C_* is positive definite and the assumption (5.1.1). This completes the proof. □

5.2 A change detection procedure for an ergodic diffusion process

Let us consider the stochastic differential equations with the state space $I = (l, r)$, where $-\infty \le l < r \le \infty$, given by

$$t \rightsquigarrow X_t = X_0 + \int_0^T S(X_s, \theta)ds + \int_0^T \sigma(X_s)dW_s, \qquad (5.2.1)$$

where $W.$ is the sandard Brownian motion and X_0 is a random variable that is independent of $W.$ and satisfies $\mathbb{E}[(X_0)^2] < \infty$. Suppose that there exist a strong solution to this SDE, that

$$\sup_{s\in[0,\infty)} \mathbb{E}[\sigma(X_s)^2] < \infty$$

and that $X.$ is ergodic in the second mean with respect to an invariant measure μ_θ for some θ, that is for any μ_θ-integrable function f, it holds that

$$\lim_{T\to\infty} \mathbb{E}\left[\left\|\frac{1}{T}\int_0^T f(X_s)ds - \int_I f(x)\mu_\theta(dx)\right\|^2\right] = 0.$$

Remark 5.2.1. *Some works consider test procedures to detect some changes in drift parameters of diffusion processes: Lee et al. (2006), Mihalache (2012), Negri and Nishiyama (2012) and Dehling et al. (2014). These previous works assume the ergodicity which guarantees the convergence in probability, so this assumption is stronger than theirs.*

Suppose that $S(x, \theta)$ is continuously differentiable with respect to θ. Consider the estimating equation

$$\Psi_T(\theta) = \frac{1}{T}\int_0^T \frac{\dot{S}(X_s, \theta)}{\sigma(X_s)^2}(dX_s - S(X_s, \theta)ds) = 0.$$

The solution of the above estimating equation is denoted by $\hat{\theta}_T$. In this case, $a_s(\theta) = S(X_s, \theta), b_s(\theta) = \sigma(X_s)^2$ and

$$H_s(\theta) = \frac{\dot{S}(X_s, \theta)}{\sigma(X_s)^2}.$$

The matrix $C_\kappa(\theta, \eta)$ is

$$C_\kappa(\theta, \eta) = \int_I \frac{\dot{S}(x, \theta)\dot{S}(x, \eta)^\top}{\sigma(x)^2}\mu_\kappa(dx).$$

Suppose the following conditions:
(I) The function $(x, \theta) \mapsto S(x, \theta)$ is continuously differentiable with respect to x and third order continuously differentiable with respect to θ in the neighborhood N of $\theta_0, \theta_1, \theta_*$ and the order of derivative is exchangeable. The function

$x \mapsto \sigma(x)$ is continuously differentiable with respect to x. The functions $\sup_{\theta\in N}|S(x,\theta)|$, $\sup_{\theta\in N}|\partial_i S(x,\theta)|$, $\sup_{\theta\in N}|\partial_{ij} S(x,\theta)|$, $\sup_{\theta\in N}|\partial_{ijk} S(x;\theta)|$, $\sigma(x)$ and $\sigma'(x)$ are bounded above by polynomial growth functions of x, that is, for example, it holds that

$$\sup_{\theta\in N}|S(x,\theta)| \leq C(1+|x|^p)$$

for some constants $C, p \geq 1$.
(II) $\inf_{x\in\mathbb{R}} \sigma(x) > 0$.
(III) For arbitrary $q \geq 1$, $\sup_{s\in[0,\infty)} \mathbb{E}\left[|X_s|^q\right] < \infty$.
(IV) Define

$$\Psi(\theta,\kappa) = \int_I \frac{(S(x,\kappa) - S(x,\theta))\dot{S}(x,\theta)}{\sigma(x)^2}\mu_\kappa(dx).$$

For all $\kappa \in \Theta$ and any $\varepsilon > 0$, $\inf_{\theta:\|\theta-\kappa\|>\varepsilon} \|\Psi(\theta,\kappa)\| > 0$ holds.
(V) The matrix $C_\kappa(\theta,\theta)$ is regular for all $\theta,\kappa \in \Theta$.
(VI) Define $x \mapsto K(x), K_d(x)$ by

$$K(\cdot) = \max_{i,j,k}\sup_{\theta\in N} \partial_i\partial_j\partial_k S(\cdot,\theta),$$
$$K_d(\cdot) = \max_{i,j,k}\sup_{\theta\in N} \partial_i\partial_j\partial_k S'(\cdot,\theta),$$

where N is the neighborhood of any θ_0. The function $K(x)$ is continuously differentiable with respect to x. The functions $|K(x)|$ and $|K_d(x)|$ are bounded above by polynomial growth functions of x.

Remark 5.2.2. *The conditions (I)-(V) are standard in order to argue inferences on ergodic diffusion processes. See Nishiyama (2011) in Japanese. See also Kutoyants (2004) for general studies on statistical inferences for ergodic diffusion processes. Note that, in the current case, since we just consider continuous processes, (C9) with $\delta = 0$ is enough.*

Proposition 5.2.1. *Assume conditions (I)-(VI). (i) Under $\mathcal{H}_0$, it holds that $\sqrt{T}(\hat{\theta}_T - \theta_0) \to^d N(0, I^{-1})$.*
(ii) Under $\mathcal{H}_1$, it holds that $\hat{\theta}_T \to^p \theta_*$, which satisfies

$$u_*\Psi(\theta_*,\theta_0) + (1-u_*)\Psi(\theta_*,\theta_1) = 0.$$

(iii) (C8) holds, that is, for $\theta = \theta_0, \theta_*$,

$$\sup_{s\in[0,\infty)} \mathbb{E}\left[\frac{\|\dot{S}(X_s,\theta)\|^4}{\sigma(X_s)^4}\right] < \infty.$$

It yields that (C9) holds for $\delta = 1$.
(iv) (C7) holds, that is, for $\theta = \theta_0, \theta_1, \theta_*$,

$$\sup_{s\in[0,\infty)} \mathbb{E}\left[\frac{\|\dot{S}(X_s,\theta_*)\|^2 (S(X_s,\theta))^2}{\sigma(X_s)^4}\right] < \infty.$$

(v) (C10) holds, that is, for $\theta = \theta_0, \theta_*$

$$\sup_{s\in[0,\infty)} \left\|\mathbb{E}\left[\frac{\dot{S}(X_s,\theta)\dot{S}(X_s,\theta)^\top}{\sigma(X_s)^2}\right]\right\|_{OP}^2 < \infty.$$

As for the proofs of (i) and (ii), see Nishiyama (2011). As for the proof of (iv), the proof is the same as (iii), essentially. Here, let us prove the rest.

Proof of the Proposition 5.2.1. Fix θ to θ_0 or θ_*.
(iii). By the assumptions, there exist constants $C, p \geq 1$ such that

$$\begin{aligned}
\sup_{s\in[0,\infty)} \mathbb{E}\left[\frac{\|\dot{S}(X_s,\theta)\|^4}{\sigma(X_s)^4}\right] &= \sup_{s\in[0,\infty)} \mathbb{E}\left[\frac{\left(\sum_{i=1}^d (\partial_i S(X_s,\theta))^2\right)^2}{\sigma(X_s)^4}\right] \\
&\leq \sup_{s\in[0,\infty)} \mathbb{E}\left[\frac{d\sum_{i=1}^d (\partial_i S(X_s,\theta))^4}{\sigma(X_s)^4}\right] \\
&\leq \sup_{s\in[0,\infty)} \mathbb{E}\left[d\sum_{i=1}^d \frac{\sup_{\theta\in N} |\partial_i S(X_s,\theta)|^4}{\inf_{x\in\mathbb{R}} \sigma(x)^4}\right] \\
&\leq \sup_{s\in[0,\infty)} \mathbb{E}\left[d\sum_{i=1}^d \frac{|C(1+|X_s|)^p|^4}{\inf_{x\in\mathbb{R}} \sigma(x)^4}\right] \\
&= \frac{C^4 d^2}{\inf_{x\in\mathbb{R}} \sigma(x)^4} \sup_{s\in[0,\infty)} \mathbb{E}\left[|1+|X_s||^{4p}\right] < \infty.
\end{aligned}$$

The latter assertion is obvious. This completes the proof.
(v). It follows from the definition of the operator norm that

$$\begin{aligned}
&\sup_{s\in[0,\infty)}\left\|\mathbb{E}\left[\frac{\dot{S}(X_s,\theta)\dot{S}(X_s,\theta)^\top}{\sigma(X_s)^2}\right]\right\|_{OP} \\
\leq\ & \frac{1}{\inf_{x\in\mathbb{R}}\sigma(x)^2}\sup_{s\in[0,\infty)}\left\|\mathbb{E}\left[\dot{S}(X_s,\theta)\dot{S}(X_s,\theta)^\top\right]\right\|_{OP} \\
=\ & \frac{1}{\inf_{x\in\mathbb{R}}\sigma(x)^2}\sup_{s\in[0,\infty)}\max_{i=1,\dots,d}\sigma_i\left(\mathbb{E}\left[\dot{S}(X_s,\theta)\dot{S}(X_s,\theta)^\top\right]\right),
\end{aligned}$$

where $\sigma_i(A)$ denotes the i-th singular value of the matrix A. By the assumptions, there exist constants $C, p \geq 1$ such that

$$\begin{aligned}
&\sup_{s\in[0,\infty)}\max_{i=1,\dots,d}\sigma_i\left(\mathbb{E}\left[\dot{S}(X_s,\theta)\dot{S}(X_s,\theta)^\top\right]\right) \\
\leq\ & \sup_{s\in[0,\infty)}\sqrt{\sum_{i=1}^d\left(\sigma_i\left(\mathbb{E}\left[\dot{S}(X_s,\theta)\dot{S}(X_s,\theta)^\top\right]\right)\right)^2} \\
=\ & \sup_{s\in[0,\infty)}\sqrt{\sum_{i=1}^d\sum_{j=1}^d\left(\mathbb{E}\left[\partial_i S(X_s,\theta)\partial_j S(X_s,\theta)\right]\right)^2} \\
\leq\ & \sup_{s\in[0,\infty)}\sqrt{\sum_{i=1}^d\sum_{j=1}^d\left(\mathbb{E}\left[\sup_{\theta\in N}\partial_i S(X_s,\theta)\sup_{\theta\in N}\partial_j S(X_s,\theta)\right]\right)^2} \\
\leq\ & \sup_{s\in[0,\infty)}\sqrt{\sum_{i=1}^d\sum_{j=1}^d\left(\mathbb{E}\left[C(1+|X_s|)^p\right]\right)^2} \\
=\ & \sup_{s\in[0,\infty)} Cd\mathbb{E}\left[(1+|X_s|)^p\right] < \infty.
\end{aligned}$$

This completes the proof. □

Proposition 5.2.2. *Let $\hat{\theta}_n$ is a random sequence such that $\sqrt{T}(\hat{\theta}_T - \vartheta) = O_P(1)$ for any $\vartheta \in \Theta$. Under $\mathcal{H}_0$, it holds that*

$$\left\|\frac{1}{\sqrt{T}}\int_0^T w_s^T\frac{\dot{S}(X_s,\theta)}{\sigma(X_s)}dW_s|_{\theta=\hat{\theta}_T} - \frac{1}{\sqrt{T}}\int_0^T w_s^T\frac{\dot{S}(X_s,\vartheta)}{\sigma(X_s)}dW_s\right\|_{L^2([0,1])} \to^p 0.$$

Remark 5.2.3. *Let us confirm the Itô formula which will be frequently used in the following proof. Let $s \rightsquigarrow X_s$ be a one dimensional continuous semi-martingale whose quadratic variation process is denoted by $s \rightsquigarrow \langle X \rangle_s$. The map $x \mapsto f(x)$ is second order continuously differentiable and its first and second derivative is denoted by f', f'', respectively. It holds that*

$$\int_{X_0}^{X_T} f'(x)dx = f(X_T) - f(X_0) = \int_0^T f'(X_s)dX_s + \frac{1}{2}\int_0^T f''(X_s)d\langle X \rangle_s.$$

Especially, when we consider the stochastic differential equation

$$X_t = X_0 + \int_0^t S(X_s, \theta_0)ds + \int_0^t \sigma(X_s)dW_s,$$

puitting $f'(\cdot) = g(\cdot)/\sigma^2(\cdot)$, it holds that

$$\begin{aligned}
&\int_{X_0}^{X_T} \frac{g(x)}{\sigma^2(x)}dx \\
= \;& \int_0^T \frac{g(X_s)}{\sigma(X_s)}dW_s + \int_0^T \left(\frac{g(X_s)S(X_s,\theta_0)}{\sigma^2(X_s)} + \frac{g'(X_s)}{2} - \frac{\sigma'(X_s)g(X_s)}{\sigma(X_s)} \right) ds.
\end{aligned}$$

We use $\dot{S}, \ddot{S}, K$ as g.

Proof of the Proposition 5.2.2. It follows from the Itô formula that

$$\begin{aligned}
&\int_0^{Tu} \frac{\dot{S}(X_s,\theta)}{\sigma(X_s)}dW_s \\
= &\int_{X_0}^{X_{Tu}} \frac{\dot{S}(x,\theta)}{\sigma^2(x)}dx - \int_0^{Tu} \left(\frac{\dot{S}(X_s,\theta)S(X_s,\theta_0)}{\sigma^2(X_s)} + \frac{\dot{S}'(X_s,\theta)}{2} - \frac{\sigma'(X_s)\dot{S}(X_s,\theta)}{\sigma(X_s)} \right) ds
\end{aligned}$$

and that

$$\begin{aligned}
&\int_{Tu}^{T} \frac{\dot{S}(X_s,\theta)}{\sigma(X_s)}dW_s \\
= &\int_{X_{Tu}}^{X_T} \frac{\dot{S}(x,\theta)}{\sigma^2(x)}dx - \int_{Tu}^{T} \left(\frac{\dot{S}(X_s,\theta)S(X_s,\theta_0)}{\sigma^2(X_s)} + \frac{\dot{S}'(X_s,\theta)}{2} - \frac{\sigma'(X_s)\dot{S}(X_s,\theta)}{\sigma(X_s)} \right) ds.
\end{aligned}$$

Noting that

$$\mathbb{M}(u,\theta) = \frac{1}{\sqrt{Tu(1-u)}} \left((1-u)\int_0^{Tu} \frac{\dot{S}(X_s,\theta)}{\sigma(X_s)}dW_s - u\int_{Tu}^{T} \frac{\dot{S}(X_s,\theta)}{\sigma(X_s)}dW_s \right),$$

the Taylor expansion around ϑ and the triangle inequality yield that

$$\begin{aligned}
&\|\mathbb{M}(u,\hat{\theta}_T)-\mathbb{M}(u,\vartheta)\| \qquad (5.2.2)\\
\leq\ & \Bigg\|(1-u)\int_{X_0}^{X_{Tu}}\frac{\ddot{S}(x,\tilde{\theta}_T)}{\sigma^2(x)}dx-u\int_{X_{Tu}}^{X_T}\frac{\ddot{S}(x,\tilde{\theta}_T)}{\sigma^2(x)}dx\\
&-(1-u)\int_0^{Tu}\left(\frac{\ddot{S}(X_s,\tilde{\theta}_T)S(X_s,\theta_0)}{\sigma^2(X_s)}+\frac{\ddot{S}'(X_s,\check{\theta}_T)}{2}-\frac{\sigma'(X_s)\ddot{S}(X_s,\tilde{\theta}_T)}{\sigma(X_s)}\right)ds\\
&+u\int_{Tu}^{T}\left(\frac{\ddot{S}(X_s,\tilde{\theta}_T)S(X_s,\theta_0)}{\sigma^2(X_s)}+\frac{\ddot{S}'(X_s,\check{\theta}_T)}{2}-\frac{\sigma'(X_s)\ddot{S}(X_s,\tilde{\theta}_T)}{\sigma(X_s)}\right)ds\Bigg\|\frac{\|\hat{\theta}_T-\vartheta\|}{\sqrt{T(u(1-u))}}\\
\leq\ & \Bigg\|(1-u)\int_{X_0}^{X_{Tu}}\frac{\ddot{S}(x,\tilde{\theta}_T)}{\sigma^2(x)}dx-u\int_{X_{Tu}}^{X_T}\frac{\ddot{S}(x,\tilde{\theta}_T)}{\sigma^2(x)}dx\Bigg\|\frac{\|\hat{\theta}_T-\vartheta\|}{\sqrt{T(u(1-u))}}\\
&+\Bigg\|-(1-u)\int_0^{Tu}\left(\frac{\ddot{S}(X_s,\tilde{\theta}_T)S(X_s,\theta_0)}{\sigma^2(X_s)}+\frac{\ddot{S}'(X_s,\check{\theta}_T)}{2}-\frac{\sigma'(X_s)\ddot{S}(X_s,\tilde{\theta}_T)}{\sigma(X_s)}\right)ds\\
&+u\int_{Tu}^{T}\left(\frac{\ddot{S}(X_s,\tilde{\theta}_T)S(X_s,\theta_0)}{\sigma^2(X_s)}+\frac{\ddot{S}'(X_s,\check{\theta}_T)}{2}-\frac{\sigma'(X_s)\ddot{S}(X_s,\tilde{\theta}_T)}{\sigma(X_s)}\right)ds\Bigg\|\frac{\|\hat{\theta}_T-\vartheta\|}{\sqrt{T(u(1-u))}},
\end{aligned}$$

where $\tilde{\theta},\check{\theta}$ are elements between $\hat{\theta}_T$ and ϑ. The second term is equal to

$$\left\|-\frac{1}{\sqrt{T}}\int_0^T w_s^T(u)\left(\frac{\ddot{S}(X_s,\tilde{\theta}_T)S(X_s,\theta_0)}{\sigma^2(X_s)}+\frac{\ddot{S}'(X_s,\check{\theta}_T)}{2}-\frac{\sigma'(X_s)\ddot{S}(X_s,\tilde{\theta}_T)}{\sigma(X_s)}\right)ds\right\|\|\hat{\theta}_T-\vartheta\|.$$

The triangle inequality yields the upper bound

$$\begin{aligned}
&\Bigg(\left\|\frac{1}{\sqrt{T}}\int_0^T w_s^T(u)\left(\frac{\ddot{S}(X_s,\vartheta)S(X_s,\theta_0)}{\sigma^2(X_s)}+\frac{\ddot{S}'(X_s,\vartheta)}{2}-\frac{\sigma'(X_s)\ddot{S}(X_s,\vartheta)}{\sigma(X_s)}\right)ds\right\|\\
+\ &\Bigg\|\frac{1}{\sqrt{T}}\int_0^T w_s^T(u)\Bigg(\frac{(\ddot{S}(X_s,\tilde{\theta}_T)-\ddot{S}(X_s,\vartheta))S(X_s,\theta_0)}{\sigma^2(X_s)}\\
&+\frac{\ddot{S}'(X_s,\check{\theta}_T)-\ddot{S}'(X_s,\vartheta)}{2}-\frac{\sigma'(X_s)(\ddot{S}(X_s,\tilde{\theta}_T)-\ddot{S}(X_s,\vartheta))}{\sigma(X_s)}\Bigg)ds\Bigg\|\Bigg)\|\hat{\theta}_T-\vartheta\|.
\end{aligned}$$

Each element in the norm of the second term is bounded above by

$$\frac{1}{\sqrt{T}}\int_0^T |w_s^T(u)| \left(\frac{|\partial_i\partial_j S(X_s,\tilde{\theta}_T) - \partial_i\partial_j S(X_s,\vartheta)| S(X_s,\theta_0)}{\sigma^2(X_s)} \right.$$
$$\left. + \frac{|\partial_i\partial_j S'(X_s,\check{\theta}_T) - \partial_i\partial_j S'(X_s,\vartheta)|}{2} + \frac{|\sigma'(X_s)||\partial_i\partial_j S(X_s,\tilde{\theta}_T) - \partial_i\partial_j S(X_s,\vartheta)|}{\sigma(X_s)} \right) ds$$
$$\leq \frac{1}{T}\int_0^T |w_s^T(u)| \left(\frac{K(X_s)S(X_s,\theta_0)}{\sigma^2(X_s)} + \frac{K_d(X_s)}{2} + \frac{|\sigma'(X_s)|K(X_s)}{\sigma(X_s)} \right) ds \cdot \sqrt{T}\|\hat{\theta}_T - \vartheta\|$$

The Schwartz inequality yields the following bound for the left part

$$\sqrt{\frac{1}{T}\int_0^T |w_s^T(u)|^2 ds \frac{1}{T}\int_0^T \left(\frac{K(X_s)S(X_s,\theta_0)}{\sigma^2(X_s)} + \frac{K_d(X_s)}{2} + \frac{|\sigma'(X_s)|K(X_s)}{\sigma(X_s)} \right)^2 ds}$$
$$= \sqrt{\frac{1}{T}\int_0^T \left(\frac{K(X_s)S(X_s,\theta_0)}{\sigma^2(X_s)} + \frac{K_d(X_s)}{2} + \frac{|\sigma'(X_s)|K(X_s)}{\sigma(X_s)} \right)^2 ds}$$

Its $L^2([0,1],du)$ norm is asymptotically tight in $\mathbb{R}$ because of the ergodicity. As for the first term in the right-hand side of (5.2.2), the triangle inequality leads that

$$\left\| (1-u)\int_{X_0}^{X_{Tu}} \frac{\ddot{S}(x,\tilde{\theta}_T)}{\sigma^2(x)} dx - u\int_{X_{Tu}}^{X_T} \frac{\ddot{S}(x,\tilde{\theta}_T)}{\sigma^2(x)} dx \right\|$$
$$\leq \left\| (1-u)\int_{X_0}^{X_{Tu}} \frac{\ddot{S}(x,\vartheta)}{\sigma^2(x)} dx - u\int_{X_{Tu}}^{X_T} \frac{\ddot{S}(x,\vartheta)}{\sigma^2(x)} dx \right\| \tag{5.2.3}$$
$$+ \left\| (1-u)\int_{X_0}^{X_{Tu}} \frac{(\ddot{S}(x,\tilde{\theta}_T) - \ddot{S}(x,\vartheta))}{\sigma^2(x)} dx - u\int_{X_{Tu}}^{X_T} \frac{(\ddot{S}(x,\tilde{\theta}_T) - \ddot{S}(x,\vartheta))}{\sigma^2(x)} dx \right\|.$$

The second term is bounded above by

$$\left\| (1-u)\int_{X_0}^{X_{Tu}} \frac{|\ddot{S}(x,\tilde{\theta}_T) - \ddot{S}(x,\vartheta)|}{\sigma^2(x)} dx + u\int_{X_{Tu}}^{X_T} \frac{|\ddot{S}(x,\tilde{\theta}_T) - \ddot{S}(x,\vartheta)|}{\sigma^2(x)} dx \right\|$$

and by the Itô formula, the first term is equal to

$$
\begin{aligned}
&\Bigg\|(1-u)\left[\int_0^{Tu}\frac{\ddot S(X_s,\vartheta)}{\sigma(X_s)}dW_s\right.\\
&\left.+\int_0^{Tu}\left(\frac{\ddot S(X_s,\vartheta)S(X_s,\theta_0)}{\sigma^2(X_s)}+\frac{\ddot S'(X_s,\vartheta)}{2}-\frac{\sigma'(X_s)\ddot S(X_s,\vartheta)}{\sigma(X_s)}\right)ds\right]\\
&-u\left[\int_{Tu}^{T}\frac{\ddot S(X_s,\vartheta)}{\sigma(X_s)}dW_s\right.\\
&\left.+\int_{Tu}^{T}\left(\frac{\ddot S(X_s,\vartheta)S(X_s,\theta_0)}{\sigma^2(X_s)}+\frac{\ddot S'(X_s,\vartheta)}{2}-\frac{\sigma'(X_s)\ddot S(X_s,\vartheta)}{\sigma(X_s)}\right)ds\right]\Bigg\|.
\end{aligned}
$$

Therefore, the right-hand side of (5.2.3) is bounded by

$$
\begin{aligned}
&\left\|\int_0^T(1\{s\le Tu\}-u)\frac{\ddot S(X_s,\vartheta)}{\sigma(X_s)}dW_s\right\|\\
&+\left\|\int_0^T(1\{s\le Tu\}-u)\left(\frac{\ddot S(X_s,\vartheta)S(X_s,\theta_0)}{\sigma^2(X_s)}+\frac{\ddot S'(X_s,\vartheta)}{2}-\frac{\sigma'(X_s)\ddot S(X_s,\vartheta)}{\sigma(X_s)}\right)ds\right\|\\
&+d\left|(1-u)\int_{X_0}^{X_{Tu}}\frac{K(x)}{\sigma^2(x)}dx+u\int_{X_{Tu}}^{X_T}\frac{K(x)}{\sigma^2(x)}dx\right|\|\tilde\theta_T-\vartheta\|
\end{aligned}
$$

because of the Lipschitz condition. Therefore, it is sufficient to prove that

$$
\left\|\frac{1}{T}\int_0^T w_s^T(\cdot)\frac{\ddot S(X_s,\vartheta)}{\sigma(X_s)}dW_s\right\|^2_{L^2([0,1])}\to^p 0 \tag{5.2.4}
$$

$$
\begin{aligned}
&\left\|\frac{1}{T}\int_0^T w_s^T(\cdot)\left(\frac{\ddot S(X_s,\vartheta)S(X_s,\theta_0)}{\sigma^2(X_s)}+\frac{\ddot S'(X_s,\vartheta)}{2}-\frac{\sigma'(X_s)\ddot S(X_s,\vartheta)}{\sigma(X_s)}\right)ds\right\|^2_{L^2([0,1])}\\
&\to^p 0
\end{aligned} \tag{5.2.5}
$$

$$
\left\|\frac{1-\cdot}{T^{3/2}\sqrt{\cdot(1-\cdot)}}\int_{X_0}^{X_{T\cdot}}\frac{K(x)}{\sigma^2(x)}dx\right\|^2_{L^2([0,1])}\to^p 0 \tag{5.2.6}
$$

$$
\left\|\frac{\cdot}{T^{3/2}\sqrt{\cdot(1-\cdot)}}\int_{X_{T\cdot}}^{X_T}\frac{K(x)}{\sigma^2(x)}dx\right\|^2_{L^2([0,1])}\to^p 0 \tag{5.2.7}
$$

As for (5.2.4), it is enough to prove

$$\begin{aligned}
&\frac{1}{T^2}\int_0^1 \mathbb{E}\left[\left|\int_0^T w_s^T(u)\frac{\partial_i\partial_j S(X_s,\vartheta)}{\sigma(X_s)}dW_s\right|^2\right]du \\
= &\int_0^1 \mathbb{E}\left[\frac{1}{T^2}\int_0^T (w_s^T(u))^2\frac{(\partial_i\partial_j S(X_s,\vartheta))^2}{\sigma^2(X_s)}ds\right]du
\end{aligned}$$

converges to 0 for any i, j. The integrand with respect to du has the following bound which do not depend on T, u:

$$\sup_{s\in[0,\infty)}\mathbb{E}\left[\frac{(\partial_i\partial_j S(X_s,\vartheta))^2}{\sigma^2(X_s)}\right] \le \frac{C^2\sup_{s\in[0,\infty)}\mathbb{E}[(1+|X_s|)^{2p}]}{\inf_{x\in\mathbb{R}}\sigma(x)^2}$$

and the integrand converges to 0 for every u by the ergodicity. Hence, the bounded convergence theorem yields that (5.2.4) holds. This way shall be also used below. As for (5.2.5), it is enough to prove the convergence of the expectation to 0. It follows from the dominated convergence theorem with the ergodicity and the Schwartz inequality. As for (5.2.6), since the Itô formula yields that

$$\begin{aligned}
&\int_{X_0}^{X_{Tu}}\frac{K(x)}{\sigma^2(x)}dx \\
= &\int_0^{Tu}\frac{K(X_s)}{\sigma^2(X_s)}dW_s + \int_0^{Tu}\left(\frac{K(X_s)S(X_s,\theta_0)}{\sigma^2(X_s)} + \frac{K'(X_s)}{2} - \frac{\sigma'(X_s)K(X_s)}{\sigma(X_s)}\right)ds,
\end{aligned}$$

it is sufficient to prove

$$\mathbb{E}\left[\int_0^1\left(\frac{1-u}{T^{3/2}\sqrt{u(1-u)}}\int_0^{Tu}\frac{K(X_s)}{\sigma^2(X_s)}dW_s\right)^2 du\right] \to 0 \tag{5.2.8}$$

and

$$\begin{aligned}
&\mathbb{E}\left[\int_0^1 \frac{u(1-u)}{T}\left(\frac{1}{Tu}\int_0^{Tu}\left(\frac{K(X_s)S(X_s,\theta_0)}{\sigma^2(X_s)}\right.\right.\right. \\
&\left.\left.\left.+\frac{K'(X_s)}{2} - \frac{\sigma'(X_s)K(X_s)}{\sigma(X_s)}\right)ds\right)^2 du\right] \to 0. \qquad (5.2.9)
\end{aligned}$$

As for (5.2.8), the left-hand side is equal to

$$\begin{aligned}
&\int_0^1 \frac{1-u}{T^3 u}\mathbb{E}\left[\int_0^{Tu} \frac{(K(X_s))^2}{\sigma^4(X_s)} ds\right] du \\
\leq\ & \frac{1}{T^2}\int_0^1 (1-u)du \sup_{s\in[0,\infty)} \mathbb{E}\left[\frac{(K(X_s))^2}{\sigma^4(X_s)}\right] \to 0.
\end{aligned}$$

As for (5.2.9), the Jensen inequality gives the upper bound of the left-hand side

$$\mathbb{E}\left[\int_0^1 \frac{1-u}{T^2}\int_0^{Tu}\left(\frac{K(X_s)S(X_s,\theta_0)}{\sigma^2(X_s)} + \frac{K'(X_s)}{2} - \frac{\sigma'(X_s)K(X_s)}{\sigma(X_s)}\right)^2 dsdu\right].$$

It is bounded above by

$$\frac{1}{T}\int_0^1 u(1-u)du \sup_{s\in[0,\infty)} \mathbb{E}\left[\left(\frac{K(X_s)S(X_s,\theta_0)}{\sigma^2(X_s)} + \frac{K'(X_s)}{2} - \frac{\sigma'(X_s)K(X_s)}{\sigma(X_s)}\right)^2\right]$$

and it converges to 0. (5.2.7) is also valid by the same reason as (5.2.6). This completes the proof. □

Remark 5.2.4. *We proved the approximation under $\mathcal{H}_0$. The assertion under $\mathcal{H}_1$ is conjectured to be true, but its proof contains more complexities because case analyses are needed.*

Now, let us introduce the test statistic

$$AD_T = \int_0^1 \frac{\mathbb{Z}_T(u,\hat{\theta}_T)^\top \hat{C}_T^{-1}\mathbb{Z}_T(u,\hat{\theta}_T)}{u(1-u)}du,$$

where

$$(u,\theta) \rightsquigarrow \mathbb{Z}_T(u,\theta) = \frac{1}{\sqrt{T}}\int_0^T w_s^T(u)\frac{\dot{S}(X_s,\theta)}{\sigma(X_s)^2}(dX_s - S(X_s,\theta)ds)$$

and

$$\hat{C}_T = \frac{1}{T}\int_0^T \frac{\dot{S}(X_s,\hat{\theta}_T)\dot{S}(X_s,\hat{\theta}_T)^\top}{\sigma^2(X_s)}ds.$$

Under $\mathcal{H}_0$, $\hat{C}_T$ converges to $C_{\theta_0}(\theta_0,\theta_0)$ in probability by the ergodicity and the continuous differentiability of $\dot{S}$. The Theorem 5.1.1, the continuous mapping theorem and the Slutsky theorem yield the following theorem.

Theorem 5.2.1. *Assume conditions (I)-(VI). Under $\mathcal{H}_0$, it holds that $AD_T \to^d \|G\|_{L^2}^2$ as $T \to \infty$.*

Chapter 6

Discrete time stochastic processes

6.1 The limit distribution of Z-process

Firstly, set a measurable space and introduce a filtration. For a one dimensional discrete time adapted process$\{X_k\}_{k=1,2,\dots}$ whose state space is $\mathbb{R}$, consider a parametric model $\{P_\theta\}$ indexed by $\theta \in \Theta$, where Θ is a bounded open subset of $\mathbb{R}^d$. The true value of θ for X_k is denoted by $\theta_{(k)}$.

For the model above, we consider the following test:

$$
\begin{aligned}
&\mathcal{H}_0\text{: } \exists \theta_0 \in \Theta \textit{ such that } \theta_{(k)} = \theta_0,\ \forall k = 1, \dots, n \\
&\mathcal{H}_1\text{: } \exists \theta_0, \theta_1 \in \Theta,\ \exists u_* \in (0,1) \textit{ such that } \theta_{(k)} = \theta_0,\ \forall k = 1, \dots, [nu_*] \\
&\qquad \textit{and that } \theta_{(k)} = \theta_1 \neq \theta_0,\ \forall k = [nu_*]+1, \dots, n
\end{aligned}
$$

In order to estimate parameter θ, the following estimating equation is considered:

$$\Psi_n(\theta) = \frac{1}{n} \sum_{k=1}^{n} H_{k-1}(\theta)\xi_k(\theta) = 0,$$

where $\{\xi_k(\theta)\}_{k=1,\dots,n}$ becomes a martingale difference sequence when $\theta = \theta_{(k)}$ which satisfies

$$\sup_{k=1,2,\dots} \mathbb{E}\left[(\xi_k(\theta))^2\right] < \infty, \quad \forall \theta \in \Theta,$$

and H_{k-1} is a d dimensional measurable-$\mathcal{F}_{k-1}$ process such that

$$\frac{1}{n}\sum_{k=1}^{n} \|H_{k-1}(\theta)\|^2 \,\mathbb{E}[(\xi_k(\theta_{(k)}))^2|\mathcal{F}_{k-1}] < \infty, \quad a.s., \quad \forall\theta \in \Theta.$$

The solution, or an approximate solution, is denoted by $\hat{\theta}_n$, i.e. it holds that $\Psi_n(\hat{\theta}_n) = o_P(1)$, and used as an estimator for θ.

Assumptions II (D1) There exists the matrix $C_\kappa(\theta,\eta)$ given by

$$C_\kappa(\theta,\eta) = \underset{n\to\infty}{\text{l.i.m.}}\frac{1}{n}\sum_{k=1}^{n} H_{k-1}(\theta)H_{k-1}(\eta)^\top E[(\xi_k(\kappa))^2|\mathcal{F}_{k-1}]$$

and $C_{\theta_0}(\theta,\theta)$ is a positive definite for any $\theta \in \Theta$.
(D2) There exists the limits

$$\underset{n\to\infty}{\text{l.i.m.}}\frac{1}{n}\sum_{k=1}^{n} H_{k-1}(\theta)(\xi_k(\theta) - \xi_k(\theta_0))$$
$$\underset{n\to\infty}{\text{l.i.m.}}\frac{1}{n}\sum_{k=1}^{n} H_{k-1}(\theta)\dot{\xi}_k(\theta_0)$$

for any $\theta_0 \in \Theta$. Moreover, it holds that for any $\theta_0 \in \Theta$ and any $\varepsilon > 0$,

$$\inf_{\theta:\|\theta-\theta_0\|>\varepsilon}\left\|\underset{n\to\infty}{\text{l.i.m.}}\frac{1}{n}\sum_{k=1}^{n} H_{k-1}(\theta)(\xi_k(\theta) - \xi_k(\theta_0))\right\| > 0. \tag{6.1.1}$$

(D3) Under $\mathcal{H}_0$, $\sqrt{n}(\hat{\theta}_n - \theta_0) \to^d N(0, C_{\theta_0}(\theta_0,\theta_0)^{-1})$.
(D4) Under $\mathcal{H}_1$, $\hat{\theta}_T \to^p \theta_*$ which satisfies

$$\begin{aligned} & u_*\left(\underset{n\to\infty}{\text{l.i.m.}}\frac{1}{[nu_*]}\sum_{k=1}^{[nu_*]} H_{k-1}(\theta_*)(\xi_k(\theta_*) - \xi_k(\theta_0))\right) \\ & +(1-u_*)\left(\underset{n\to\infty}{\text{l.i.m.}}\frac{1}{(n-[nu_*])}\sum_{k=[nu_*]+1}^{n} H_{k-1}(\theta_*)(\xi_k(\theta_*) - \xi_k(\theta_1))\right) \\ = \;& u_* D^\infty(\theta_0) + (1-u_*)D^\infty(\theta_1) = 0, \end{aligned}$$

where
$$D^{\infty}(\theta) = \underset{n\to\infty}{\text{l.i.m.}} \frac{1}{n}\sum_{k=1}^{n} H_{k-1}(\theta_*)(\xi_k(\theta_*) - \xi_k(\theta)).$$

(D5) $H_k(\theta)$ is continuously differentiable with respect to θ.
(D6) $\xi_k(\theta)$ is second order continuously differentiable with respect to θ
(D7) Under $\mathcal{H}_1$, it holds that

$$\sup_{k=1,2,\ldots} \mathbb{E}\left[\|H_{k-1}(\theta_*)\|^2 \left(\xi_k(\theta)\right)^2\right] < \infty,$$

for $\theta \in \{\theta_0, \theta_1, \theta_*\}$.
(D8) Under $\mathcal{H}_0$, it holds that

$$\sup_{k=1,2,\ldots} \mathbb{E}\left[\|H_{k-1}(\theta_0)\|^2 \left(\partial_i \xi_k(\theta_0)\right)^2\right] < \infty,$$

for $i = 1, \ldots, d$. Under $\mathcal{H}_1$, it holds that

$$\sup_{k=1,2,\ldots} \mathbb{E}\left[\|H_{k-1}(\theta_*)\|^2 \left(\partial_i \xi_k(\theta_*)\right)^2\right] < \infty,$$

for $i = 1, \ldots, d$.
(D9) Under $\mathcal{H}_0$, it holds that

$$\sup_{k=1,2,\ldots} \mathbb{E}\left[\|H_{k-1}(\theta_0)\|^{2+\delta} (\xi_k(\theta_0))^2\right] < \infty,$$

for some $\delta > 0$. Under $\mathcal{H}_1$, it holds that

$$\sup_{k=1,2,\ldots} \mathbb{E}\left[\|H_{k-1}(\theta_*)\|^{2+\delta} (\xi_k(\theta_{(k)}))^2\right] < \infty,$$

for some $\delta > 0$.
(D10) Under $\mathcal{H}_0$, it holds that

$$\sup_{k=1,2,\ldots} \left\|\mathbb{E}\left[H_{k-1}(\theta_0) H_{k-1}(\theta_0)^{\top} (\xi_k(\theta_0))^2\right]\right\|_{OP}^2 < \infty.$$

Under $\mathcal{H}_1$, it holds that

$$\sup_{k=1,2,\ldots} \left\|\mathbb{E}\left[H_{k-1}(\theta_*) H_{k-1}(\theta_*)^{\top} (\xi_k(\theta_{(k)}))^2\right]\right\|_{OP}^2 < \infty.$$

(D11) It holds that

$$\left|(\dot{H}_{k-1}(\theta_1) - \dot{H}_{k-1}(\theta_2))_{(i,j)}\right| \leq K_{k-1}^{(i,j)} \|\theta_1 - \theta_2\|, \quad \forall \theta_1, \theta_2 \in N,$$

where under $\mathcal{H}_0$, N is a neighborhood of θ_0, and under $\mathcal{H}_1$, N is a neighborhood of θ_*. Moreover, under $\mathcal{H}_0$, it holds that

$$\sup_{k=1,2,\ldots} \mathbb{E}[(K_{k-1}^{(i,j)})^2(\xi_k(\theta_0))^2] < \infty,$$

and under $\mathcal{H}_1$, it holds that

$$\sup_{k=1,2,\ldots} \mathbb{E}[(K_{k-1}^{(i,j)})^2(\xi_k(\theta_{(k)}))^2] < \infty.$$

(D12) Under $\mathcal{H}_0$, it holds that

$$\sup_{k=1,2\ldots} \mathbb{E}\left[\|\partial_i H_{k-1}(\theta_0)\|^2(\xi_k(\theta_0))^2\right] < \infty,$$

for all $i = 1, \ldots, d$.
Under $\mathcal{H}_1$, it holds that

$$\sup_{k=1,2\ldots} \mathbb{E}\left[\|\partial_i H_{k-1}(\theta_*)\|^2(\xi_k(\theta_{(k)}))^2\right] < \infty,$$

for all $i = 1, \ldots, d$.
(D13) Under $\mathcal{H}_0$, there exists the limits

$$\operatorname*{l.i.m.}_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n}\left(K_{k-1}^{(i,j)}(\xi_k(\theta_0))\right)^2$$
$$\operatorname*{l.i.m.}_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n}\left(\partial_i H_{k-1}(\theta_0)\xi_k(\theta_0))\right)^2$$

for all $i, j = 1, \ldots, d$. Under $\mathcal{H}_1$, there exists the limits

$$\operatorname*{l.i.m.}_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n}\left(K_{k-1}^{(i,j)}(\xi_k(\theta_{(k)}))\right)^2$$
$$\operatorname*{l.i.m.}_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n}\left(\partial_i H_{k-1}(\theta_*)\xi_k(\theta_{(k)}))\right)^2$$

for all $i, j = 1, \ldots, d$.

Assumptions II-S Assume (D1)-(D8) and (D11)-(D13). (D14) Under $\mathcal{H}_0$, it holds that

$$\sup_{k=1,2,\ldots} \mathbb{E}\left[\|H_{k-1}(\theta_0)\|^4(\xi_k(\theta_0))^4\right] < \infty.$$

Under $\mathcal{H}_1$, it holds that

$$\sup_{k=1,2,\ldots} \mathbb{E}\left[\|H_{k-1}(\theta_*)\|^4(\xi_k(\theta_{(k)}))^4\right] < \infty.$$

(D15) Under $\mathcal{H}_0$, it holds that

$$\sup_{k=1,2,\ldots} \mathbb{E}\left[\|H_{k-1}(\theta_0)\|^4\right] < \infty.$$

Under $\mathcal{H}_1$, it holds that

$$\sup_{k=1,2,\ldots} \mathbb{E}\left[\|H_{k-1}(\theta_*)\|^4\right] < \infty.$$

Proposition 6.1.1. *Assumptions II-S is a sufficient condition for Assumptions II. Especially,*
(D14) implies (D10). (D14) and (D15) imply (D9).

Proof of the Proposition 6.1.1 (i) In general it holds that for d dimensional random vector a, b,

$$\begin{aligned}
\left\|\mathbb{E}[ab^\top]\right\|_{OP}^2 &\leq \operatorname{tr}\left(\left(\mathbb{E}[ab^\top]\right)^\top \mathbb{E}[ab^\top]\right) \\
&\leq \operatorname{tr}\left(\mathbb{E}[ba^\top ab^\top]\right) = \mathbb{E}[\operatorname{tr}\left(ba^\top ab^\top\right)] = \mathbb{E}[\|a\|^2\|b\|^2] \\
&\leq \mathbb{E}[\|a\|^4 + \|b\|^4],
\end{aligned}$$

where the first inequality is led by the definition of the operator norm, the second inequality is led by the fact that $\mathbb{E}[A^\top A] - (\mathbb{E}[A])^\top \mathbb{E}[A] = \mathbb{E}[(A - \mathbb{E}[A])^\top(A - \mathbb{E}[A])]$ is non negative definite matrix. It follows from this inequality that

$$\begin{aligned}
&\sup_{k=1,2,\ldots} \left\|\mathbb{E}\left[H_{k-1}(\theta)H_{k-1}(\theta)^\top(\xi_k(\theta_{(k)}))^2\right]\right\|_{OP}^2 \\
\leq &\sup_{k=1,2,\ldots} \mathbb{E}\left[\left(\|H_{k-1}(\theta)\|^4 + \|H_{k-1}(\theta)\|^4\right)\left(\xi_k(\theta_{(k)})\right)^4\right] < \infty.
\end{aligned}$$

This completes the proof. □

Introduce the random field $\{\mathbb{Z}_n(u,\theta); (u,\theta) \in (0,1) \times \Theta\}$ given by

$$\mathbb{Z}_n(u,\theta) = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} w_k^n(u) H_{k-1}(\theta) \xi_k(\theta),$$

where

$$w_k^n(u) = \begin{cases} 0, & u \in \left(0, \frac{1}{n}\right), \\ \frac{1\{k \leq nu\} - [nu]/n}{\sqrt{[nu]/n(1-[nu]/n)}}, & u \in \left[\frac{1}{n}, 1\right), \ k = 1, \dots, n \end{cases}.$$

Its "predictable projection" to the true model is

$$\mathbb{Z}_n^p(u,\theta) = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} w_k^n(u) H_{k-1}(\theta) (\xi_k(\theta) - \xi_k(\theta_{(k)})).$$

The difference between $\mathbb{Z}$ and $\mathbb{Z}^p$ is denoted by $\{\mathbb{M}_n(u,\theta); (u,\theta) \in (0,1) \times \Theta\}$, say, it is given by

$$\mathbb{M}_n(u,\theta) = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} w_k^n(u) H_{k-1}(\theta) \xi_k(\theta_{(k)}).$$

Under $\mathcal{H}_0$, it holds that

$$\mathbb{Z}_n^p(\cdot, \theta_0) = 0,$$

so

$$\mathbb{M}_n(\cdot, \theta_0) = \mathbb{Z}_n(\cdot, \theta_0).$$

This relationship gives us the idea to use functions of $\mathbb{Z}_n$ as a test statistic. However, since we cannot know the true value θ_0, it is crucial to hold that under $\mathcal{H}_0$,

$$\mathbb{Z}_n(\cdot, \hat{\theta}_n) - \mathbb{Z}_n(\cdot, \theta_0) \to^p 0$$

which enables us to apply the limit theorem in the preceding chapter. Moreover, in order to ensure the power of the test, it is crucial to hold that under $\mathcal{H}_1$,

$$\frac{1}{\sqrt{n}} \mathbb{Z}_n^p(\cdot, \hat{\theta}_n) \not\to^p 0.$$

Lemma 6.1.1. *Under $\mathcal{H}_0$, it holds that*

$$\left\| \mathbb{Z}_n^p(\cdot, \hat{\theta}_n) \right\|_{L^2}^2 \to^p 0$$

as $n \to \infty$

Proof of the Lemma 6.1.1 The Taylor expansion yields that

$$\begin{aligned}\mathbb{Z}_n^p(u,\hat{\theta}_T) &= \frac{1}{\sqrt{n}}\sum_{k=1}^n w_k^n(u)H_{k-1}(\hat{\theta}_n)(\xi_k(\hat{\theta}_n)-\xi_k(\theta_{(k)})) \\ &= \frac{1}{n}\sum_{k=1}^n w_k^n(u)H_{k-1}(\hat{\theta}_n)\dot{\xi}_k(\tilde{\theta}_n)^\top\sqrt{n}(\hat{\theta}_n-\theta_0),\end{aligned}$$

where $\tilde{\theta}_n$ is a value between θ_0 and $\hat{\theta}_n$. Because of the assumption $\sqrt{n}(\hat{\theta}_n-\theta_0)=O_P(1)$, we argue the convergence to 0 in probability in $L^2([0,1],du)$ of the all elements in the following matrix:

$$\frac{1}{n}\sum_{k=1}^n w_k^n(\cdot)H_{k-1}(\hat{\theta}_n)\dot{\xi}_k(\tilde{\theta}_n)^\top = \frac{1}{n}\sum_{k=1}^n w_k^n(\cdot)H_{k-1}(\theta_0)\dot{\xi}_k(\theta_0)^\top + o_P(1),$$

by the continuous differentiability of $H_{k-1}(\theta)$ and $\dot{\xi}_k(\theta)$ with respect to θ. The terms in the right-hand side converge to 0 in the second mean for any u because of the assumption (D2). Moreover, by the Schwartz inequality and assumption (D8), it holds that

$$\begin{aligned}&\mathbb{E}\left[\left\|\frac{1}{n}\sum_{k=1}^n w_k^n(u)H_{k-1}(\theta_0)\partial_i\xi_k(\theta_0)\right\|^2\right] \\ \le\ & \mathbb{E}\left[\frac{1}{n}\sum_{k=1}^n(w_k^n(u))^2\frac{1}{n}\sum_{k=1}^n H_{k-1}(\theta_0)^\top H_{k-1}(\theta_0)\left(\partial_i\xi_k(\theta_0)\right)^2\right] \\ \le\ & \sup_{k=1,2,\ldots}\mathbb{E}\left[H_{k-1}(\theta_0)^\top H_{k-1}(\theta_0)\left(\partial_i\xi_k(\theta_0)\right)^2\right]<\infty,\end{aligned}$$

for all $i=1,\ldots,d$ and $u\in[0,1]$. Hence, the Fubini theorem and the bounded convergence theorem yield that

$$\begin{aligned}&\lim_{n\to\infty}\mathbb{E}\left[\int_0^1\left\|\frac{1}{n}\sum_{k=1}^n w_k^n(u)H_{k-1}(\theta_0)\partial_i\xi_k(\theta_0)\right\|^2 du\right] \\ =\ & \lim_{n\to\infty}\int_0^1\mathbb{E}\left[\left\|\frac{1}{n}\sum_{k=1}^n w_k^n(u)H_{k-1}(\theta_0)\partial_i\xi_k(\theta_0)\right\|^2\right]du \\ =\ & \int_0^1\lim_{n\to\infty}\mathbb{E}\left[\left\|\frac{1}{n}\sum_{k=1}^n w_k^n(u)H_{k-1}(\theta_0)\partial_i\xi_k(\theta_0)\right\|^2\right]du = 0\end{aligned}$$

for all $i = 1, \ldots, d$. This completes the proof. $\square$

Lemma 6.1.2. *Under $\mathcal{H}_0$, it holds that*

$$\left\|\mathbb{M}_n(\cdot, \hat{\theta}_n) - \mathbb{M}_n(\cdot, \theta_0)\right\|_{L^2}^2 \to^p 0$$

as $n \to \infty$

Proof of the Lemma 6.1.2 The Taylor expansion yields that

$$\mathbb{M}_n(u, \hat{\theta}_n) - \mathbb{M}_n(u, \theta_0) = \frac{1}{n}\sum_{k=1}^{n} w_k^n(u)\dot{H}_{k-1}(\tilde{\theta}_n)^\top \xi_k(\theta_0)\sqrt{n}(\hat{\theta}_n - \theta_0),$$

where $\tilde{\theta}_n$ is a value between θ_0 and $\hat{\theta}_n$. Because of the assumption $\sqrt{n}(\hat{\theta}_n - \theta_0) = O_P(1)$, we argue the convergence to 0 in probability in $L^2([0,1], du)$ of the all elements in the following matrix:

$$\frac{1}{n}\sum_{k=1}^{n} w_k^n(\cdot)\dot{H}_{k-1}(\tilde{\theta}_n)^\top \xi_k(\theta_0) = \frac{1}{n}\sum_{k=1}^{n} w_k^n(\cdot)\dot{H}_{k-1}(\theta_0)^\top \xi_k(\theta_0) + o_P(1),$$

where the last equality is followed by the Schwartz inequality and the assumption (D11). Indeed, it holds that

$$\begin{aligned}
&\int \left(\frac{1}{n}\sum_{k=1}^{n} w_k^n(u)\left(\dot{H}_{k-1}(\tilde{\theta}_n) - \dot{H}_{k-1}(\theta_0)\right)_{(i,j)} \xi_k(\theta_0)\right)^2 du \\
\leq\ & \int \frac{1}{n^2}\sum_{k=1}^{n}(w_k^n(u))^2 \sum_{k=1}^{n}\left(\dot{H}_{k-1}(\tilde{\theta}_n) - \dot{H}_{k-1}(\theta_0)\right)_{(i,j)}^2 (\xi_k(\theta_0))^2 du \\
=\ & \frac{1}{n}\sum_{k=1}^{n}\left(\dot{H}_{k-1}(\tilde{\theta}_n) - \dot{H}_{k-1}(\theta_0)\right)_{(i,j)}^2 (\xi_k(\theta_0))^2 \\
\leq\ & \frac{1}{n}\left(\frac{1}{n}\sum_{k=1}^{n}\left(K_{k-1}^{(i,j)}\xi_k(\theta_0)\right)^2\right)\left\|\sqrt{n}(\tilde{\theta}_n - \theta_0)\right\|^2 \to^p 0,
\end{aligned}$$

for all $i, j = 1, \ldots, d$, where the last convergence is followed by the Slustsky theorem since $\sqrt{n}(\hat{\theta}_n - \theta_0) = O_P(1)$ and

$$\mathbb{E}\left[\frac{1}{n}\left(\frac{1}{n}\right)\sum_{k=1}^{n}(K_{k-1}^{(i,j)}\xi_k(\theta_0))^2\right] \leq \frac{1}{n}\sup_{k=1,2,\ldots}\mathbb{E}\left[(K_{k-1}^{(i,j)}\xi_k(\theta_0))^2\right] \to 0$$

hold. It holds that

$$
\begin{aligned}
& \mathbb{E}\left[\left\|\frac{1}{n}\sum_{k=1}^{n} w_k^n(u)\partial_i H_{k-1}(\theta_0)\xi_k(\theta_0)\right\|^2\right] \\
\leq\ & \mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n}(w_k^n(u))^2\frac{1}{n}\sum_{k=1}^{n}\partial_i H_{k-1}(\theta_0)^\top \partial_i H_{k-1}(\theta_0)(\xi_k(\theta_0))^2\right] \\
\leq\ & \sup_{k=1,2\dots} \mathbb{E}\left[\partial_i H_{k-1}(\theta_0)^\top \partial_i H_{k-1}(\theta_0)(\xi_k(\theta_0))^2\right] < \infty.
\end{aligned}
$$

In consequence, it follows from the Fubini theorem and the bounded convergence theorem that

$$
\mathbb{E}\left[\left\|\frac{1}{n}\sum_{k=1}^{n} w_k^n(\cdot)\dot{H}_{k-1}(\theta_0)^\top \xi_k(\theta_0)\right\|_{L^2}^2\right] \to 0,
$$

because the assumption (D13) yields the pointwise convergence to 0. This completes the proof. □

Next, we discuss a limit theorem for $\mathbb{M}_n(\cdot,\theta_0)$ which is taking values in $L^2([0,1],du)$ spaces, which is a consequence of Theorem 4.2.2.

Lemma 6.1.3. *Under $\mathcal{H}_0$, it holds that the random field $u \rightsquigarrow \mathbb{M}_n(u,\theta_0)$ converges weakly to $u \rightsquigarrow C_{\theta_0}(\theta_0,\theta_0)B_d^\circ(u)/\sqrt{u(1-u)}$ in $L^2([0,1],du)$, where B_d° is the d dimensional Brownian bridge.*

It is ready to propose a test statistic AD_n defined by

$$
\begin{aligned}
AD_n &= \int_0^1 \mathbb{Z}_n(u,\hat{\theta}_n)^\top \hat{C}_n^{-1}\mathbb{Z}_n(u,\hat{\theta}_n)du \\
&= \left\|\hat{C}_n^{-1/2}\mathbb{Z}_n(\cdot,\hat{\theta}_n)\right\|_{L^2}^2,
\end{aligned}
$$

where $\hat{C}_n$ is a consistent estimator for $C_{\theta_0}(\theta_0,\theta_0)$ under $\mathcal{H}_0$. By the preceding lemma, the Slutsky theorem and the continuous mapping theorem yield the former assertion of the following theorem.

Theorem 6.1.1. *(i) Under $\mathcal{H}_0$, it holds that*

$$
AD_n \to^d \|G\|_{L^2}^2
$$

as $n \to \infty$, where $u \rightsquigarrow G(u) = B_d^\circ(u)/\sqrt{u(1-u)}$. (ii) Under $\mathcal{H}_1$, AD_n diverges to positive infinity as $n \to \infty$.

PROOF OF THE THEOREM 6.1.1.(ii). Since $\hat{C}_n^{-1}$ is non-negative definite matrix, it holds that

$$2v_1^\top \hat{C}_n^{-1} v_1 + 2v_2^\top \hat{C}_n^{-1} v_2 \geq (v_1 - v_2)^\top \hat{C}_n^{-1}(v_1 - v_2)$$

for arbitrary d-dimensional vector v_1, v_2. This property and the inequality $\sqrt{[nu]/n(1-[nu]/n)} \leq 1/2$ yield that

$$\begin{aligned} AD_n &\geq \frac{1}{2}\int_0^1 (Z_n^p)^\top(u,\hat{\theta}_n)\hat{C}_n^{-1} Z_n^p(u,\hat{\theta}_n)du - \int_0^1 \mathbb{M}_n(u,\hat{\theta}_n)\hat{C}_n^{-1}\mathbb{M}_n(u,\hat{\theta}_n)du \\ &\geq 2n\int_0^1 A_n^\top(u,\hat{\theta}_n)\hat{C}_n^{-1}A_n(u,\hat{\theta}_n)du - \int_0^1 \mathbb{M}_n(u,\hat{\theta}_n)\hat{C}_n^{-1}\mathbb{M}_n(u,\hat{\theta}_n)du, \end{aligned}$$

where

$$\begin{aligned} A_n(u) &= \frac{\sqrt{[nu]/n(1-[nu]/n)}}{\sqrt{n}}\mathbb{Z}_n^p(u,\hat{\theta}_n) \\ &= \frac{1}{n}\sum_{k=1}^n \left(1\{k \leq nu\} - \frac{[nu]}{n}\right) H_{k-1}(\hat{\theta}_n)(\xi_k(\hat{\theta}_n) - \xi_k(\theta_0)1\{k \leq nu_*\} \\ &\quad -\xi_k(\theta_1)1\{k \geq nu_*\}) \\ &= \frac{1}{n}\sum_{k=1}^n \left(1\{k \leq nu\} - \frac{[nu]}{n}\right) H_{k-1}(\theta_*)(\xi_k(\theta_*) - \xi_k(\theta_0)1\{k \leq nu_*\} \\ &\quad -\xi_k(\theta_1)1\{k \geq nu_*\}) + o_P(1) \\ &= \tilde{A}_n(u) + o_P(1), \quad (say). \end{aligned}$$

The second last equality can be obtained by the same reason as the Lemma 6.1.1. It holds that, for $u < [nu_*]/n$,

$$\begin{aligned} \tilde{A}_n(u) &= \frac{1}{n}\left(1 - \frac{[nu]}{n}\right)\sum_{k=1}^{[nu]} H_{k-1}(\theta_*)(\xi_k(\theta_*) - \xi_k(\theta_0)) \\ &\quad +\frac{1}{n}\left(-\frac{[nu]}{n}\right)\sum_{k=[nu]+1}^{[nu_*]} H_{k-1}(\theta_*)(\xi_k(\theta_*) - \xi_k(\theta_0)) \\ &\quad +\frac{1}{n}\left(-\frac{[nu]}{n}\right)\sum_{k=[nu_*]+1}^{n} H_{k-1}(\theta_*)(\xi_k(\theta_*) - \xi_k(\theta_1)) \end{aligned}$$

so,

$$\begin{aligned}\underset{n\to\infty}{\text{l.i.m.}}\tilde{A}_n(u) &= (u(1-u)-u(u_*-u))D^\infty(\theta_0)-u(1-u_*)D^\infty(\theta_1)\\ &= u(1-u_*)(D^\infty(\theta_0)-D^\infty(\theta_1)),\end{aligned}$$

for $[nu_*]/n \le u < ([nu_*]+1)/n$,

$$\begin{aligned}\tilde{A}_n(u) &= \frac{1}{n}\left(1-\frac{[nu_*]}{n}\right)\sum_{k=1}^{[nu_*]} H_{k-1}(\theta_*)(\xi_k(\theta_*)-\xi_k(\theta_0))\\ &\quad +\frac{1}{n}\left(-\frac{[nu_*]}{n}\right)\sum_{k=[nu_*]+1}^{n} H_{k-1}(\theta_*)(\xi_k(\theta_*)-\xi_k(\theta_1))\end{aligned}$$

so,

$$\underset{n\to\infty}{\text{l.i.m.}}\tilde{A}_n(u) = u_*(1-u_*)(D^\infty(\theta_0)-D^\infty(\theta_1)),$$

and for $u \ge ([nu_*]+1)/n$,

$$\underset{n\to\infty}{\text{l.i.m.}}\tilde{A}_n(u) = u_*(1-u)(D^\infty(\theta_0)-D^\infty(\theta_1)).$$

Let us denote

$$A_\infty(u) = \begin{cases} u(1-u_*)(D^\infty(\theta_0)-D^\infty(\theta_1)), & u\in(0,u_*)\,,\\ u_*(1-u)(D^\infty(\theta_0)-D^\infty(\theta_1)), & u\in[u_*,1)\,.\end{cases}$$

Next, we shall prove

$$\mathbb{E}\left[\|\tilde{A}_n - A_\infty\|_{L^2}^2\right]\to 0. \tag{6.1.2}$$

It holds that for all u,

$$\mathbb{E}\left[\left(\tilde{A}_n(u)-A_\infty(u)\right)^2\right] \le 2\mathbb{E}\left[(\tilde{A}_n(u))^2\right]+2(A_\infty(u))^2$$

and the first term in the right-hand side is bounded above by

$$\begin{aligned}&2\mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n}\left(1\{k\le nu\}-\frac{[nu]}{n}\right)^2 H_{k-1}(\theta_*)^\top H_{k-1}(\theta_*)(\xi_k(\theta_*)\right.\\ &\qquad \left.-\xi_k(\theta_0)1\{k\le nu_*\}-\xi_k(\theta_1)1\{k\ge nu_*\})^2\right]\\ \le\ & \frac{2}{n}\sum_{k=1}^{n}\left(1\{k\le nu\}-\frac{[nu]}{n}\right)^2 \sup_{k=1,2,\ldots}\mathbb{E}\left[H_{k-1}(\theta_*)^\top H_{k-1}(\theta_*)(\xi_k(\theta_*)\right.\\ &\qquad \left.-\xi_k(\theta_0)1\{k\le nu_*\}-\xi_k(\theta_1)1\{k\ge nu_*\})^2\right]\\ \le\ & 2\sup_{k=1,2,\ldots}\mathbb{E}\left[H_{k-1}(\theta_*)^\top H_{k-1}(\theta_*)(\xi_k(\theta_0)^2+\xi_k(\theta_1)^2+\xi_k(\theta_*)^2)\right]<\infty.\end{aligned}$$

Since the left-hand side of (6.1.2) is equal to

$$\int_0^1 \mathbb{E}\left[\left(\tilde{A}_n(u) - A_\infty(u)\right)^2\right] du$$

and $(A_\infty(u))^2$ is integrable with respect to u, the dominated convergence theorem yields (6.1.2), and (6.1.2) yields that $\tilde{A}_n \to^p A_\infty$ in L^2. This result, the Slutsky theorem and the continuous mapping theorem yields that

$$\int_0^1 A_n^\top(u)\hat{C}_n^{-1}A_n(u)du \to^p \int_0^1 A_\infty^\top(u)C_*^{-1}A_\infty(u)du,$$

where $C_* := u_*C_{\theta_0}(\theta_*, \theta_*) + (1-u_*)C_{\theta_1}(\theta_*, \theta_*)$. By simple calculations, the right-hand side is equal to

$$\frac{u_*^2(1-u_*)^2}{3}(D^\infty(\theta_0) - D^\infty(\theta_1))^\top C_*^{-1}(D^\infty(\theta_0) - D^\infty(\theta_1)).$$

Finally, $\mathbb{M}_n(\cdot, \hat{\theta}_n)$ is asymptotically tight in $L^2([0,1], du)$ by the assumption (D11), which guarantees an approximation of θ_* by consistent estimator, and the Theorem 4.2.2. The last assertion is followed since C_* is positive definite and the assumption (6.1.1). This completes the proof. □

6.2 A change detection procedure for for an ergodic time series

Consider the time series model given by

$$X_k = \tilde{S}(X_{k-1}, \ldots, X_{k-q_1}; \theta) + \tilde{\sigma}(X_{k-1}, \ldots, X_{k-q_2}; \theta)\varepsilon_k$$

for $k = 1, \ldots, n$, where $\{\varepsilon_k\}_{k=1}^n$ are independently, identically distributed with $N(0, 1^2)$ and $\theta \in \Theta \subset \mathbb{R}^d$. Assume that $\tilde{S}$ and $\tilde{\sigma}$ do not have the same elements of the parameter θ. By putting $q = q_1 \vee q_2$ and changing the domain of the functions $\tilde{S}$ and $\tilde{\sigma}$, we can write

$$X_k = S(\mathbb{X}_k; \theta) + \sigma(\mathbb{X}_k; \theta)w_k,$$

where $\mathbb{X}_k = (X_{k-1}, \ldots, X_{k-q})$ and $S(\cdot;\theta)$ and $\sigma(\cdot;\theta)$ are some measurable functions on $\mathbb{R}^q$. We can substitute $X_0. \ldots, X_{-q}$ with zero, or other finite value. Suppose that X is ergodic in the second mean with respect to an invariant measure μ, that is, for any μ- measurable function f, it holds that

$$\lim_{n\to\infty} \mathbb{E}\left[\left\|\frac{1}{n}\sum_{k=1}^{n} f(X_k, \mathbb{X}_k) - \int_{\mathbb{R}^{q+1}} f(x)\mu(dx)\right\|^2\right] = 0.$$

Assume that $S(x,\theta)$ and $\sigma(x,\theta)$ are third order continuously differentiable with respect to θ.

Consider the estimating equation

$$\begin{aligned}
\Psi_n(\theta) &= \frac{1}{n}\sum_{k=1}^{n} \frac{1}{\sigma(\mathbb{X}_k;\theta)} \left(\dot{S}(\mathbb{X}_k;\theta)\left(\frac{X_k - S(\mathbb{X}_k;\theta)}{\sigma(\mathbb{X}_k;\theta)}\right)\right. \\
&\quad \left. +\dot{\sigma}(\mathbb{X}_k;\theta)\left(\left(\frac{X_k - S(\mathbb{X}_k;\theta)}{\sigma(\mathbb{X}_k;\theta)}\right)^2 - 1\right)\right) \\
&= \frac{1}{n}\sum_{k=1}^{n} \frac{\dot{S}(\mathbb{X}_k;\theta)}{(\sigma(\mathbb{X}_k;\theta))^2}\xi_k^a(\theta) + \frac{1}{n}\sum_{k=1}^{n} \frac{\dot{\sigma}(\mathbb{X}_k;\theta)}{(\sigma(\mathbb{X}_k;\theta))^3}\xi_k^b(\theta) \\
&= \Psi_n^a(\theta) + \Psi_n^b(\theta) \quad (say) \\
&= 0,
\end{aligned}$$

where ξ_k^a and ξ_k^b are defined by

$$\xi_k^a(\theta) = X_k - S(\mathbb{X}_k;\theta)$$

and

$$\xi_k^b(\theta) = (X_k - S(\mathbb{X}_k;\theta))^2 - (\sigma(\mathbb{X}_k;\theta))^2,$$

respectively. The solution of the above estimating equation is denoted by $\hat{\theta}_n$.

Define the random fields:

$$(u,\theta) \rightsquigarrow \mathbb{Z}_n(u,\theta) = \mathbb{Z}_n^a(u,\theta) + \mathbb{Z}_n^b(u,\theta),$$

where

$$\begin{aligned}
\mathbb{Z}_n^a(u,\theta) &= \frac{1}{n}\sum_{k=1}^{n} w_k^n(u)\frac{\dot{S}(\mathbb{X}_k;\theta)}{(\sigma(\mathbb{X}_k;\theta))^2}\xi_k^a(\theta) \\
\mathbb{Z}_n^b(u,\theta) &= \frac{1}{n}\sum_{k=1}^{n} w_k^n(u)\frac{\dot{\sigma}(\mathbb{X}_k;\theta)}{(\sigma(\mathbb{X}_k;\theta))^3}\xi_k^b(\theta);
\end{aligned}$$

$$(u,\theta) \rightsquigarrow \mathbb{Z}_n^p(u,\theta) = \mathbb{Z}_n^{p,a}(u,\theta) + \mathbb{Z}_n^{p,b}(u,\theta),$$

where

$$\begin{aligned}
\mathbb{Z}_n^{p,a}(u,\theta) &= \frac{1}{n}\sum_{k=1}^{n} w_k^n(u)\frac{\dot{S}(\mathbb{X}_k;\theta)}{(\sigma(\mathbb{X}_k;\theta))^2}(\xi_k^a(\theta)-\xi_k^a(\theta_{(k)}))\\
\mathbb{Z}_n^{p,b}(u,\theta) &= \frac{1}{n}\sum_{k=1}^{n} w_k^n(u)\frac{\dot{\sigma}(\mathbb{X}_k;\theta)}{(\sigma(\mathbb{X}_k;\theta))^3}(\xi_k^b(\theta)-\xi_k^b(\theta_{(k)}));
\end{aligned}$$

$$(u,\theta) \rightsquigarrow \mathbb{M}_n(u,\theta) = \mathbb{M}_n^a(u,\theta) + \mathbb{M}_n^b(u,\theta),$$

where

$$\begin{aligned}
\mathbb{M}_n^a(u,\theta) &= \frac{1}{n}\sum_{k=1}^{n} w_k^n(u)\frac{\dot{S}(\mathbb{X}_k;\theta)}{(\sigma(\mathbb{X}_k;\theta))^2}\xi_k^a(\theta_{(k)})\\
\mathbb{M}_n^b(u,\theta) &= \frac{1}{n}\sum_{k=1}^{n} w_k^n(u)\frac{\dot{\sigma}(\mathbb{X}_k;\theta)}{(\sigma(\mathbb{X}_k;\theta))^3}\xi_k^b(\theta_{(k)}).
\end{aligned}$$

As for the Lemmas 6.1.1 and 6.1.2, it is sufficient to treat the first term and the second term separately by the ineqaulity $(a+b)^2 \le 2(a^2+b^2)$. In order to prove the Lemma 6.1.3, it is sufficient to prove the weak convergences of $\mathbb{M}_n^a(\cdot,\theta_0)$ and $\mathbb{M}_n^b(\cdot,\theta_0)$ since the limits are the Gaussian and it follows from the tower property that

$$\mathbb{E}[\mathbb{M}_n^a(\cdot,\theta_0)^\top \mathbb{M}_n^b(\cdot,\theta_0)] = 0.$$

The matrix $C_\kappa(\theta,\eta)$ is

$$C_\kappa(\theta,\eta) = C_\kappa^a(\theta,\eta) + C_\kappa^b(\theta,\eta),$$

where

$$C_\kappa^a(\theta,\eta) = \int_{\mathbb{R}^q} \frac{\dot{S}(x;\theta)\dot{S}(x;\eta)^\top}{(\sigma(x;\theta))^2(\sigma(x;\eta))^2}(\sigma(x;\kappa))^2\mu_\kappa(dx)$$

and

$$C_\kappa^a(\theta,\eta) = \int_{\mathbb{R}^q} \frac{\dot{\sigma}(x;\theta)\dot{\sigma}(x;\eta)^\top}{(\sigma(x;\theta))^3(\sigma(x;\eta))^3}(\sigma(x;\kappa))^4\mu_\kappa(dx).$$

Suppose the following conditions:
(I) There exists measurable function Λ on $\mathbb{R}^q$ such that for $\theta \in \{\theta_0, \theta_1, \theta_*\}$ and i=1,. . . , d,

$$\begin{aligned} |S(x,\theta)| &\leq \Lambda(x), \\ \|\dot{S}(x,\theta)\| &\leq \Lambda(x), \\ \|\partial_i \dot{S}(x,\theta)\| &\leq \Lambda(x), \\ \sigma^2(x,\theta) &\leq \Lambda(x), \\ \|\dot{\sigma}\|^2(x,\theta) &\leq \Lambda(x), \\ \|\partial_i \dot{\sigma}(x,\theta)\|^2 &\leq \Lambda(x). \end{aligned}$$

(II) It holds that

$$\inf_{\theta\in\Theta} \inf_{x\in\mathbb{R}^q} \sigma(x;\theta) > 0.$$

(III) It holds that

$$\sup_{k=1,2,\ldots} \mathbb{E}[(\Lambda(\mathbb{X}_k))^6] < \infty.$$

(IV) Define $\Psi(\theta, \theta_0) = \Psi^a(\theta, \theta_0) + \Psi^b(\theta, \theta_0)$, where

$$\Psi^a(\theta, \theta_0) = \int_{\mathbb{R}^q} \frac{(S(x,\theta_0) - S(x,\theta))\dot{S}(x,\theta)}{\sigma^2(x,\theta)} \mu_{\theta_0}(dx)$$

and

$$\Psi^b(\theta, \theta_0) = \int_{\mathbb{R}^q} \frac{((\sigma(x,\theta_0))^2 - (\sigma(x,\theta))^2)\dot{\sigma}(x,\theta)}{\sigma^3(x,\theta)} \mu_{\theta_0}(dx).$$

For all $\theta \in \Theta$ and any ε, it holds that

$$\sup_{\theta\in\Theta} \|\Psi_n(\theta,\theta_0) - \Psi(\theta,\theta_0)\| \to^p 0$$

and that

$$\inf_{\theta:\|\theta-\theta_0\|>\varepsilon} \|\Psi(\theta,\theta_0)\| > 0.$$

(V) The matrix $C_\kappa(\theta, \theta)$ is invertible for all $\theta, \kappa \in \Theta$.
(VI) Under $\mathcal{H}_0$, $\sqrt{n}(\hat{\theta} - \theta_0) \to^d N(0, C^{-1}(\theta_0, \theta_0))$ and under $\mathcal{H}_1$, $\hat{\theta}_n \to^p \theta_*$ which satisfies

$$u_* \Psi(\theta_*, \theta_0) + (1 - u_*)\Psi(\theta_*, \theta_1) = 0.$$

(VII) It holds that

$$\left|(\ddot{S}(\mathbb{X}_k, \theta_1) - \ddot{S}(\mathbb{X}_k, \theta_2))_{(i,j)}\right| \leq K^{a,(i,j)}(\mathbb{X}_k) \|\theta_1 - \theta_2\|, \quad \forall \theta_1, \theta_2 \in N,$$

and

$$\left|(\ddot{\sigma}(\mathbb{X}_k,\theta_1)-\ddot{\sigma}(\mathbb{X}_k,\theta_2))_{(i,j)}\right| \leq K^{b,(i,j)}(\mathbb{X}_k)\left\|\theta_1-\theta_2\right\|, \quad \forall \theta_1,\theta_2 \in N,$$

where under $\mathcal{H}_0$, N is a neighborhood of θ_0, and under $\mathcal{H}_1$, N is a neighborhood of θ_*. Moreover, it holds that

$$\sup_{k=1,2,\ldots} \mathbb{E}[(K^{a,(i,j)}(\mathbb{X}_k))^2\sigma^2(\mathbb{X}_k,\theta_{(k)})] < \infty$$

and that

$$\sup_{k=1,2,\ldots} \mathbb{E}[(K^{b,(i,j)}(\mathbb{X}_k))^2\sigma^4(\mathbb{X}_k,\theta_{(k)})] < \infty.$$

With the conditions above, the following proposition holds. Its proof is basic, so we omit.

Proposition 6.2.1. *Suppose the conditions (I)-(III). Fix $\theta \in \{\theta_0,\theta_1,\theta_*\}$.*
(i) (D7) holds, that is it holds that

$$\sup_{k=1,2,\ldots} \mathbb{E}\left[\frac{\|\dot{S}(\mathbb{X}_k,\theta_*)\|^2}{\sigma^4(\mathbb{X}_k,\theta_*)}\left((S(\mathbb{X}_k,\theta_{(k)})-S(\mathbb{X}_k,\theta))^2+\sigma^2(\mathbb{X}_k,\theta)\right)\right] < \infty$$

and that

$$\begin{aligned}\sup_{k=1,2,\ldots} \mathbb{E}\Bigg[&\frac{\|\dot{\sigma}(\mathbb{X}_k,\theta_*)\|^2}{\sigma^6(\mathbb{X}_k,\theta_*)}\big((S(\mathbb{X}_k,\theta_{(k)})-S(\mathbb{X}_k,\theta))^4 \\ &+4((S(\mathbb{X}_k,\theta_{(k)})-S(\mathbb{X}_k,\theta))^2\sigma^2(\mathbb{X}_k,\theta)+2\sigma^4(\mathbb{X}_k,\theta)\big)\Bigg] < \infty.\end{aligned}$$

(ii) (D8) holds, that is, it holds that

$$\sup_{k=1,2,\ldots} \mathbb{E}\left[\frac{\left\|\dot{S}(\mathbb{X}_k,\theta)\right\|^4}{\sigma^4(\mathbb{X}_k,\theta)}\right] < \infty$$

and that

$$\begin{aligned}\sup_{k=1,2,\ldots} \mathbb{E}\Bigg[&\frac{\|\dot{\sigma}(\mathbb{X}_k,\theta_*)\|^2}{\sigma^6(\mathbb{X}_k,\theta_*)}\Big(\|\dot{S}(\mathbb{X}_k,\theta)\|^2(S(\mathbb{X}_k,\theta_{(k)})-S(\mathbb{X}_k,\theta))^2 \\ &+\sigma^2(\mathbb{X}_k,\theta)(\|\dot{S}(\mathbb{X}_k,\theta)\|^2+\|\dot{\sigma}(\mathbb{X}_k,\theta))\|^2 \\ &+2\dot{S}(\mathbb{X}_k,\theta)^\top\dot{\sigma}(\mathbb{X}_k,\theta)(S(\mathbb{X}_k,\theta_{(k)})-S(\mathbb{X}_k,\theta))\sigma(\mathbb{X}_k,\theta)\Big)\Bigg] < \infty.\end{aligned}$$

(iii) (D9) holds for $\delta = 1$, that is, it holds that

$$\sup_{k=1,2,\ldots} \mathbb{E}\left[\frac{\left\|\dot{S}(\mathbb{X}_k,\theta)\right\|^4}{\sigma^8(\mathbb{X}_k,\theta)}\sigma^2(\mathbb{X}_k,\theta_{(k)})\right] < \infty$$

and that

$$\sup_{k=1,2,\ldots} \mathbb{E}\left[\frac{\|\dot{\sigma}(\mathbb{X}_k,\theta)\|^4}{\sigma^{12}(\mathbb{X}_k,\theta)}\sigma^4(\mathbb{X}_k,\theta_{(k)})\right] < \infty.$$

(iv) (D10) holds. Especially, by the Proposition 6.1.1, it holds that

$$\sup_{k=1,2,\ldots} \mathbb{E}\left[\frac{\left\|\dot{S}(\mathbb{X}_k,\theta)\right\|^4}{\sigma^8(\mathbb{X}_k,\theta)}\sigma^4(\mathbb{X}_k,\theta_{(k)})\right] < \infty$$

and that

$$\sup_{k=1,2,\ldots} \mathbb{E}\left[\frac{\|\dot{\sigma}(\mathbb{X}_k,\theta)\|^4}{\sigma^{12}(\mathbb{X}_k,\theta)}\sigma^8(\mathbb{X}_k,\theta_{(k)})\right] < \infty.$$

(v) (D12) holds, that is, it holds that

$$\sup_{k=1,2\ldots} \mathbb{E}\left[\frac{\|\partial_i\dot{S}(\mathbb{X}_k,\theta)\|^2}{\sigma^4(\mathbb{X}_k,\theta)}\sigma^2(\mathbb{X}_k,\theta_{(k)})\right] < \infty$$

and that

$$\sup_{k=1,2,\ldots} \mathbb{E}\left[\frac{\|\partial_i\dot{\sigma}(\mathbb{X}_k,\theta)\sigma(\mathbb{X}_k,\theta) - 3\partial_i\sigma(\mathbb{X}_k,\theta)\dot{\sigma}(\mathbb{X}_k,\theta)\|^2}{\sigma^8(\mathbb{X}_k,\theta)}\sigma^4(\mathbb{X}_k,\theta_{(k)})\right] < \infty.$$

for all $i = 1,\ldots,d$.

Now, let us introduce the test statistic

$$AD_n = \sum_{j=1}^{n-1}\frac{nU_{n,j}^\top\hat{C}_n^{-1}U_{n,j}}{j(n-j)},$$

where

$$
\begin{aligned}
j \rightsquigarrow U_{n,j} &= \frac{1}{\sqrt{n}}\sum_{k=1}^{j}\frac{\dot{S}(\mathbb{X}_k,\hat{\theta}_n)}{(\sigma(\mathbb{X}_k,\hat{\theta}_n))^2}\xi_k^a(\hat{\theta}_n)+\frac{1}{n}\sum_{k=1}^{j}\frac{\dot{\sigma}(\mathbb{X}_k,\hat{\theta}_n)}{(\sigma(\mathbb{X}_k,\hat{\theta}_n))^3}\xi_k^b(\hat{\theta}_n),\\
\hat{C}_n &= \frac{1}{n}\sum_{k=1}^{n}\frac{\dot{S}(\mathbb{X}_k,\hat{\theta}_n)\dot{S}(\mathbb{X}_k,\hat{\theta}_n)^\top}{(\sigma(\mathbb{X}_k,\hat{\theta}_n))^2(\sigma(\mathbb{X}_k,\hat{\theta}_n))^2}(\sigma(\mathbb{X}_k,\hat{\theta}_n))^2\\
&\quad +\frac{1}{n}\sum_{k=1}^{n}\frac{\dot{\sigma}(\mathbb{X}_k,\hat{\theta}_n)\dot{\sigma}(\mathbb{X}_k,\hat{\theta}_n)^\top}{(\sigma(\mathbb{X}_k,\hat{\theta}_n))^3(\sigma(\mathbb{X}_k,\hat{\theta}_n))^3}(\sigma(\mathbb{X}_k,\hat{\theta}_n))^4.
\end{aligned}
$$

It holds that

$$
\frac{U_{n,[nu]}}{\sqrt{[nu]/n(1-[nu]/n)}}=\mathbb{Z}_n(u,\hat{\theta}_n)+o_P(1),
$$

where

$$
\begin{aligned}
(u,\theta) &\rightsquigarrow \mathbb{Z}_n(u,\theta)\\
&= \frac{1}{\sqrt{n}}\sum_{k=1}^{n}w_k^n(u)\left(\frac{\dot{S}(\mathbb{X}_k;\theta)}{(\sigma(\mathbb{X}_k;\theta))^2}\xi_k^a(\theta)+\frac{\dot{\sigma}(\mathbb{X}_k;\theta)}{(\sigma(\mathbb{X}_k;\theta))^3}\xi_k^b(\theta)\right).
\end{aligned}
$$

With the setting above, the following proposition directly from the Theorem 6.1.1.

Theorem 6.2.1. *Assume conditions (I)-(VII). Under $\mathcal{H}_0$, it holds that $AD_n \to^d \|G\|_{L^2}^2$ as $n\to\infty$ and under $\mathcal{H}_1$, the test is consistent.*

Chapter 7

Other topics

7.1 Z-process method and likelihood ratio process method for independent data

Recall that the notation in Section 3.2. In order to test change point hypotheses, it can be used the KS (Kolmogorov-Smirnov) type, the CM (Crámer-von Mises) type and the AD type test statistics written as follows:

$$KS = \sup_{u \in (0,1)} |n\Psi_n^\circ(u, \hat{\theta}_n)^\top \hat{I}_n^{-1} \Psi_n^\circ(u, \hat{\theta}_n)|,$$

$$CM = \int_0^1 n\Psi_n^\circ(u, \hat{\theta}_n)^\top \hat{I}_n^{-1} \Psi_n^\circ(u, \hat{\theta}_n) du$$

and

$$AD = \int_0^1 \frac{n\Psi_n^\circ(u, \hat{\theta}_n)^\top \hat{I}_n^{-1} \Psi_n^\circ(u, \hat{\theta}_n)}{s_n(u)(1 - s_n(u))} du.$$

Now we compare the suprema type statistics of Z-process method and likelihood ratio test, which is considered as a standard method for parametric change point problems, from the viewpoint of local asymptotic powers when we consider likelihood equations as estimating equations. Only this section, let us set a local alternative hypothesis

$$\mathcal{H}_1 : \theta_0 \neq \theta_1 = \theta_0 + v_n^{-1} h,$$

where $v_n \to \infty$ and $n^{-1/2} v_n \to 0$ as $n \to \infty$. Moreover, we define h_0 and h_1 as follows: $h_0 = v_n(\theta_* - \theta_0) = h(1 - u_*)$ and $h_1 = v_n(\theta_1 - \theta_*) = hu_*$, respectively. We consider a change point estimator $\hat{k} = n\hat{u}$ which satisfy

$$\hat{u} - u_* = O_p(n^{-1}v_n^2);$$

we refer to, for example, Bhattacharya (1987) and Gombay and Horváth (1996) for the construction of such estimators and proofs. So, it is enough to calculate the asymptotic distribution of

$$\sqrt{n}(\Psi_n^\circ(u_*, \hat{\theta}_n)^\top \hat{I}_n^{-1} \Psi_n^\circ(u_*, \hat{\theta}_n) - \mu_n^\top I_{\theta_*}^{-1} \mu_n),$$

where

$$\mu_n = u_*(1 - u_*)(\mathbb{E}_{\theta_0}[\dot{l}_{\theta_*}] - \mathbb{E}_{\theta_1}[\dot{l}_{\theta_*}]).$$

The Taylor expansion yields that

$$\mathbb{E}_{\theta_0}[\dot{l}_{\theta_*}(X_1)] = \mathbb{E}_{\theta_* - v_n^{-1} h_0}[\dot{l}_{\theta_*}(X_1)] = -v_n^{-1} I_{\theta_*} h_0 + o(v_n^{-1})$$

and

$$\mathbb{E}_{\theta_1}[\dot{l}_{\theta_*}(X_1)] = \mathbb{E}_{\theta_* + v_n^{-1} h_1}[\dot{l}_{\theta_*}(X_1)] = v_n^{-1} I_{\theta_*} h_1 + o(v_n^{-1}),$$

so it holds that

$$\mu_n = -v_n^{-1} u_*(1 - u_*) I_{\theta_*} h + o(v_n^{-1}) = O(v_n^{-1}).$$

Let us derive the asymptotic distribution of

$$\sqrt{n}(\hat{I}_n^{-\frac{1}{2}} \Psi_n^\circ(u_*, \hat{\theta}_n) - I_{\theta_*}^{-\frac{1}{2}} \mu_n) = \sqrt{n} \hat{I}_n^{-\frac{1}{2}} (\Psi_n^\circ(u_*, \hat{\theta}_n) - \mu_n) + \sqrt{n}(\hat{I}_n^{-\frac{1}{2}} - I_{\theta_*}^{-\frac{1}{2}}) \mu_n.$$

The second term converges to zero in probability. As for the first term, the central limit theorem and Slutsky's theorem yield that

$$\sqrt{n} \hat{I}_n^{-\frac{1}{2}} (\Psi_n^\circ(u_*, \hat{\theta}_n) - \mu_n) \to^d N\left(0, u_*(1 - u_*)\right).$$

This is true because $\Psi_n^\circ(u_*, \hat{\theta}_n)$ can be replaced by $\Psi_n^\circ(u_*, \theta_*)$ and the variance of $\sqrt{n}\Psi_n^\circ(u_*, \theta_*)$ is

$$u_*(1-u_*)((1-u_*)Var_{\theta_* - v_n^{-1} h_0}[\dot{l}_{\theta_*}(X_1)] + u_* Var_{\theta_* + v_n^{-1} h_1}[\dot{l}_{\theta_*}(X_1)]) \to u_*(1-u_*) I_{\theta_*}$$

as $n \to \infty$. Hence, the delta method yields that

$$\frac{\sqrt{n}(\Psi_n^\circ(u_*, \hat{\theta}_n)^\top \hat{I}_n^{-1} \Psi_n^\circ(u_*, \hat{\theta}_n) - \mu_n^\top I_{\theta_*}^{-1} \mu_n)}{\sqrt{4 \mu_n^\top I_{\theta_*}^{-1} \mu_n}} \to^d N\left(0, u_*(1 - u_*)\right),$$

and, since $\hat{u} - u_* = O_p(n^{-1}v_n^2)$, it leads that

$$\frac{KS - \mu_{(Z)}}{\sigma_{(Z)}} \to^d N(0,1),$$

where

$$\mu_{(Z)} = nv_n^{-2} \lim_{n\to\infty} v_n^2 \mu_n^\top I_{\theta_*}^{-1} \mu_n,$$

and

$$\sigma_{(Z)}^2 = 4nv_n^{-2} u_*(1-u_*) \lim_{n\to\infty} v_n^2 \mu_n^\top I_{\theta_*}^{-1} \mu_n.$$

On the other hand, a likelihood ratio statistic is that

$$LR = \sup_u [u(1-u)(-2\log \Lambda_{[nu]})],$$

where

$$\Lambda_{[nu]} = \frac{\sup_\theta \prod_{k=1}^n f(X_k,\theta)}{\sup_\theta \prod_{k=1}^{[nu]} f(X_k,\theta) \sup_\theta \prod_{k=[nu]+1}^n f(X_k,\theta)}. \tag{7.1.1}$$

For likelihood ratio methods, see, for example, Gombay and Horváth (1994, 1996). It holds that

$$\frac{LR - \mu_{(LR)}}{\sigma_{(LR)}} \to^d N(0,1),$$

where $\mu_{(LR)}$ and $\sigma_{(LR)}^2$ is $2u_*(1-u_*)\mu^*$ and $4nu_*^2(1-u_*)^2\delta^2\sigma_2^2$ in Gombay and Horváth (1996), respectively, if v_n satisfies $nv_n^{-2}(\log\log n)^{-1} \to \infty$ as $n \to \infty$. Hereafter, this condition is assumed. Under $\mathcal{H}_0$, LR and KS converges to the same limit if we consider likelihood equations as estimating equations. Under $\mathcal{H}_1$, both statistics are thought to compete well asymptotically, at least in the following examples.

(i) Normal observations, change in the mean when the variance is unity. The probability density function is

$$f(x,\theta) = \exp\left(-\frac{(x-\theta)^2}{2} - \frac{1}{2}\log(2\pi)\right).$$

It holds that

$$\mu_{(Z)} = \mu_{(LR)} = nv_n^{-2} u_*^2 (1-u_*)^2 h^2$$

and

$$\sigma_{(Z)}^2 = \sigma_{(LR)}^2 = 4nv_n^{-2} u_*^3 (1-u)^3 h^2.$$

If we use $\sum -\ddot{l}_{\hat{\theta}_n}(X_k)/n$ as $\hat{I}_n$, both test statistics are identical.

(ii) Exponential observations. The probability density function is

$$f(x,\theta) = \exp(-\theta x + \log\theta)1\{x \geq 0\}.$$

It holds that

$$\mu_{(Z)} = nv_n^{-2}u_*^2(1-u_*)^2h^2/\theta_0^2$$

and

$$\sigma_{(Z)}^2 = \sigma_{(LR)}^2 = 4nv_n^{-2}u_*^3(1-u_*)^3h^2/\theta_0^2.$$

On the other hand, it holds that

$$\mu_{(LR)} = 2nu_*(1-u_*)(-u_*\log(1+v_n^{-1}h/\theta_0) + \log(1+u_*v_n^{-1}h/\theta_0)) \approx \mu_{(Z)}.$$

The last approximation is $\log(1+x) \approx x - x^2/2$ for small x.

(iii) Poisson observations. The probability mass function in the natural form is

$$f(x,\theta) = \exp(\theta x - \exp(\theta) - \log x!)1\{x \text{ is nonnegative integer}\}.$$

It holds that

$$\mu_{(Z)} = nv_n^{-2}u_*^2(1-u_*)^2\exp(\theta_0)h^2$$

and that

$$\sigma_{(Z)}^2 = 4nv_n^{-2}u_*^3(1-u_*)^3\exp(\theta_0)h^2.$$

On the other hand, it holds that

$$\begin{aligned}\mu_{(LR)} &= 2nu_*(1-u_*)\exp(\theta_0)[(\exp(v_n^{-1}h)-1)v_n^{-1}h - u_* \\ &\quad -(1-u_*)\exp(v_n^{-1}h) + \exp((1-u_*)v_n^{-1}h)] \\ &\approx \mu_{(Z)}\end{aligned}$$

and that

$$\sigma_{(LR)}^2 = 4nu_*^3(1-u_*)^3\exp(\theta_0)(1-\exp(v_n^{-1}h))^2 \approx \sigma_{(Z)}^2.$$

The two approximations here are $\exp(x) \approx 1 + x - x^2/2$ for small x.

Two local asymptotic powers are identical in these cases. Numerical comparisons by Monte Carlo simulations will be done in the next section.

7.2 Monte Carlo simulations

From the viewpoints of approximation accuracies under $\mathcal{H}_0$ and powers under $\mathcal{H}_1$, we compare the tests with the test statistics KS_n, CM_n and AD_n defined by

$$\begin{aligned} KS_n &= n \max_{j=1,\dots,n-1} \Phi_{n,j}(\hat{\theta}_n)^\top \hat{I}_n^{-1} \Phi_{n,j}(\hat{\theta}_n) \\ CM_n &= \sum_{j=1}^{n-1} \Phi_{n,j}(\hat{\theta}_n)^\top \hat{I}_n^{-1} \Phi_{n,j}(\hat{\theta}_n), \\ AD_n &= \sum_{j=1}^{n-1} \frac{n^2 \Phi_{n,j}(\hat{\theta}_n)^\top \hat{I}_n^{-1} \Phi_{n,j}(\hat{\theta}_n)}{j(n-j)}, \end{aligned}$$

where $\hat{I}_n = -\sum_{k=1}^n \ddot{l}_{\hat{\theta}_n}(X_k)/n$, by Monte Carlo simulation In addition, we compare these statistics based on Z-process methods and the corresponding three types of statistics based on likelihood ratio methods:

$$\begin{aligned} LR_n^{KS} &= \max_{j=1,\dots,n-1} j(n-j)(-2\log\Lambda_j)/n^2, \\ LR_n^{CM} &= \sum_{j=1}^{n-1} j(n-j)(-2\log\Lambda_j)/n^3, \\ LR_n^{AD} &= \sum_{j=1}^{n-1} -2\log\Lambda_j/n, \end{aligned}$$

where $\{\Lambda_\cdot\}$ is (7.1.1). Under $\mathcal{H}_0$, (A) KS_n and LR_n^{KS}, (B) CM_n and LR_n^{CM}, and (C) AD_n and LR_n^{AD} converge to

$$\begin{aligned} &\text{(A)} \quad \sup_{u\in(0,1)} \|B_d^\circ(u)\|^2, \\ &\text{(B)} \quad \int_0^1 \|B_d^\circ(u)\|^2\, du, \\ &\text{(C)} \quad \int_0^1 \left\| \frac{B_d^\circ(u)}{\sqrt{u(1-u)}} \right\|^2 du \end{aligned}$$

in distribution, respectively. Moreover, under $\mathcal{H}_1$, these tests are consistent.

Let us consider (i) normal observations, change in the mean when the variance is unity; (ii) exponential observations; (iii) Poisson observations;

(iv) observations generated by the following simple regression model: $y_k = \theta^{(0)} + \theta^{(1)} x_k + \varepsilon_k$, $x_k \sim N(0,1)$, $\varepsilon_k \sim N(0,\sigma^2)$, where x_k and ε_k $(k = 1,\ldots,n)$ are mutually independent random variables. Observed variables are x_k and y_k, and ε_k is unobserved. The parameter σ^2 is a nuisance parameter and we will consider tests for $\theta = (\theta^{(0)}, \theta^{(1)})$.

Firstly, under $\mathcal{H}_0$, we evaluate approximation accuracies of six tests. By setting the sample sizes n to 10, 50, 100, and 500 and the significance level α to 0.10, we count the number of type I errors in 5000 times trial. Models and their parameters are set as follows: (i) normal observations with unit variance, $\theta_0 = 0$; (ii) exponential observations, $\theta_0 = 1$; (iii) Poisson observations, $\theta_0 = \log 5$; (iv) regression, $\theta_0 = (1,1)$ and $\sigma^2 = 1$. The type I error rates are tabulated in Table 7.1. This result shows that the type I errors of the CM type and the AD type statistics are almost the same as values of α and not seen much change due to n, especially when it is larger than 50. On the other hand, the tests by the KS type statistics are too conservative with practical sample size such as 10, 50 or 100. Comparing likelihood ratio and Z-process methods, in case (ii), the results are contrasting: the CM type and the AD type statistics of likelihood ratio methods are liberal and Z-process methods are conservative with n up to 100. In cases (iii) and (iv), the results are similar.

Secondly, under $\mathcal{H}_1$, let us compare powers of six tests. The combination of parameters are set to $(\theta_0, \theta_1) =$(i)$(0,1)$, (ii)$(1,2)$, (iii)$(\log 5, \log 6)$, (iv-a) $((1,1),(2,1))$, (iv-b) $((1,1),(1,2))$, and (iv-c) $((1,1),(1+1/\sqrt{2}, 1+1/\sqrt{2}))$, n to 100, α to 0.1 and we repeat 1000 times. The change point, it is denoted by $[nu_*]$, varies in 5, 10, 20, 30, 40, 50, 60, 70, 80, 90 and 95. The powers calculated from the trials are tabulated in Table 7.2-7.3. These results show AD type statistics have remarkable power when a change point is close to the first or the last sample and tolerable power otherwise. Comparing likelihood ratio and Z-process methods in various cases, it does not seem possible to state that one method is always better than the other.

7.3 A test for a raw moment change

By a similar way as Section 3.2, we can treat other change point tests. In this section, we shall consider a test for a change in moments.

Let $X_1, \ldots, X_n$ be mutually independent random variables and we consider detecting a change in the moments up to $r (r \in \mathbb{N})$ th moment in the

Table 7.1: Simulation results under $\mathcal{H}_0$: (i) normal observations, $\theta_0 = 0$; (ii) exponential observations, $\theta_0 = 1$; (iii) Poisson observations, $\theta_0 = \log 5$; (iv) regression, $\theta_0 = (1, 1)$

	n	LR^{KS}	LR^{CM}	LR^{AD}	KS	CM	AD
(i)	10	0.038	0.097	0.085	0.038	0.097	0.085
	50	0.069	0.102	0.101	0.069	0.102	0.101
	100	0.089	0.111	0.109	0.089	0.111	0.109
	500	0.087	0.098	0.095	0.087	0.098	0.095
(ii)	10	0.046	0.118	0.109	0.028	0.085	0.075
	50	0.070	0.108	0.109	0.067	0.100	0.101
	100	0.076	0.104	0.104	0.071	0.098	0.100
	500	0.084	0.101	0.100	0.085	0.100	0.100
(iii)	10	0.034	0.102	0.092	0.033	0.099	0.087
	50	0.073	0.104	0.102	0.071	0.104	0.102
	100	0.074	0.101	0.099	0.073	0.101	0.097
	500	0.090	0.104	0.104	0.090	0.104	0.104
(iv)	10	0.007	0.102	0.036	0.002	0.088	0.074
	50	0.056	0.110	0.098	0.050	0.097	0.101
	100	0.069	0.095	0.093	0.066	0.092	0.096
	500	0.095	0.110	0.109	0.094	0.111	0.111

following formulation of the following hypothetical testing:

$\mathcal{H}_0$: $\mathbb{E}[X_k^i] = \mu_0^i \ \forall k = 1, \ldots, n, \ \forall i = 1 \ldots, r$
$\mathcal{H}_1$: $\exists u_* \in (0,1), \ \exists i \in \{1, 2, \ldots, r\}$ such that $\mathbb{E}[X_k^i] = \mu_0^i$ for $k = 1, \ldots, [nu_*]$ and $\mathbb{E}[X_k^i] = \mu_1^i$ for $k = [nu_*] + 1, \ldots, n$

We assume the existence of up to $2r + 1$ th moments. For i th moment, the discrete time stochastic process $(\phi_{nj}^{(i)})_{j=1,\ldots,n-1}$ is given by

$$\phi_{nj}^{(i)} = \frac{1}{n^2} \left\{ (n - j) \sum_{k=1}^{j} X_k^i - j \sum_{k=j+1}^{n} X_k^i \right\}.$$

The r-dimension vector $\Phi_{n,j}$ is defined by $\phi_{nj}^{(i)}$ by

$$\Phi_{n,j} := (\phi_{nj}^{(1)}, \ldots, \phi_{nj}^{(r)})^\top.$$

Table 7.2: Simulation results under $\mathcal{H}_1$: (θ_0, θ_1) = (i) $(0,1)$, (ii) $(1,2)$ and (iii) $(\log 5, \log 6)$

	$[nu_*]$	LR^{KS}	LR^{CM}	LR^{AD}	KS	CM	AD
	5	0.150	0.207	0.282	0.150	0.207	0.282
	10	0.445	0.531	0.644	0.445	0.531	0.644
	20	0.917	0.919	0.940	0.917	0.919	0.940
	30	0.989	0.991	0.987	0.989	0.991	0.987
	40	1.000	0.998	0.997	1.000	0.998	0.997
(i)	50	0.997	1.000	1.000	0.997	1.000	1.000
	60	0.994	0.996	0.996	0.994	0.996	0.996
	70	0.992	0.988	0.987	0.992	0.988	0.987
	80	0.907	0.916	0.936	0.907	0.916	0.936
	90	0.443	0.503	0.616	0.443	0.503	0.616
	95	0.145	0.214	0.298	0.145	0.214	0.298
	5	0.135	0.211	0.251	0.182	0.233	0.320
	10	0.271	0.402	0.443	0.356	0.426	0.511
	20	0.634	0.712	0.730	0.701	0.738	0.757
	30	0.821	0.858	0.848	0.847	0.868	0.859
	40	0.898	0.924	0.905	0.902	0.923	0.909
(ii)	50	0.902	0.919	0.898	0.897	0.915	0.895
	60	0.893	0.923	0.906	0.875	0.905	0.872
	70	0.824	0.835	0.838	0.747	0.793	0.783
	80	0.581	0.575	0.633	0.418	0.474	0.509
	90	0.182	0.240	0.302	0.112	0.190	0.203
	95	0.110	0.147	0.171	0.094	0.132	0.133
	5	0.097	0.120	0.134	0.094	0.118	0.124
	10	0.126	0.185	0.201	0.115	0.180	0.192
	20	0.271	0.309	0.320	0.256	0.300	0.310
	30	0.433	0.486	0.488	0.416	0.481	0.475
	40	0.545	0.593	0.566	0.537	0.590	0.564
(iii)	50	0.574	0.614	0.584	0.568	0.616	0.580
	60	0.505	0.560	0.547	0.509	0.565	0.553
	70	0.422	0.487	0.469	0.427	0.490	0.477
	80	0.274	0.358	0.384	0.290	0.363	0.398
	90	0.150	0.179	0.208	0.154	0.183	0.217
	95	0.095	0.123	0.129	0.094	0.126	0.136

Table 7.3: Simulation results under $\mathcal{H}_1$: $(\theta_0, \theta_1) =$ (iv-a) $((1,1),(2,1))$, (iv-b) $((1,1),(1,2))$ and (iv-c) $((1,1),(1+1/\sqrt{2}, 1+1/\sqrt{2}))$

	$[nu_*]$	LR^{KS}	LR^{CM}	LR^{AD}	KS	CM	AD
	5	0.092	0.155	0.200	0.092	0.156	0.213
	10	0.249	0.374	0.485	0.242	0.364	0.497
	20	0.786	0.810	0.842	0.781	0.804	0.853
	30	0.954	0.956	0.954	0.952	0.953	0.955
	40	0.991	0.992	0.983	0.990	0.990	0.984
(iv-a)	50	0.990	0.994	0.988	0.989	0.995	0.987
	60	0.995	0.992	0.988	0.995	0.992	0.987
	70	0.952	0.952	0.954	0.950	0.953	0.952
	80	0.809	0.829	0.864	0.801	0.830	0.869
	90	0.243	0.374	0.461	0.257	0.375	0.482
	95	0.102	0.177	0.213	0.100	0.168	0.219
	5	0.100	0.184	0.237	0.103	0.193	0.269
	10	0.250	0.376	0.450	0.263	0.381	0.482
	20	0.716	0.763	0.795	0.672	0.733	0.765
	30	0.944	0.951	0.948	0.930	0.938	0.932
	40	0.986	0.987	0.982	0.979	0.984	0.976
(iv-c)	50	0.987	0.987	0.981	0.985	0.982	0.977
	60	0.986	0.990	0.980	0.981	0.984	0.976
	70	0.938	0.944	0.939	0.918	0.931	0.929
	80	0.769	0.805	0.833	0.728	0.790	0.812
	90	0.255	0.363	0.446	0.256	0.364	0.463
	95	0.095	0.176	0.234	0.105	0.182	0.271
	5	0.114	0.178	0.223	0.114	0.183	0.253
	10	0.253	0.364	0.444	0.260	0.370	0.464
	20	0.773	0.774	0.815	0.747	0.754	0.796
	30	0.942	0.954	0.952	0.922	0.940	0.937
	40	0.985	0.983	0.979	0.982	0.982	0.975
(iv-c)	50	0.986	0.988	0.985	0.985	0.989	0.985
	60	0.983	0.984	0.973	0.980	0.980	0.971
	70	0.942	0.936	0.941	0.925	0.928	0.932
	80	0.789	0.802	0.835	0.758	0.787	0.815
	90	0.282	0.373	0.472	0.276	0.390	0.481
	95	0.127	0.199	0.247	0.126	0.204	0.253

If there is no change in moments, $\mathbb{E}[\Phi_{n,j}] = 0$. The covariance matrix of

$$\left(\frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k, \ldots, \frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k^r \right)^\top$$

is denoted by Σ, that is, whose r_1-th row and r_2-th column element is

$$(\Sigma)_{(r_1, r_2)} = \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} Cov(X_k^{r_1}, X_k^{r_2})$$

We can calculate the covariance of $X_k^{r_1}$ and $X_k^{r_2}$ $(r_1, r_2 \in (1, \ldots, r))$ by

$$Cov(X_k^{r_1}, X_k^{r_2}) = \mathbb{E}[X_k^{r_1 + r_2}] - \mathbb{E}[X_k^{r1}]\mathbb{E}[X_k^{r2}].$$

The estimator $\hat{\Sigma}_n$ of Σ is a sample covariance matrix whose r_1-th row and r_2-th column element is calculated by

$$(\hat{\Sigma}_n)_{(r_1, r_2)} = \frac{1}{n} \sum_{k=1}^{n} X_k^{r_1 + r_2} - \frac{1}{n} \sum_{k=1}^{n} X_k^{r_1} \frac{1}{n} \sum_{k=1}^{n} X_k^{r_2}.$$

Then, we propose the following test statistic:

$$MM_n := \sum_{j=1}^{n-1} \frac{n^2 \Phi_{n,j}^\top \hat{\Sigma}_n^{-1} \Phi_{n,j}}{j(n-j)}.$$

In this problem setting, we have the following theorem.

Theorem 7.3.1. *(i) Under $\mathcal{H}_0$, the asymptotic distribution of MM_n is*

$$\int_0^1 \left\| \frac{B_d^\circ(u)}{\sqrt{u(1-u)}} \right\|^2 du.$$

(ii) Under $\mathcal{H}_1$, the test is consistent.

Remark 7.3.1. *In this case, the pinned Z-process is*

$$\Psi_n^\circ(u) = (\psi_{n,1}^\circ(u), \ldots, \psi_{n,r}^\circ(u))^\top,$$

where

$$\psi^{\circ}_{n,i}(u) = \frac{1}{n}\left[(1-s_n(u))\sum_{k=1}^{ns_n(u)} X_k^i - s_n(u)\sum_{k=ns_n(u)+1}^{n} X_k^i\right], \quad u \in (0,1).$$

A direct computation yields that

$$MM_n = \int_0^1 \frac{n\Psi^{\circ}_n(u)^{\top}\hat{\Sigma}_n^{-1}\Psi^{\circ}_n(u)}{s_n(u)(1-s_n(u))}du.$$

We can derive the asymptotic distribution and prove the consistency for MM_n *by the similar way as before.*

Part II

Applications to functional limit theorems for random combinatorial structures

Chapter 8

Introduction

Consider non negative integer-valued random variables $(C_1^n, C_2^n, \ldots, C_n^n)$ which satisfy $C_1^n + 2C_2^n + \ldots + nC_n^n = n$, that is to say, C_j^n is the number of components whose size is j of a random partition of a given natural number n. In many cases, it is known that finite dimensional marginals of $(C_1^n, C_2^n, \ldots)$ converges in distribution to corresponding ones of $(Z_1, Z_2, \ldots)$, where Z_j's are mutually independent Poisson random variables. For example, in case of the Ewens sampling formula with the parameter θ, such Poisson variables satisfy

$$\mathbb{E}[Z_j] = \frac{\theta}{j}, \quad j = 1, \ldots, n$$

while in case of n^n uniform random mappings from the set $\{1, \ldots, n\}$ to itself they satisfy

$$\mathbb{E}[Z_j] = \frac{e^{-j}}{j} \sum_{i=0}^{j-1} \frac{j^i}{i!}, \quad j = 1, \ldots, n.$$

Such a weak convergence ensures that, for any fixed b,

$$(C_1^n, \ldots, C_b^n) \to^d (Z_1, \ldots, Z_b)$$

as $n \to \infty$. In some cases, it is important to establish this kind of approximation in the case where $b = b(n)$ gets large as n increases. Indeed, it has been recognized to be interesting to prove functional CLTs

$$u \rightsquigarrow \frac{\sum_{j=1}^{[n^u]} C_j^n - \theta u \log(n)}{\sqrt{\theta \log(n)}} \to^d B \quad \text{in } D[0,1]$$

as $n \to \infty$, where D denotes the Skorokhod space, which is the space of càdlàg functions, and B is the standard Brownian motion. In this part, we shall prove new functional CLTs not in $D[0,1]$ but in $L^2([0,1], du)$. To be more specific, we will prove the weak convergence

$$u \rightsquigarrow \frac{\sum_{j=1}^{[n^u]} C_j^n - \sum_{j=1}^{[n^u]} \mathbb{E}[Z_j]}{\sqrt{\sum_{j=1}^{[n^u]} \mathbb{E}[Z_j]}} \to^d G \quad \text{in } L^2([0,1], du),$$

where $u \rightsquigarrow G(u) = B(u)/\sqrt{u}$, for both the number of partitions of the Ewens sampling formula and random mappings through showing Poisson process approximations are still working well in the current settings.

Functional CLTs in D were originally proved by Hansen (1990) for the Ewens sampling formula and by Hansen (1989) for random mappings through the direct ways to check the tightnesses and convergences of finite dimensional marginal distributions (about the weak convergence theory in the Skorokhod spaces, see the well-known book: Billingsley (1999)), though DeLaurentis and Pittel (1985) had proved a functional CLT for random permutations earlier. After that, they were proved by different, elegant ways via Poisson process approximations by Arratia and Tavaré (1992). Arratia et al. (2000) proved that it is possible to apply such approaches to general problem settings. See Arratia et al. (2003) for overall arguments in this field.

Here, let us recall an sophisticated, unified approach by Arratia et al. (2000) to treat asymptotic behavior of general logarithmic structures, which satisfy the conditioning relation

$$\mathbb{P}[C_1^n = c_1, \ldots, C_n^n = c_n] = \mathbb{P}\left[Z_1 = c_1, \ldots, Z_n = c_n \,\middle|\, \sum_{j=1}^{n} jZ_j = n \right]$$

and the logarithmic condition

$$\lim_{j \to \infty} j\mathbb{P}[Z_j = 1] = \lim_{j \to \infty} j\mathbb{E}[Z_j] = \theta,$$

where $\{Z_\cdot\}$ are not necessarily Poisson variables. They used the total variation distances to prove limit theorems including functional CLTs. The Ewens sampling formula and random mappings, which are argued in this part, satisfy these conditions. We believe that two important problems which are studied in L^2 spaces in this part would be the prototypes to open the new windows to undiscovered problems with general logarithmic structures.

Chapter 9

The Ewens sampling formula

9.1 The result

Let us introduce the Ewens sampling formula with some results proven by Arratia et al. (1992), see also Arratia and Tavaré (1992). Introduce the sequence $\{\xi_k\}_{k=1}^n$ of Bernoulli random variables whose distribution is given by

$$\mathbb{P}[\xi_j = 1] = p_j = \frac{\theta}{\theta + j - 1}, \quad j = 1, \ldots, n,$$

where $\theta \in \mathbb{R}$. Define C_j^n by

$$C_j^n = \begin{cases} \sum_{i=1}^{n-1} \xi_i(1-\xi_{i+1})\cdots(1-\xi_{i+j-1})\xi_{i+j} + \xi_n, & (j=1) \\ \sum_{i=1}^{n-j} \xi_i(1-\xi_{i+1})\cdots(1-\xi_{i+j-1})\xi_{i+j} & \\ \quad +\xi_{n-j+1}(1-\xi_{n-j+2})\cdots(1-\xi_n), & (1 < j \le n) \\ 0, & (j > n) \end{cases}$$

for $j = 1, 2, \ldots$ and define C_j^∞ by

$$C_j^\infty = \sum_{i=1}^{\infty} \xi_i(1-\xi_{i+1})\cdots(1-\xi_{i+j-1})\xi_{i+j}.$$

It holds that $\mathbb{E}[C_j^\infty] = \theta/j$, so it is almost surely finite for all j, and that

$$\sum_{j=1}^{n} C_j^n = \sum_{j=1}^{n} \xi_j.$$

The law of $\{C^n_\cdot\}_{j=1}^n$ is

$$\mathbb{P}[(C_1^n,\dots,C_n^n)=(c_1,\dots,c_n)]=\frac{n!}{(\theta)_n}\prod_{j=1}^n\left(\frac{\theta}{j}\right)^{c_j}\frac{1}{c_j!}1\left\{\sum_{j=1}^n jc_j=n\right\},$$

where $(\theta)_n$ denotes the rising factorial $\theta\times(\theta+1)\times\cdots\times(\theta+n-1)$. This formula is firstly derived by Ewens (1972) in the context of the population genetics and it is called the Ewens sampling formula. Consider a sequence $\{Z_\cdot\}$ of mutually independent Poisson variables such that

$$\mathbb{E}[Z_j]=\frac{\theta}{j},\quad j=1,2,\dots$$

and that $C_j^n\to^d C_j^\infty =^d Z_j$. A coupling

$$\sum_{j=1}^n\mathbb{E}[|C_j^n-Z_j|]=O(1)$$

can be constructed, and we fix this coupling through this chapter. For the expectation of the sum $\sum_{j=1}^n C_j^n$, it holds that

$$\mathbb{E}\left[\sum_{j=1}^n C_j^n\right]=\sum_{j=1}^n\frac{\theta}{\theta+j-1}$$

and

$$\theta\log\left(\frac{n}{\theta}\right)\le\mathbb{E}\left[\sum_{j=1}^n C_j^n\right]\le 1+\theta+\theta\log(n).$$

Here, it is known that the asymptotic normality

$$\frac{\sum_{j=1}^n C_j^n-\theta\log(n)}{\sqrt{\theta\log(n)}}\to^d N(0,1)\tag{9.1.1}$$

as $n\to\infty$ holds. This result is generalized to a functional CLT in the Skorokhod space $D[0,1]$ by Hansen (1990). That is to say, the standardized partial sum random field

$$u\rightsquigarrow\frac{\sum_{j=1}^{[n^u]}C_j^n-\theta u\log(n)}{\sqrt{\theta\log(n)}},\ 0\le u\le 1$$

converges in distribution to $u \rightsquigarrow B(u)$ in $D[0,1]$ by the usage of a limit theorem in $D[0,1]$. This statement subsumes the asymptotic normality (9.1.1) if we fix u to 1. This theorem has some other proofs: see, for example, Arratia and Tavaré (1992) and Donnelly et al. (1991b). Especially, Arratia and Tavaré (1992) proved it via a Poisson process approximation and basically we follow their approach. In the case of $\theta = 1$, which means random permutations, the functional CLT is proven by DeLaurentis and Pittel (1985) earlier.

The first goal of this part is to prove a new functional CLT described in the following theorem.

Theorem 9.1.1. *Define a random field*

$$u \rightsquigarrow X_n(u) = \frac{\sum_{j=1}^{[n^u]} C_j^n - \ell([n^u])}{\sqrt{\ell([n^u])}}, \quad 0 \leq u \leq 1$$

where

$$\ell(n) = \theta \sum_{j=1}^{n} \frac{1}{j}$$

for a natural number n. It holds that $X_n \to^d G$ in $L^2([0,1], du)$ and the limit G is

$$u \rightsquigarrow G(u) = \frac{B(u)}{\sqrt{u}}.$$

This theorem shall be proven in section 9.4 after the preparation argued in section 9.2 and section 9.3.

Remark 9.1.1. *It is conjectured to be true that the same property holds for other assemblies which satisfy the conditioning relation and the logarithmic condition argued in Arratia et al. (2000).*

The functional space $L^2([0,1], du)$, which is equivalence classes of square integrable functions on $[0,1]$ with respect to the Lebesgue measure du, is a typical example of a separable Hilbert space. For the limit theory in a separable Hilbert space, see section 1.8 of van der Vaart and Wellner (1996). The standardization in the theorem seems to be odd, since it is different from the standardization in Hansen's functional CLT, in which the partial sum is divided by $\sqrt{\theta \log(n)}$. The reason why we use $\ell(n)$ is that it is not clear that functional CLT also holds if we define a random field by

$$u \rightsquigarrow \frac{\sum_{j=1}^{[n^u]} C_j^n - \theta u \log(n)}{\sqrt{\theta u \log(n)}}, \; 0 \le u \le 1,$$

whose behavior is severely different from $X_n(u)$ when u is nearly 0. About standardizations, see also the following remark.

Remark 9.1.2. *For the asymptotic normality, Yamato (2013) showed that the approximation accuracy of*

$$\frac{\sum_{j=1}^{n} C_j^n - \theta(\log(n) - \psi(\theta))}{\sqrt{\theta(\log(n) - \psi(\theta))}} \to^d N(0,1)$$

is better than (9.1.1) and derived the Edgeworth expansions, where ψ is the digamma function $\psi(x) = d \log \Gamma(x)/dx = \Gamma'(x)/\Gamma(x)$. It has a representation

$$\psi(x) = -\gamma - \frac{1}{x} + \sum_{j=1}^{\infty} \frac{x}{j(x+j)},$$

where γ is the Euler's γ which is defined by

$$\gamma = \lim_{n\to\infty} \left| \sum_{j=1}^{n} \frac{1}{j} - \log(n) \right|.$$

This result is similar to ours in the sense that it is not standardized by $\theta \log(n)$. Moreover, Arratia et al. (2000) proved the functional CLT of

$$u \rightsquigarrow \sum_{j=1}^{[n^u]} \frac{\left(C_j^n - \mathbb{E}[Z_j]\right)}{\sqrt{\theta \log(n)}},$$

which is the equation (3.14) in their paper.

9.2 A Poisson process approximation

Arratia and Tavaré (1992) showed that

$$u \rightsquigarrow \frac{\sum_{j=1}^{[n^u]} (C_j^n - Z_j)}{\sqrt{\theta \log(n)}}$$

converges to 0 in probability uniformly with respect to u, which means $\sum_{j=1}^{[n^u]} C_j^n$ is approximated by the sum of independent Poisson random variables. Corresponding to this, we prove the L^2 norm of

$$u \rightsquigarrow \frac{\sum_{j=1}^{[n^u]}(C_j^n - Z_j)}{\sqrt{\ell([n^u])}} \tag{9.2.1}$$

converges to 0 in probability by proving the following lemma.

Lemma 9.2.1. *It holds that*

$$\mathbb{E}\left[\int_0^1 \left|\frac{\sum_{j=1}^{[n^u]}(C_j^n - Z_j)}{\sqrt{\ell([n^u])}}\right|^2 du\right] = O\left(\frac{\log\log(n)}{\log(n)}\right). \tag{9.2.2}$$

Remark 9.2.1. *This order is different from one without u in the denominator:*

$$\mathbb{E}\left[\sup_{0\le u\le 1}\left|\frac{\sum_{j=1}^{[n^u]}(C_j^n - Z_j)}{\sqrt{\theta\log(n)}}\right|\right] = O\left(\frac{1}{\sqrt{\log(n)}}\right),$$

see Arratia and Tavaré (1992). In the equation (9.2.1), u in the denominator changes the numerator of the right-hand side in (9.2.2). By the same argument we can prove

$$\mathbb{E}\left[\sup_{0\le u\le 1}\left|\frac{\sum_{j=1}^{[n^u]}(C_j^n - Z_j)}{\sqrt{\theta\log(n)}}\right|^2\right] = O\left(\frac{1}{\log(n)}\right).$$

Proof of the Lemma 9.2.1. Since it holds that

$$\begin{aligned} C_j^n - Z_j &=^d C_j^n - C_j^\infty \\ &= -\sum_{l\ge n-j+1} \xi_l(1-\xi_{l+1})\cdots(1-\xi_{l+j-1})\xi_{l+j} \\ &\quad +\xi_{n-j+1}(1-\xi_{n-j+2})\cdots(1-\xi_n), \end{aligned}$$

the inequality $(a+b)^2 \le 2(a^2+b^2)$ yields that

$$\begin{aligned} &\frac{1}{2}\mathbb{E}[|C_j^n - Z_j|^2] \\ \le\ & \mathbb{E}\left[\left(\sum_{l\ge n-j+1} \xi_l(1-\xi_{l+1})\cdots(1-\xi_{l+j-1})\xi_{l+j}\right)^2\right] \\ &+\mathbb{E}\left[\xi_{n-j+1}^2(1-\xi_{n-j+2})^2\cdots(1-\xi_{n+1})^2\right]. \end{aligned} \tag{9.2.3}$$

By the monotone convergence theorem, the first term is equal to

$$
\begin{aligned}
& \mathbb{E}\left[\left(\lim_{N\to\infty}\sum_{l=n-j+1}^{N}\xi_l(1-\xi_{l+1})\cdots(1-\xi_{l+j-1})\xi_{l+j}\right)^2\right] \\
= \ & \mathbb{E}\left[\lim_{N\to\infty}\left(\sum_{l=n-j+1}^{N}\xi_l(1-\xi_{l+1})\cdots(1-\xi_{l+j-1})\xi_{l+j}\right)^2\right] \\
= \ & \lim_{N\to\infty}\mathbb{E}\left[\left(\sum_{l=n-j+1}^{N}\xi_l(1-\xi_{l+1})\cdots(1-\xi_{l+j-1})\xi_{l+j}\right)^2\right] \\
\leq \ & \lim_{N\to\infty}\left(\mathbb{E}\left[\sum_{l=n-j+1}^{N}\xi_l^2\xi_{l+j}^2+\sum_{l=n-j+1}^{N}\xi_l\xi_{l+j}^2\xi_{l+2j}\right.\right. \\
& \left.\left.+\sum_{l=n-j+1}^{N}\xi_l\xi_{l+j}\sum_{k=n-j+1;k\neq l}^{N}\xi_k\xi_{k+j}\right]\right) \\
= \ & \lim_{N\to\infty}\left(\sum_{l=n-j+1}^{N}p_lp_{l+j}+\sum_{l=n-j+1}^{N}p_lp_{l+j}p_{l+2j}\right. \\
& \left.+\sum_{l=n-j+1}^{N}p_lp_{l+j}\sum_{k=n-j+1;k\neq l}^{N}p_kp_{k+j}\right) \\
\leq \ & \sum_{l\geq n-j+1}(p_lp_{l+j}+p_lp_{l+j}p_{l+2j})+\left(\sum_{l=n-j+1}^{\infty}p_lp_{l+j}\right)^2
\end{aligned}
$$

and the second term of (9.2.3) is bounded above by $p_{n-j+1}(1-p_{n+1})$, since the sequence $\{\xi_\cdot\}$ is a sequence of mutually independent random variables, and $(1-\xi_\cdot)$ are bounded above by 1. It holds that

$$\begin{aligned}
&\sum_{l\geq n-j+1}(p_lp_{l+j}+p_lp_{l+j}p_{l+2j})+\left(\sum_{l=n-j+1}^{\infty}p_lp_{l+j}\right)^2+p_{n-j+1}(1-p_{n+1})\\
=\ &\theta^2\sum_{l>n-j+1}\frac{1}{(\theta+l-1)(\theta+l+j-1)}\\
&+\theta^3\sum_{l>n-j+1}\frac{1}{(\theta+l-1)(\theta+l+j-1)(\theta+l+2j-1)}\\
&+\left(\theta^2\sum_{l>n-j+1}\frac{1}{(\theta+l-1)(\theta+l+j-1)}\right)^2+\frac{\theta n}{(\theta+n-j)(n+\theta)}\\
=\ &\frac{\theta^2}{j}\sum_{l=n-j+2}^{n+1}\frac{1}{\theta+l-1}+\frac{\theta^3}{2j}\sum_{l=n-j+2}^{n+1}\left(\frac{1}{(\theta+l-1)(\theta+l+j-1)}\right)\\
&+\left(\frac{\theta^2}{j}\sum_{l=n-j+2}^{n+1}\frac{1}{\theta+l-1}\right)^2+\frac{\theta n}{(\theta+n-j)(n+\theta)}.
\end{aligned}$$

It is bounded above by

$$\begin{aligned}
&\frac{\theta^2}{\theta+n-j+1}+\frac{\theta^3}{2(\theta+n-j+1)(\theta+n+1)}\\
&+\left(\frac{\theta^2}{\theta+n-j+1}\right)^2+\frac{\theta n}{(\theta+n-j)(n+\theta)}\\
\leq\ &\frac{\theta^2(1+\theta)}{\theta+n-j+1}\left(1+\frac{\theta}{2(\theta+n+1)}\right)+\frac{\theta n}{(\theta+n-j)(n+\theta)}.
\end{aligned}$$

Next, It holds that

$$\begin{aligned}
&\sum_{j=1}^{n}\mathbb{E}[|C_j^n-Z_j|^2]\\
\leq\ &\sum_{j=1}^{b}\mathbb{E}[|C_j^n-Z_j|^2]+\sum_{j=b+1}^{n}\mathbb{E}[(C_j^n)^2]+\sum_{j=b+1}^{n}\mathbb{E}[(Z_j)^2] \qquad (9.2.4)
\end{aligned}$$

for any $1 \le b \le n$. The first term in (9.2.4) is bounded above by

$$\begin{aligned}
&\sum_{j=1}^{b} \frac{\theta}{\theta+n-j}\left(\theta(1+\theta)+\frac{n}{n+\theta}+\frac{\theta^2}{2(\theta+n+1)}\right) \\
< \ &\frac{\theta b}{\theta+n-b}\left(\theta(1+\theta)+\frac{n}{n+\theta}+\frac{\theta^2}{2(\theta+n+1)}\right).
\end{aligned}$$

Moreover, the Lemma 1 of Arratia et al. (1992) yields that

$$C_j^n \le C_j^\infty + 1\{J_n = j\},$$

where the random variable $J_n \in \{1, \ldots, n\}$ is defined by

$$J_n = \min(j \ge 1 : \xi_{n+1-j} = 1),$$

so, because of the property of Poisson distribution, the second and third terms of the right-hand side of (9.2.4) are bounded above by

$$1 + 2\sum_{j=b+1}^{n} \mathbb{E}[Z_j] + 2\sum_{j=b+1}^{n} \mathbb{E}[(Z_j)^2] = 1 + 4\sum_{j=b+1}^{n} \mathbb{E}[Z_j] + 2\sum_{j=b+1}^{n} \left(\mathbb{E}[Z_j]\right)^2,$$

because it holds that

$$\sum_{j=b+1}^{n} \mathbb{E}[1\{J_n = j\}] \le \sum_{j=1}^{n} \mathbb{E}[1\{J_n = j\}] = 1.$$

Since $\mathbb{E}[Z_j] = \theta/j$, the second and third terms in the right-hand side are bounded above by $4\theta\log(n/b)$ and $4\theta^2$, respectively, because it holds that

$$\sum_{j=1}^{n} \frac{1}{j^2} = \frac{\pi^2}{6} < 2.$$

Therefore, for $n \ge 2$, the right-hand side of (9.2.4) is bounded above by

$$\frac{n\theta}{2\theta+n-2}\left(\theta(1+\theta)+\frac{n}{n+\theta}+\frac{\theta^2}{2(\theta+n+1)}\right) + 1 + 4\theta^2 + 4\theta\log\left(\frac{2n}{n-1}\right),$$

where we set b to $[n/2]$. As $n \to \infty$, it converges to

$$1 + (1 + 4\log 2)\theta + 5\theta^2 + \theta^3,$$

so is $O(1)$. Hence, it holds that

$$\begin{aligned}
&\mathbb{E}\left[\int_0^1 \left|\frac{\sum_{j=1}^{[n^u]}(C_j^n - Z_j)}{\sqrt{\ell([n^u])}}\right|^2 du\right] \qquad (9.2.5)\\
= \ & \mathbb{E}\left[\int_0^1 \left|\frac{\sum_{j=1}^{n} 1\{j \le n^u\}(C_j^n - Z_j)}{\sqrt{\ell([n^u])}}\right|^2 du\right]\\
\le \ & \mathbb{E}\left[\int_0^1 \frac{\sum_{j=1}^{n} 1\{j \le n^u\} \left|C_j^n - Z_j\right|^2}{\ell([n^u])} du\right]\\
= \ & \mathbb{E}\left[\int_0^1 \frac{|C_1^n - Z_1|^2 + \sum_{j=2}^{n} 1\{j \le n^u\} \left|C_j^n - Z_j\right|^2}{\ell([n^u])} du\right]\\
= \ & \mathbb{E}\left[\int_0^1 \frac{du}{\ell([n^u])}|C_1^n - Z_1|^2 + \sum_{j=2}^{n}\int_0^1 \frac{1\{j \le n^u\}}{\ell([n^u])} du \left|C_j^n - Z_j\right|^2\right].
\end{aligned}$$

For the right-hand side, the inequality

$$\ell([n^u]) = \theta \sum_{j=1}^{[n^u]} \frac{1}{j} > \theta \log([n^u]+1) > \theta \log(n^u),$$

yields that

$$\begin{aligned}
\int_0^1 \frac{du}{\ell([n^u])} &< \frac{1}{\theta}\int_0^1 \frac{du}{\log(1+[n^u])}\\
&< \frac{1}{\theta}\int_0^{\frac{1}{\log(n)}} \frac{du}{\log(1+[n^u])} + \frac{1}{\theta}\int_{\frac{1}{\log(n)}}^1 \frac{du}{\log(n^u)}\\
&< \frac{1}{\theta\log(n)}\left(\frac{1}{\log 2} + \log\log(n)\right)
\end{aligned}$$

and that for $j \ge 2$

$$\begin{aligned}
\int_0^1 \frac{1\{j \le n^u\}}{\ell([n^u])} du &= \int_{\frac{\log(j)}{\log(n)}}^1 \frac{du}{\ell([n^u])}\\
&< \frac{1}{\theta\log(n)}\int_{\frac{\log(j)}{\log(n)}}^1 \frac{du}{u}\\
&= \frac{\log\log(n) - \log\log(j)}{\theta\log(n)}.
\end{aligned}$$

Therefore, the right-hand side of (9.2.5) is bounded above by

$$\begin{aligned}
& \frac{1}{\theta\log(n)}\mathbb{E}\left[\left(\frac{1}{\log 2}+\log\log(n)\right)|C_1^n - Z_1|^2 \right. && (9.2.6)\\
& \left. +\sum_{j=2}^{n}(\log\log(n)-\log\log(j))\left|C_j^n - Z_j\right|^2\right] \\
\leq\ & \frac{1}{\theta\log(n)}\mathbb{E}\left[\frac{|C_1^n - Z_1|^2}{\log 2} - \log\log 2\,|C_2^n - Z_2|^2 \right.\\
& \left. +\log\log(n)\sum_{j=1}^{n}\left|C_j^n - Z_j\right|^2\right] \\
\leq\ & \frac{1}{\theta\log(n)}\left(\frac{1}{\log 2}-\log\log 2+\log\log(n)\right)\sum_{j=1}^{n}\mathbb{E}\left[\left|C_j^n - Z_j\right|^2\right] \\
=\ & O\left(\frac{\log\log(n)}{\log(n)}\right)
\end{aligned}$$

This completes the proof. □

9.3 A functional CLT for a Poisson process

Here we prove a functional CLT for a homogeneous Poisson process in L^2 space. That is because, to prove the theorem, it is of the same importance as a functional CLT for the Poisson process in $D[0,1]$ which plays a key role in Arratia and Tavaré (1992).

Lemma 9.3.1. *For the homogeneous Poisson process $\{N_\cdot\}$ whose intensity is λ, it holds that*

$$u \rightsquigarrow \frac{N_{s([n^u])}-\lambda s([n^u])}{\sqrt{\lambda s([n^u])}} \to^d G \text{ in } L^2([0,1],du),$$

where $s(n) = K\sum_{j=1}^{n} 1/j$ for any positive constant K.

Proof of the Lemma 9.3.1. First of all, it holds that

$$\frac{(N_{s([n^u])}-\lambda s([n^u]))}{\sqrt{\lambda s([n^u])}} = \frac{1}{\sqrt{\lambda}}\int_0^\infty \frac{1\{t\leq s([n^u])\}}{\sqrt{s([n^u])}}(dN_t - \lambda dt).$$

Let us denote the right-hand side by $u \rightsquigarrow \mathbb{M}_n(u)$. For the proof, we use the tightness criterion by Prohorov (1956) called asymptotic finite dimensionality, see section 1.8 of van der Vaart and Wellner (1996).

(A) Convergence of the inner product. Fix an arbitrary $h \in L^2([0,1], du)$. The Fubini theorem yields that

$$\langle \mathbb{M}_n, h\rangle_{L^2} = \frac{1}{\sqrt{\lambda}} \int_0^\infty \left(\int_0^1 \frac{1\{t \le s([n^u])\}}{\sqrt{s([n^u])}} h(u)du \right) (dN_t - \lambda dt).$$

This is a terminal value of a square integrable martingale with an adequate filtration and its predictable quadratic variation process can be written as follows:

$$\begin{aligned}
&\langle \langle \mathbb{M}_n, h\rangle_{L^2} \rangle \\
=\ & \frac{1}{\lambda} \int_0^\infty \left(\int_0^1 \frac{1\{t \le s([n^u])\}}{\sqrt{s([n^u])}} h(u)du \right)^2 \lambda dt \\
=\ & \int_0^\infty \left(\int_0^1 \int_0^1 \frac{1\{t \le s([n^u])\}1\{t \le s([n^v])\}}{\sqrt{s([n^u])s([n^v])}} h(u)h(v)dudv \right) dt \\
=\ & \int_0^1 \int_0^1 \left(\frac{s([n^u]) \wedge s([n^v])}{\sqrt{s([n^u])s([n^v])}} \right) h(u)h(v)dudv
\end{aligned}$$

For the integrand in the right-most side, it holds that

$$\left| \frac{s([n^u]) \wedge s([n^v])}{\sqrt{s([n^u])s([n^v])}} h(u)h(v) \right| \le |h(u)h(v)|,$$

$$\int_0^1 \int_0^1 |h(u)h(v)|dudv \le \int_0^1 h(u)^2 du < \infty$$

and

$$\lim_{n\to\infty} \frac{s([n^u]) \wedge s([n^v])}{\sqrt{s([n^u])s([n^v])}} = \lim_{n\to\infty} \frac{\log(n^u) \wedge \log(n^v)}{\sqrt{\log n^u \log n^v}} = \frac{u \wedge v}{\sqrt{uv}}.$$

Therefore, the dominated convergence theorem yields that

$$\lim_{n\to\infty} \langle \langle \mathbb{M}_n, h\rangle_{L^2} \rangle = \int_0^1 \int_0^1 \frac{u \wedge v h(u)h(v)}{\sqrt{uv}} dudv.$$

Let us check the Lyapunov condition. The Schwartz inequality yields that

$$\begin{aligned}
&\int_0^\infty \frac{1}{\lambda^{3/2}} \left(\int_0^1 \frac{1\{t \le s([n^u])\}}{\sqrt{s([n^u])}} h(u)du \right)^3 \lambda dt \\
\le \ & \int_0^\infty \frac{1}{\sqrt{\lambda}} \left(\int_0^1 \frac{1\{t \le s([n^u])\}}{s([n^u])} du \right)^{3/2} \left(\int_0^1 h(u)^2 du \right)^{3/2} dt.
\end{aligned}$$

The Jensen inequality gives the upper bound

$$\begin{aligned}
&\int_0^\infty \frac{1}{\sqrt{\lambda}} \int_0^1 \frac{1\{t \le s([n^u])\}}{(s([n^u]))^{3/2}} du \left(\int_0^1 h(u)^2 du \right)^{3/2} dt \\
= \ & \frac{1}{\sqrt{\lambda}} \int_0^1 \int_0^\infty 1\{t \le s([n^u])\} dt \frac{1}{(s([n^u]))^{3/2}} du \left(\int_0^1 h(u)^2 du \right)^{3/2} \\
= \ & \frac{1}{\sqrt{\lambda}} \int_0^1 \frac{1}{\sqrt{s([n^u])}} du \left(\int_0^1 h(u)^2 du \right)^{3/2}.
\end{aligned}$$

Because it holds that $s([n^u]) = K \sum_{j=1}^{[n^u]} 1/j > K \log([n^u] + 1) > K \log(n^u)$, it is bounded above by

$$\begin{aligned}
&\frac{1}{\sqrt{\lambda K \log(n)}} \int_0^1 \frac{1}{\sqrt{u}} du \left(\int_0^1 h(u)^2 du \right)^{3/2} \\
= \ & \frac{2}{\sqrt{\lambda K \log(n)}} \left(\int_0^1 h(u)^2 du \right)^{3/2} \to 0.
\end{aligned}$$

Therefore, the convergence of the inner product is proved by the martingale CLT.

(B) The asymptotic finite dimensionality. It is sufficient to prove

$$\lim_{J \to \infty} \limsup_{n \to \infty} \mathbb{E} \left[\sum_{j > J} \langle \mathbb{M}_n, e_j \rangle_{L^2}^2 \right] = 0,$$

where $\{e_\cdot\}$ is a complete orthonormal system of $L^2([0,1], du)$. The Fubini theorem yields that

$$\langle \mathbb{M}_n, e_j \rangle_{L^2} = \frac{1}{\sqrt{\lambda}} \int_0^\infty \left(\int_0^1 \frac{1\{t \le s([n^u])\}}{\sqrt{s([n^u])}} e_j(u) du \right) (dN_t - \lambda dt).$$

It holds that

$$\limsup_{n\to\infty} \mathbb{E}\left[\sum_{j>J} \langle \mathbb{M}_n, e_j\rangle_{L^2}^2\right] = \limsup_{n\to\infty} \mathbb{E}\left[\|\mathbb{M}_n\|_{L^2}^2 - \sum_{j=1}^{J} \langle \mathbb{M}_n, e_j\rangle_{L^2}^2\right]. \quad (9.3.1)$$

For the first term of the integrand in the right-hand side, it follows from $\mathbb{E}[\|\mathbb{M}_n\|_{L^2}^2] = 1$ that

$$\limsup_{n\to\infty} \mathbb{E}[\|\mathbb{M}_n\|_{L^2}^2] = 1,$$

and, for the second term, that

$$\begin{aligned}
& \limsup_{n\to\infty} \mathbb{E}\left[-\sum_{j=1}^{J} \langle \mathbb{M}_n, e_j\rangle_{L^2}^2\right] \\
= & -\liminf_{n\to\infty} \mathbb{E}\left[\sum_{j=1}^{J} \langle \mathbb{M}_n, e_j\rangle_{L^2}^2\right] \\
= & -\liminf_{n\to\infty} \sum_{j=1}^{J} \mathbb{E}\left[\left(\frac{1}{\sqrt{\lambda}} \int_0^\infty \left(\int_0^1 \frac{1\{t \le s([n^u])\}}{\sqrt{s([n^u])}} e_j(u) du\right) (dN_t - \lambda dt)\right)^2\right] \\
= & -\liminf_{n\to\infty} \sum_{j=1}^{J} \frac{1}{\lambda} \int_0^\infty \left(\int_0^1 \frac{1\{t \le s([n^u])\}}{\sqrt{s([n^u])}} e_j(u) du\right)^2 \lambda dt \\
= & -\liminf_{n\to\infty} \sum_{j=1}^{J} \int_0^1 \int_0^1 \int_0^\infty \frac{1\{t \le s([n^u]) \wedge s([n^v])\}}{\sqrt{s([n^u]) s([n^v])}} dt e_j(u) e_j(v) du dv \\
= & -\liminf_{n\to\infty} \int_0^1 \int_0^1 \frac{s([n^u]) \wedge s([n^v])}{\sqrt{s([n^u]) s([n^v])}} \left(\sum_{j=1}^{J} e_j(u) e_j(v)\right) du dv
\end{aligned}$$

Applying the Fatou-Lebesgue theorem, the right-hand side is bounded above by

$$-\int_0^1 \int_0^1 \liminf_{n\to\infty} \frac{s([n^u]) \wedge s([n^v])}{\sqrt{s([n^u]) s([n^v])}} \left(\sum_{j=1}^{J} e_j(u) e_j(v)\right) du dv,$$

since

$$\left|\frac{s([n^u]) \wedge s([n^v])}{\sqrt{s([n^u]) s([n^v])}} \left(\sum_{j=1}^{J} e_j(u) e_j(v)\right)\right| \le \sum_{j=1}^{J} |e_j(u) e_j(v)|$$

and $\sum_{j=1}^J |e_j(u)e_j(v)|$ is integrable with respect to $dudv$. It is equal to

$$-\int_0^1\int_0^1 \frac{u\wedge v}{\sqrt{uv}}\left(\sum_{j=1}^J e_j(u)e_j(v)\right)dudv = -\mathbb{E}\left[\sum_{j=1}^J \left\langle \frac{B}{\sqrt{\cdot}}, e_j\right\rangle_{L^2}^2\right]$$

The Bessel inequality yields that

$$\sum_{j=1}^J \left\langle \frac{B}{\sqrt{\cdot}}, e_j\right\rangle_{L^2}^2 \leq \left\|\frac{B}{\sqrt{\cdot}}\right\|_{L^2}^2,$$

so the dominated convergence theorem yields that

$$\lim_{J\to\infty}\mathbb{E}\left[\sum_{j=1}^J \left\langle \frac{B}{\sqrt{\cdot}}, e_j\right\rangle_{L^2}^2\right] = \mathbb{E}\left[\left\|\frac{B}{\sqrt{\cdot}}\right\|_{L^2}^2\right] = 1.$$

Hence, (9.3.1) converges to 0 as $J\to\infty$.

Because (A) and (B) hold, the conclusion follows from the Theorem 1.8.4 in van der Vaart and Wellner (1996). □

9.4 Proof of the Theorem 9.1.1.

Let $\{N_\cdot\}$ be the homogeneous Poisson process whose intensity is unity, then it holds that

$$\sum_{j=1}^{[n^\cdot]} Z_j =^d \sum_{j=1}^{[n^\cdot]}(N_{\ell(j)} - N_{\ell(j-1)}) = N_{\ell([n^\cdot])}.$$

Lemma 9.2.1 and Lemma 9.3.1 yield that

$$d\left(\frac{\sum_{j=1}^{[n^\cdot]} C_j^n - \ell([n^\cdot])}{\sqrt{\ell([n^\cdot])}}, \frac{\sum_{j=1}^{[n^\cdot]} Z_j - \ell([n^\cdot])}{\sqrt{\ell([n^\cdot])}}\right) \to^p 0,$$

where $d(,)$ is the L^2 distance, and that

$$u \rightsquigarrow \frac{\sum_{j=1}^{[n^u]} Z_j - \ell([n^u])}{\sqrt{\ell([n^u])}} \to^d G \quad \text{in } L^2([0,1], du).$$

Theorem 2.7 (iv) in van der Vaart (1998) yields the conclusion. □

Chapter 10

Random mappings

10.1 The result

As the other goal of this part, let us discuss functional CLT for random mappings in L^2 space. The strategy for the proof is basically the same as one of the Ewens sampling formula, namely, approximating the partial sum of C_j^n by the one of independent Poisson random variables and using the functional CLT for the Poisson process. The difference is that, in this chapter, we do not calculate the second moment of $\sum(C_j^n - Z_j)$ but evaluate the approximation accuracy by the total variation distance between the law of $(C_1^n, \ldots, C_b^n)$ and the law of $(Z_1, \ldots, Z_b)$. First of all, let us introduce the problem of random mappings.

Consider the n^n mappings of the set $\{1, \ldots, n\}$ to itself and assume the probabilities of such events are equally likely. Each mapping partitions the set $\{1, \ldots, n\}$. Let us argue the number C_j^n of elements whose size is equal to j. Its law is given by

$$\mathbb{P}[(C_1^n, \ldots, C_n^n) = ((c_1, \ldots, c_n))] = \frac{n! e^n}{n^n} \prod_{j=1}^{n} \frac{\lambda_j^{c_j}}{c_j!} 1\left\{\sum_{j=1}^{n} j c_j = n\right\},$$

for example see Donnelly et al. (1991a), where

$$\lambda_j = \frac{e^{-j}}{j} \sum_{i=0}^{j-1} \frac{j^i}{i!}.$$

It holds that

$$(C_1^n, C_2^n, \ldots) \to^d (Z_1, Z_2, \ldots),$$

where $\{Z_\cdot\}$ is a sequence of mutually independent Poisson variables such that

$$\mathbb{E}[Z_j] = \lambda_j,\ j = 1, \dots, n.$$

Also in this case, the asymptotic normality

$$\frac{\sum_{j=1}^{n} C_j^n - \frac{1}{2}\log(n)}{\sqrt{\frac{1}{2}\log(n)}} \to^d N(0,1)$$

holds (Stepanov (1969)) and a functional CLT in $D[0,1]$ holds (Hansen (1989)).

The goal in this chapter is the following theorem.

Theorem 10.1.1. *Define a random field*

$$u \rightsquigarrow Y_n(u) = \frac{\sum_{j=1}^{[n^u]} C_j^n - \ell'([n^u])}{\sqrt{\ell'([n^u])}},\ 0 \le u \le 1,$$

where

$$\ell'(n) = \sum_{j=1}^{n} \lambda_j = \sum_{j=1}^{n} \frac{e^{-j}}{j} \sum_{i=0}^{j-1} \frac{j^i}{i!}$$

for a natural number n. It holds that $Y_n \to^d G$ in $L^2([0,1], du)$, and the limit $u \rightsquigarrow G(u)$ is

$$u \rightsquigarrow G(u) = \frac{B(u)}{\sqrt{u}}.$$

10.2 A Poisson process approximation

Lemma 10.2.1. *It holds that*

$$\int_0^1 \left| \frac{\sum_{j=1}^{[n^u]}(C_j^n - Z_j)}{\sqrt{\ell([n^u])}} \right|^2 du \to^p 0,$$

where $\ell(n) = \sum_{j=1}^{n} 1/(2j)$.

Proof of the Lemma 10.2.1. The left-hand side is evaluated by

$$\begin{aligned}
&\int_0^1 \left| \frac{\sum_{j=1}^n 1\{j \le n^u\}(C_j^n - Z_j)}{\sqrt{\ell([n^u])}} \right|^2 du \\
= &\int_0^1 \sum_{j=1}^n \sum_{k=1}^n \frac{1\{j \le n^u\}(C_j^n - Z_j) 1\{k \le n^u\}(C_k^n - Z_k)}{\ell([n^u])} du \\
= &\sum_{j=1}^n \sum_{k=1}^n \int_0^1 \frac{1\{j \vee k \le n^u\}(C_j^n - Z_j)(C_k^n - Z_k)}{\ell([n^u])} du \\
\le &\sum_{j=1}^n \sum_{k=1}^n \int_0^1 \frac{1\{j \vee k \le n^u\}}{\ell([n^u])} du |(C_j^n - Z_j)(C_k^n - Z_k)|
\end{aligned}$$

By the similar way as (9.2.6), the right-hand side is bounded above by

$$\begin{aligned}
&\frac{2}{\log(n)} \left(\frac{1}{\log 2} + \log\log(n) \right) |(C_1^n - Z_1)(C_1^n - Z_1)| \\
- &\frac{2 \log\log 2}{\log(n)} |(C_2^n - Z_2)((C_1^n - Z_1) + (C_2^n + Z_2))| \\
+ &\sum_{j=1}^n \sum_{k=2}^n \frac{2 \log\log(n)}{\log(n)} |(C_j^n - Z_j)(C_k^n - Z_k)| \\
< &\frac{2}{\log(n)} \left(\frac{1}{\log 2} - \log\log 2 + \log\log(n) \right) \left(\sum_{j=1}^n |C_j^n - Z_j| \right)^2
\end{aligned}$$

So, it is sufficient to prove that

$$\sqrt{\frac{\log\log(n)}{\log(n)}} \sum_{j=1}^n |C_j^n - Z_j| \to^p 0.$$

For any $1 \le b = b(n) \le n$, the triangle inequality yields that

$$\sum_{j=1}^n |C_j^n - Z_j| \le \sum_{j=1}^b |C_j^n - Z_j| + \sum_{j=b+1}^n C_j^n + \sum_{j=b+1}^n Z_j.$$

By the way similar to the equation (22) in Arratia et al. (1995), fix a good coupling such that the first term in the right-hand side is bounded by the total variation distance which converges to 0 and both of the expectation of the second term and third term are $O(\log\log(n))$ where we let $b(n) = n/\log(n)$. This completes the proof. □

This lemma yields the following corollary, which is a Poisson process approximation.

Corollary 10.2.1. *If Lemma 10.2.1 holds, then it holds that*

$$\int_0^1 \left| \frac{\sum_{j=1}^{[n^u]}(C_j^n - Z_j)}{\sqrt{\ell'([n^u])}} \right|^2 du \to^p 0.$$

Proof of the Corollary 10.2.1. Consider random variables $P_j \sim Pois(\lambda_j = j) \quad j = 1, 2, \ldots$. By the definition of the median, it holds that

$$\mathbb{P}(P_j < med(P_j)) < \frac{1}{2}.$$

On the other hand, it holds that

$$\mathbb{P}(P_j < j) = e^{-j} \sum_{i=0}^{j-1} \frac{j^i}{i!} = j\lambda_j.$$

Teicher (1955) proves that $j\lambda_j$ is increasing as j goes larger, which is stated in their second inequality of (8). The convergence (20) in Donnelly et al. (1991a)

$$j\lambda_j \to \frac{1}{2}$$

yields that

$$\frac{1}{2j} > \lambda_j, \quad \forall j = 1, 2, \ldots. \tag{10.2.1}$$

Since it holds that

$$\sum_{j=1}^{\infty} \left(\frac{1}{2j} - \lambda_j \right) = \frac{1}{2} \log(2),$$

which is the equation (31) in Donnelly et al. (1991a) and (10.2.1), it holds that

$$0 < \sup_{u \in [0,1]} \left(\ell([n^u]) - \ell'([n^u]) \right) = \sum_{j=1}^{n} \left(\frac{1}{2j} - \lambda_j \right) = O(1)$$

and $\inf_{u\in[0,1]} \ell'([n^u]) = \ell'(1) = 1/e$. Thence, we have

$$
\begin{aligned}
& \int_0^1 \left| \frac{\sum_{j=1}^{[n^u]}(C_j^n - Z_j)}{\sqrt{\ell'([n^u])}} \right|^2 du \\
= & \int_0^1 \left| \frac{\sum_{j=1}^{[n^u]}(C_j^n - Z_j)}{\sqrt{\ell([n^u])}} \right|^2 \left(1 + \frac{\ell([n^u]) - \ell'([n^u])}{\ell'([n^u])}\right) du \\
\leq & \int_0^1 \left| \frac{\sum_{j=1}^{[n^u]}(C_j^n - Z_j)}{\sqrt{\ell([n^u])}} \right|^2 \left(1 + \frac{\sup_{u\in[0,1]}\left(\ell([n^u]) - \ell'([n^u])\right)}{\inf_{u\in[0,1]} \ell'([n^u])}\right) du \\
= & \left(1 + \frac{\ell(n) - \ell'(n)}{\ell'(1)}\right) \int_0^1 \left| \frac{\sum_{j=1}^{[n^u]}(C_j^n - Z_j)}{\sqrt{\ell([n^u])}} \right|^2 du \\
= & \left(1 + e\sum_{j=1}^{n}\left(\frac{1}{2j} - \lambda_j\right)\right) \int_0^1 \left| \frac{\sum_{j=1}^{[n^u]}(C_j^n - Z_j)}{\sqrt{\ell([n^u])}} \right|^2 du \to^p 0.
\end{aligned}
$$

This completes the proof. □

This corollary and the functional CLT for a Poisson process yields the Theorem 10.1.1.

Bibliography

ANDERSON, T.W. and DARLING, D.A. (1952). Asymptotic theory of certain "goodness of fit criteria" based on stochastic processes, *Ann. Math. Statist.* **23** 193–212.

ARRATIA, R., BARBOUR, A.D. and TAVARÉ, S. (1992). Poisson process approximations for the Ewens sampling formula, *Ann. Appl. Probab.* **2** 519–535.

ARRATIA, R., BARBOUR, A.D. and TAVARÉ, S. (2000). Limits of logarithmic combinatorial structures, *Ann. Probab.* **28** 1620–1644.

ARRATIA, R., BARBOUR, A.D. and TAVARÉ, S.(2003). *Logarithmic Combinatorial Structures: a Probabilistic Approach.* EMS Monographs in Mathematics. European Mathematical Society (EMS), Zürich.

ARRATIA, R., STARK, D. and TAVARÉ, S.(1995). Total variation asymptotics for Poisson process approximations of logarithmic combinatorial assemblies, *Ann. Probab.* **23** 1347–1388.

ARRATIA, R. and TAVARÉ, S. (1992). Limit theorems for combinatorial structures via discrete process approximations, *Random Structures Algorithms* **3** 321–345.

BHATTACHARYA, P.K., (1987). Maximum likelihood estimation of a change-point in the distribution of independent random variables: general multiparameter case, *J. Multivariate Anal.* **23** 183–208.

BILLINGSLEY, P. (1999). *Convergence of Probability Measures. Second edition.* John Wiley & Sons, Inc., New York.

BILLINGSLEY, P. (2012). *Probability and Measure. Anniversary edition.* John Wiley & Sons, Inc., Hoboken, NJ.

BRODSKY, B.E. and DARKHOVSKY, B.S. (2000). *Non-parametric Statistical Diagnosis: Problems and Methods*. Kluwer Academic Publishers, Dordrecht.

CSÖRGŐ, M., CSÖRGŐ, S., HORVÁTH, L. and MASON, D.M. (1986). Weighted empirical and quantile processes, *Ann. Probab.* **14** 31–85.

CHEN, J. and GUPTA, A. K. (2012). *Parametric statistical change point analysis. With applications to genetics, medicine, and finance. Second edition.* Birkhuser/Springer, New York.

CSÖRGŐ, M. and HORVÁTH, L. (1997). *Limit Theorems in Change-point Analysis*. John Wiley & Sons, Ltd., Chichester.

CSÖRGŐ, M., HORVÁTH, L. and SHAO, Q.-M. (1993). Convergence of integrals of uniform empirical and quantile processes, *Stochastic Process Appl.* **45** 283–294.

DEDECKER, J. and MERLEVÈDE, F.(2003). The conditional central limit theorem in Hilbert spaces, *Stochastic Process Appl.* **108** 229-262.

DEGREGORIO, A. and IACUS, S. M. (2008). Least squares volatility change point estimation for partially observed diffusion processes, *Comm. Statist. Theory Meth.* **37** 2342–2357.

DEHLING, H., FRANKE, B., KOTT, T. and KULPERGER, R. (2014). Change point testing for the drift parameters of a periodic mean reversion process, *Statist. Inference Stoch. Process.* **17**, 1, 1–18.

DELAURENTIS, J. M. and PITTEL, B. (1985). Random permutations and Brownian motion, *Pac. J. Math.* **119** 287–301.

DONNELLY, P., EWENS, W.J. and PADMADISASTRA, S. (1991a). Functionals of random mappings: exact and asymptotic results, *Adv. Appl. Probab.* **23** 437–455.

DONNELLY, P., KURTZ, T.G. and TAVARÉ, S. (1991b). On the functional central limit theorem for the Ewens sampling formula, *Ann. Appl. Probab.* **1** 539–545.

EWENS (1972) The sampling theory of selectively neutral alleles, *Theoret. Population Biol.* **3** 87–112.

GOMBAY, E. (2008). Change detection in autoregressive time series, *J. Multivariate Anal.* **99** 451–464.

GOMBAY, E. and HORVÁTH, L. (1994). An application of the maximum likelihood test to the change-point problem, *Stochastic Process Appl.* **50** 161–171.

GOMBAY, E. and HORVÁTH, L. (1996). Approximations for the time of change and the power function in change-point models, *J. Statist. Plann. Inference* **50** 161–171.

HANSEN, J.C. (1989). A functional central limit theorem for random mappings, *Ann. Probab.* **17** 317–332. Correction: (1991). *Ann. Probab.* **19** 1393–1396.

HANSEN, J.C. (1990). A functional central limit theorem for the Ewens sampling formula, *J. Appl. Probab.* **27** 28–43.

HORVÁTH, L. and PARZEN, E. (1994). Limit theorems for fisher-score change processes. In: Carlstein, E., Müller, H.-G., Siegmund, D. (eds.) *Change-point Problems*, IMS Lecture Notes - Monogr. Ser. **23** 157–169.

HORVÁTH, L. and RICE, G. (2014). Extensions of some classical methods in change point analysis (with discussions), *TEST* **23** 219–290.

JAKUBOWSKI, A. (1980). On limit theorems for sums of dependent Hilbert space valued random variables. In Lecture Notes in Statist. **2** 178–187, Springer-Verlag, New York.

KHMALADZE, E.V. (1979). The use of ω^2 tests for testing parametric hypothesis, *Theory Probab. Appl.* **24** 283–301.

KUTOYANTS, Y. A. (2004). *Statistical Inference for Ergodic Diffusion Processes.* Springer Series in Statistics. Springer-Verlag London, Ltd., London.

LARICCIA, V. and MASON, D.M. (1986). Cramér-von Mises statistics based on the sample quantile function and estimated parameters, *J. Multivariate Anal.* **18** 93–106.

LEE, S., NISHIYAMA, Y. and YOSHIDA, N. (2006). Test for parameter change in diffusion processes by cusum statistics based on one-step estimators, *Ann. Inst. Statist. Math.* **58** 211-222.

LIANG, K.Y., SELF, S. and LIU, X. (1990). The Cox proportional hazards model with change point: an epidemiologic application, *Biometrics* **46** 783–793.

LIPTSER, R.S. and SHIRYAEV, A.N. (2001). *Statistics of Random Processes I General Theory.* Springer-Verlag, Berlin.

MASON, D.M. (1984). Weak convergence of the weighted empirical quantile process in $L^2(0,1)$, *Ann. Probab.* **12** 243–255.

MATTHEWS, D.E., FAREWELL, V.T. and PYKE, R. (1985). Asymptotic score-statistic processes and tests for constant hazard against a change-point alternative, *Ann. Statist.* **13** 583–591.

MERLEVÈDE, F. (2003). On the central limit theorem and its weak invariance principle for strongly mixing sequences with values in a Hilbert space via martingale approximation, *J. Theoret. Probab.* **16** 625–653.

MIHALACHE, S. (2012). Strong approximations and sequential change-point analysis for diffusion processes, *Statist. Probab. Lett.* **8**2, 464–472.

MOREL, B. and SUQUET, C. (2002). Hilbertian invariance principles for the empirical process under association, *Math. Methods Statist.* **11** 203–220.

NEGRI, I. and NISHIYAMA, Y. (2012). Asymptotically distribution free test for parameter change in a diffusion process model, *Ann. Inst. Statist. Math.* **64** 911–918.

NEGRI, I. and NISHIYAMA, Y. (2014). Z-process method for change point problems, *Quaderni del Dipartimento di Ingegneria dell'informazione e metodi matematici. Serie "Matematica e Statistica"* n. 5/MS 嫑 2014. Dalmine: Universit degli studi di Bergamo. Facolt di Ingegneria. Retrieved from http://hdl.handle.net/10446/30761

NISHIYAMA, Y. (1999). A maximal inequality for continuous time martingales and M-estimation in a Gaussian white noise model, *Ann. Statist.* **27** 675–696.

NISHIYAMA, Y. (2000). *Entropy Methods for Martingales.* CWI Tract **128** Centrum voor Wiskunde en Informatica, Amsterdam.

NISHIYAMA, Y. (2009). Asymptotic theory of semiparametric Z-estimators for stochastic processes with applications to ergodic diffusions and time series, *Ann. Statist.* **37** 3555–3579.

NISHIYAMA, Y. (2011). *Martingale riron ni yoru toukeikaiseki.* (In Japanese; English title: *Statistical Analysis by the Theory of Martingales.*) Kindaikagakusha, Tokyo.

OLIVEIRA, P.E. (2012). *Asymptotics for associated random variables.* Springer, Heidelberg.

OLIVEIRA, P.E. and SUQUET, C. (1995). $L^2(0,1)$ weak convergence of the empirical process for dependent variables, In Antoniadis, A. and Oppenheim, G. (eds.) *Wavelets and Statistics*, Lecture Notes in Statistics **103** 331–344.

OLIVEIRA, P.E. and SUQUET, C. (1996). An $L^2[0,1]$ invariance principle for LPQD random variables, *Port. Math.* **53** 367–379.

OLIVEIRA, P.E. and SUQUET, C. (1998). Weak convergence in $L^p[0,1]$ of the uniform empirical process under dependence, *Statist. Probab. Lett.* **39** 363–370.

PARTHASARATHY, K.R. (1967). *Probability Measures on Metric Spaces.* Academic Press, New York.

PROHOROV, Y.V. (1956). Convergence of random processes and limit theorems in probability, *Theory Probab. Appl.* **1** 157–214.

SONG, J and LEE, S. (2009). Test for parameter change in discretely observed diffusion processes, *Statist. Inference Stoch. Process.* **12** 165–183.

STEPANOV, V.E. (1969). Limit distributions for certain characteristics of random mappings, *Theory Probab. Appl.* **14**, 612–626.

SUQUET, C. and VIANO, M.C. (1998). Change point detection in dependent sequences: invariance principles for some quadratic statistics, *Math. Methods Statist.* **7** 157–191.

TEICHER, H. (1955). An inequality on Poisson probabilities, *Ann. Math. Statist.* **26**, 147–149.

TSUKUDA, K. (2014). New functional central limit theorems for the Ewens sampling formula and random mappings, preprint.

TSUKUDA, K. (2015). A change detection procedure for an ergodic diffusion process, manuscript in preparation.

TSUKUDA, K. and NISHIYAMA, Y. (2014). On L^2 space approach to change point problems, *J. Statist. Plann. Inference* **149** 46–59.

TSUKUDA, K. and NISHIYAMA, Y. (2015). Manuscript in preparation.

VAN DER VAART, A.W. (1998). *Asymptotic Statistics.* Cambridge University Press, Cambridge.

VAN DER VAART, A.W. and WELLNER, J.A. (1996). *Weak Convergence and Empirical Processes: with Applications to Statistics.* Springer-Verlag, New York.

YAMATO, H. (2013). Edgeworth expansions for the number of distinct components associated with the Ewens sampling formula, *J. Japan Statist. Soc.* **43** 17–28.

Acknowledgements

The author would like to state his sincere thanks to those who gave comments to the contents and who supported him in various senses.

The author is a Research Fellow of Japan Society for the Promotion of Science. This work is partly supported by JSPS KAKENHI Grant Number 26-1487 (Grant-in-Aid for JSPS Fellows).