Feature Sensitive and Automated Curve Registration

Technical Report No. ASU/2013/4
Dated: 24 October 2013

Dibyendu Bhaumik, Reserve Bank of India, Mumbai
Radhendushka Srivastava, Indian Institute of Technology, Bombay
and
Debasis Sengupta, Indian Statistical Institute, Kolkata

Indian Statistical Institute
Applied Statistics Unit
Kolkata 700 108

Feature Sensitive and Automated Curve Registration

Dibyendu Bhaumik, Reserve Bank of India, Mumbai
Radhendushka Srivastava, Indian Institute of Technology, Bombay
Debasis Sengupta, Indian Statistical Institute, Kolkata

Abstract

Given two sets of functional data having a common underlying mean function but different degrees of distortion in time measurements, we provide a method of estimating the time transformation necessary to align (or 'register') them. The estimator is shown to be consistent under fairly general conditions. Simulation results show superiority of the performance of the proposed method over its competitors. The method is illustrated through the analysis of a paleo-climatic data set.

1 Introduction

Functional data are typically thought to consist of noisy observations from an underlying function. Each set of observations consists of ordered pairs of the form $\{(t_j, Y_{t_j}),\, j = 1, \cdots, n\}$ where $\{t_1,\, t_2, \cdots, t_n\}$ indicate sample points of the continuum over which the function is defined and $\{Y_{t_1},\, Y_{t_2}, \cdots, Y_{t_n}\}$ represent values of the function at these sample points. Examples of functional data include measurements of biometric attributes such as height, weight etc. at different points of time, grey-scale images of scenes generated through computer vision, three-dimensional images of the brain created through various radio-imaging techniques, and so on. A good overview of techniques for analysing functional data is given in Ramsay (1982), Ramsay and Silverman (2002, 2005) and Ramsay *et al.* (2009).

Functional data may include multiple sets of measurements of a single function. However, sometimes there is a non-uniformity among the sets of observations in terms of the scales of measurement of the continuum. In a paleo-climatic study based on ice-cores, ice samples are collected at various depths of the core. Observations on climatic variables, e.g., temperature and atmospheric concentrations of greenhouse gases like carbon dioxide and methane, are measured from chemical analysis of the ice itself or of air-bubbles trapped in it. These observations are recorded against age, which is determined by dating techniques involving radioactive isotopes. Such ordered pairs of data can be regarded as functional data sampled at irregularly spaced time-points. Note that dating techniques have certain imperfections. An account of statistical issues relating to paleoclimatic dating techniques is available in Mudelsee (2010). When observations on a particular variable obtained from different ice cores from a common geographic region are available, it seems reasonable to assume that the data sets arose from the same underlying process. Collation and joint analysis of these comparable sets of data present difficulties, as imperfections in the exercise of dating may produce a certain non-uniformity of the time scales of different datasets. This problem is sometimes resolved by identifying one data set as reference and aligning others with it by using a technique of *curve registration*.

An intuitive technique of curve registration consists of identifying certain salient features (also called landmarks or markers) in the curves and then aligning the curves by transforming time so that marker events occur at the same values of the transformed times. This technique, referred to as *marker* registration (or sometimes as *landmark* registration), is discussed in detail in Bookstein (1991), Kneip and Gasser (1988, 1992) and Gasser and Kneip (1995). A major drawback of the marker based approach is the need for manual specification of the markers by a domain expert. Bigot (2003, 2005) proposed a method of detecting markers through wavelets, for eventual use in marker registration.

Using an idea of Silverman (1995), Ramsay and Li (1998) developed a technique for curve registration that does not require identification of markers. This estimation method involves optimizing a global fitting criterion with respect to a restricted parametric family of time transformations. Further research in this direction was conducted by Ramsay (1998). Kneip *et al.* (2000) mentioned that the numerical optimization techniques used in these methods may sometimes be plagued by ill-conditioning. They proposed curve registration with locally estimated monotone time transformations

Wang and Gasser (1997, 1999) built on the idea of *dynamic time warping* found in the engineering literature [see Shakoe and Chiba (1978)] in curve registration, and extended the idea to align multiple curves.

Some authors have explored curve registration using self-modelling regression (SEMOR). The original version of SEMOR (see Lawton *et al.*, 1972 and Kniep and Gasser, 1988) is based on a semiparametric model, in which subject-specific regression functions are parametric transformations of a common regression function. Specifically, let $y_i(t)$ denote the observed curve for subject i at time t. The general SEMOR model can be expressed as $y_i(t) = m_i(t) + e_i$ with $m_i(t) = g(m, \theta_i)(t)$ where m is the common shape function, g is the transformation that generates the individual regression functions, and e_i is the random error. Both the shape function g and the model parameter θ_i are estimated from the data. A common class of SEMOR models with transformations of both x and y axes and a shape invariant shift has $m_i(t) = \alpha_i + \beta_i m(\gamma_i t + \delta_i)$, referred to as the shape-invariant model (SIM). Rønn (2001), considering a version of the foregoing SIM model in which $m_i(t) = m(t + \delta_i)$, where the shift δ_i are random effects, proposed nonparametric maximum likelihood registration of random time-shifted curves. Brumback and Lindstrom (2004) built on SIM models, replacing the linear time transformation function to allow more flexible, monotone random time transformations. Specifically, in their formulation, curve-specific amplitude parameters as well as curve-specific time transformations are assumed to be random. Moreover, the time transformation and common shape function are modeled with B-splines.

With the aim of developing flexible alignment functions, Gervini and Gasser (2004) proposed a self-modelling alignment function with the components of the function estimated from the data using B-spline basis functions. In their formulation, the time transformations are not random. Liu and Müller (2004) proposed time synchronization of random processes where the observed random curves are assumed to be generated by a bivariate stochastic process consisting of a stochastic monotone time transformation function and an unrestricted random amplitude function.

In Bugli *et al.* (2005), a registration technique based on fractional polynomials is proposed. James (2007) developed another curve registration technique based on equating the "moments" of a given set of curves. Bayesian approach to curve registration can be found in Telesca and Inoue (2008).

Central to most of the above methods is the notion that the functions being registered ought to be smooth. However, in some situations, data exhibit sharp variations. Making use of such information may yield better curve registration. The aforesaid methods are unable to take advantage of this information, as smoothing tends to blur the features. On the other hand, marker based methods require explicit identification of markers, which may not be desirable as a pre-condition for registration. An automated alternative method can potentially be developed by utilizing the idea of edge-preserving smoothing techniques [see McDonald and Owen (1986), Qiu *et al.* (1991), Hall and Titterington (1992), Müller (1992), Wu and Chu (1993) and Chu *et al.* (1998)].

In this paper, we propose a new nonparametric statistical measure for the alignment of two time transformed functional data sets. We show that this measure of alignment possesses some desired properties. We then estimate the time transformation by maximizing this measure over an appropriate class of transformations. Subsequently, we establish the consistency of the estimator. Results of a simulation study, to demonstrate the performance of a particular estimator of the proposed class, are reported. The method is illustrated through the analysis of a real data set.

2 Methodology

Let $\{(t_j, Y_{t_j}),\ j = 1, \cdots, n_1\}$ and $\{(s_j, Y'_{s_j}),\ j = 1, \cdots, n_2\}$ be two sets of functional data, arising from the model

$$\begin{aligned} Y_t &= m(t) + \epsilon_t, \\ Y'_s &= m(g_0(s)) + \epsilon'_s, \end{aligned} \tag{1}$$

where m is an underlying location function, and g_0 is a time transformation (alignment) function. The terms ϵ_t and ϵ'_s represent additive random error.

If s and t are time points such that $g_0(s) = t$, one would expect Y'_s to be close to Y_t. In a real data sets, the equivalence of t and $g_0(s)$ is unlikely to occur frequently. On the other hand, if m is a continuous function, then Y'_s should be close to Y_t even when $g_0(s)$ is close to t but not equal to it.

In a pair of real data sets following model (1), the function g_0 is not known. However, if it had been known and $g_0(s_j)$ is close to t_i for some $i \in \{1, \dots, n_1\}$ and $j \in \{1, \dots, n_2\}$, then both the differences $t_i - g_0(s_j)$ and $Y_{t_i} - Y'_{s_j}$ are expected to be small. If K_1 and K_2 are kernel functions and h_1 and h_2 are appropriate bandwidth parameters, then the product

$$\frac{1}{h_1} K_1 \left(\frac{t_i - g_0(s_j)}{h_1} \right) \frac{1}{h_2} K_2 \left(\frac{Y_{t_i} - Y'_{s_j}}{h_2} \right),$$

should be large for such a pair of data points but not so large for other pairs. Moreover, the product should be small when g_0 is replaced by another function.

In view of the above discussion, for any given transformation function g, we propose the measure of alignment between the two data sets

$$C_n(g) = \frac{\dfrac{1}{n_1 n_2} \displaystyle\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \frac{1}{h_1} K_1 \left(\frac{t_i - g(s_j)}{h_1} \right) \frac{1}{h_2} K_2 \left(\frac{Y_{t_i} - Y'_{s_j}}{h_2} \right)}{\dfrac{1}{n_1 n_2} \displaystyle\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \frac{1}{h_1} K_1 \left(\frac{t_i - g(s_j)}{h_1} \right)} \tag{2}$$

where $n = n_1 + n_2$, K_1 and K_2 are kernel functions and h_1 and h_2 are bandwidth parameters. Under model (1), this measure of alignment should be large when $g = g_0$. On the other hand, alignment through a wrong time transformation would produce fewer pairs of matching points in the two series, and consequently fewer large summands in the numerator would produce a smaller value of the measure.

We propose to estimate the function g_0 by the maximizer of the measure of alignment, C_n, over an appropriate set of candidate functions.

In several applications (e.g., paleo-climatic data sets), time transformation is expected to be a monotonic and continuous function, which is not very far from the identity function. The search space for the optimizing function may be chosen in view of these considerations.

As for the choices for bandwidths, it is clear that h_1 should be determined by the average seperation between sucessive samples in the iwo sets of data (say R_1 and R_2), and h_2 should depend on the range of the combined set of observed values of the underlying function(say R_3). A simple set of choices would be $h_1 = c_1 \min\{R_1, R_2\}$ and $h_2 = R_3/c_2$ where c_1 and c_2 are some constants. Values of c_1 and c_2 around 10 and 10, respectively, worked well in the simulations and the data analysis reported in Sections 4 and 5.

3 Theoretical Results

We begin with a clear specification of model (1). Let m be a continuous function (smooth?) and g_0 be a monotonically increasing and continuous function. Let the errors $\{\epsilon_{1j};\, j = 1, \ldots n_1\}$ and $\{\epsilon_{2j};\, j = 1, \ldots n_2\}$ and the time points $\{t_j;\, j = 1, \ldots n_1\}$ and $\{s_j;\, j = 1, \ldots n_2\}$ be mutually independent sets of samples from four distributions having probability density functions $f_{\epsilon 1}$, $f_{\epsilon 2}$, f_1 and f_2, respectively, with the last two densities defined over $[0, \infty)$.

We formally define the proposed estimator of the function g_0 as

$$\widehat{g_n} = \arg\max_{g \in \mathbb{G}} C_n(g), \tag{3}$$

where $C_n(g)$ is as defined in (2) and $\mathbb{G}$ is a class of functions that includes the true transformation function.

We make the following assumptions.

Assumption 1. The densities $f_{\epsilon 1}$, $f_{\epsilon 2}$, f_1 and f_2 are continuous and bounded. $f_{\epsilon 1}$ and $f_{\epsilon 2}$ are symmetric about zero and having modes at zero.

Assumption 2. The kernels K_1 and K_2 are positive valued, bounded, continuous, and integrable functions defined over the real line. Further $\int K_i(x)dx = 1$, $i = 1, 2$.

Assumption 3. The sample sizes n_1 and n_2 are such that $n_1/n \to \xi$ for some $\xi \in (0, 1)$, as

$n \to \infty$.

Assumption 4. The bandwidths h_i's are such that $h_i \to 0$ and $n_i h_i \to \infty$ as $n \to \infty$.

Theorem 3.1 *Under assumptions 1–4, for any $g \in \mathbb{G}$ as $n \to \infty$, $C_n(g) \xrightarrow{P} C(g)$, where*

$$C(g) = \frac{\int_{-\infty}^{\infty} \int_0^{\infty} f_1(g(y)) f_2(y) f_{\epsilon 1}(v - m(g(y)) + m(g_0(y))) f_{\epsilon 2}(v) dy dv}{\int_0^{\infty} f_1(g(y)) f_2(y) dy}. \tag{4}$$

Let $C[a, b]$ be the metric space of all continuous functions defined over the interval $[a, b]$, with respect to the *supremum* norm.

Assumption 5. The class $\mathbb{G}$ is a compact subset of $C[a, b]$, consisting of monotonically increasing functions with first derivative bounded away from zero.

Assumption 2A. The kernels K_1 and K_2 are bounded away from 0 and have bounded first order derivatives.

Theorem 3.2 *Under the assumptions 1–4, 2A, and 5, as $n \to \infty$,*

$$\sup_{g \in \mathbb{G}} |C_n(g) - C(g)| \xrightarrow{P} 0.$$

Theorem 3.3 *Under the assumptions 1–4, 2A, and 5, as $n \to \infty$, $C_n(\widehat{g}) \xrightarrow{P} C(g_0)$.*

Assumption 6. The supremum of the functional C on $\mathbb{G}$ defined as in (4) is attained only at g_0.

Theorem 3.4 *Under the assumptions 1–4, 2A, 5, and 6, as $n \to \infty$, $\widehat{g} \xrightarrow{P} g_0$.*

4 Simulation Study

In order to compare the performance of the proposed method with that proposed by Ramsay and Li (1998), we generated two sets of data from model (1). We chose the function m plotted in Figure 1, and the time transformation function g_0 defined as

$$g_0(t) = 0.01t^2 + 0.9t.$$

The time points for both the datasets were generated as samples from the uniform distribution over the interval [0, 20]. The distributions of ϵ_t and ϵ'_s were choosen as normal with zero mean and standard deviation 0.02. The sample sizes for the two datasets were $n_1 = n_2 = 500$.

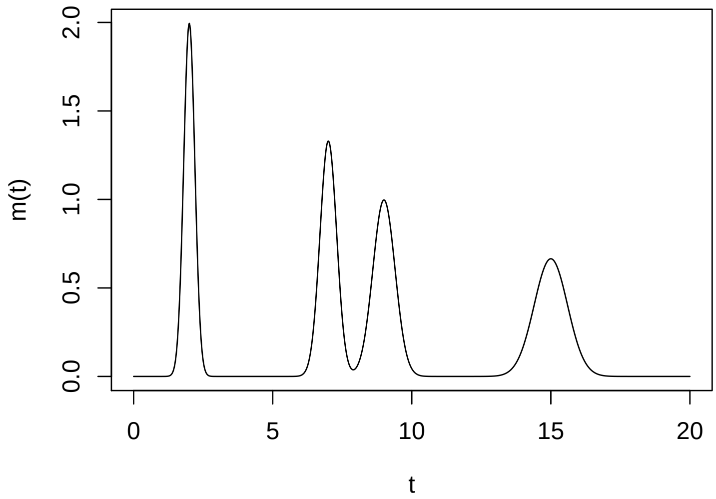

Figure 1: Plot of the function m used in the Simulation Study

For the method described in Section 2, we chose K_1 and K_2 as standar normal densities. We chose the bandwidths in the manner mentioned in Section 2 with $c_1 = 10$ and $c_2 = 10$. The search for the transformation function g was made over the class of all linear and quadratic function defined over the interval [0, 20]. The method of Steepest Ascent was used for optimization while using the identity map as initial iterate.

For implementing the method of Ramsay and Li (1998) we use a function *register.fd* available in fda library of the R Statistical Package (see Ramsay, Hooker and Graves (2009)). The smoothing parameter and the order of the polynomial spline were choosen as .001 and 4, respectively. Following the pattern used in the examples given in Ramsay and Li (1998) we choose number of knots as 6.

The performance of the two estimators of the time-transformation functions are studied in terms of the following quantities: (a) point-wise bias, (b) point-wise standard deviation, (c) point-wise mean squared error and (d) Average of the integrated mean square error normalized by the squared norm of the true function, defined for each simulation run as

$$\frac{\frac{1}{S}\int_0^{20}\sum_{j=0}^{S}(\hat{g}_j(t)-g(t))^2dt}{\int_0^{20}g^2(t)dt}$$

where S is the number of independent runs of the simulation and $\hat{g}_j$ is the estimate of g at the jth run, with the numerator evaluated numerically over a uniform grid of 1000 points. The results of the simulations based on 500 independent runs are shown in Figures 2, 3 and 4. It is found that the

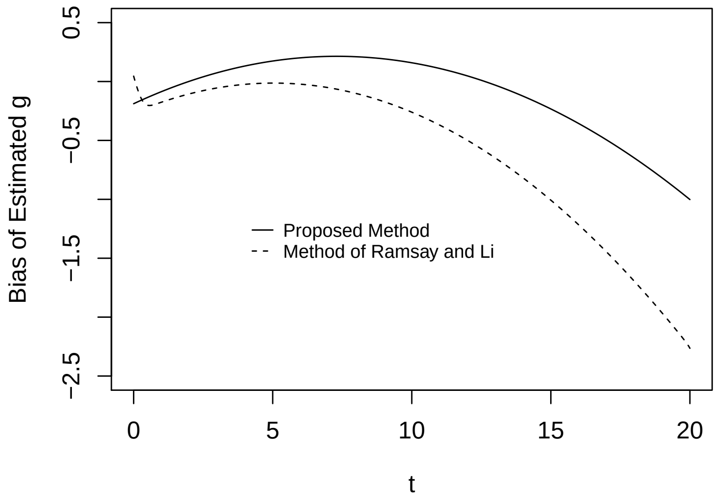

Figure 2: Point-wise bias of the two estimators

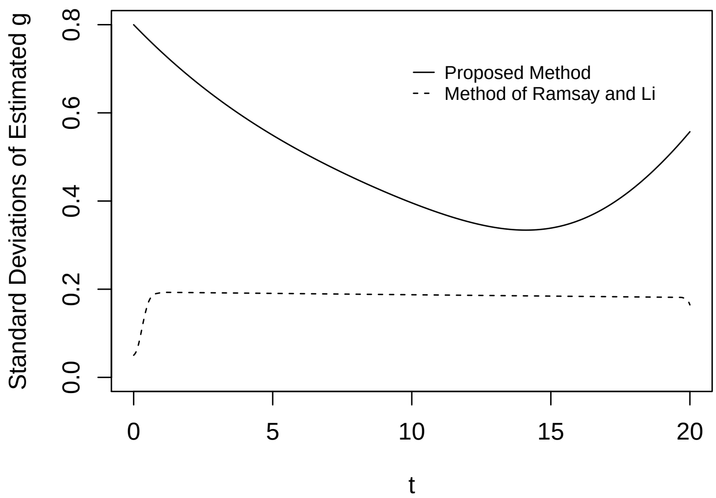

Figure 3: Point-wise standard deviations of the two estimators

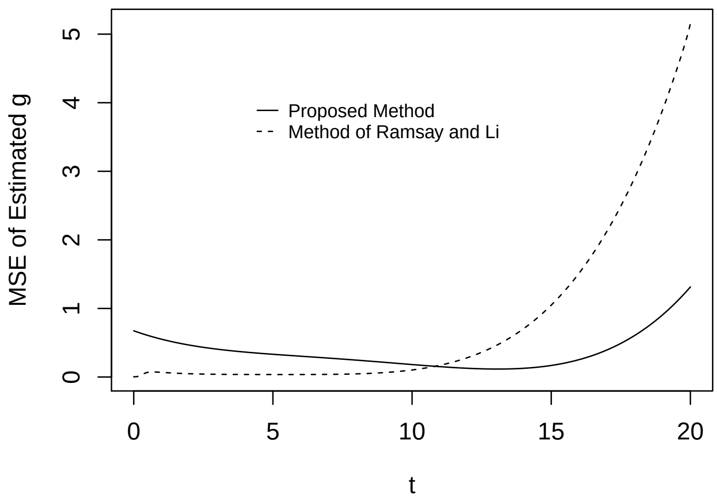

Figure 4: Point-wise mean squared errors of the two estimators

proposed method generally has higher variance but smaller bias and mean squared error than the method of Ramsay and Li. The normalized integrated mean squared errors of the estomator proposed in section 2 and that of Ramsay and Li turned out to be 0.002427 and 0.005473, respectively. The superior performance of the proposed estimator is evident.

5 Data Analysis

We now illustrate the proposed method with paleoclimatico data on the atmospheric concentration of carbon-di-oxide as determined from air-bubbles trapped in ice cores collected over lake Vostok and Epica Dome of Antarctica. The data were downloaded from the website of the World Data Center for Paleoclimatology, Boulder and NOAA Paleoclimatology Program (www.ncdc.noaa.gov). The data used for analysis corresponded to the nominal age range of 215,000 to 385,000 years. Since the age of the sample was determined by radio-isotope dating, there is a possibility of mismatch. The two sets of functional data, plotted in Figure 5 indeed show a considerable amount of mismatch. We choose to align the Epica Dome dataset with the Vostok dataset as reference.

The method proposed by Ramsay and Li, with choices of the first two parameters as in the foregoing section and number of knot points equal to 15, led to the alignment shown in Figure 6. The average squared difference between the two curves (computed from time points of the first data set) turned out to be 308.20. The corresponding value obtained from the method of section 2,

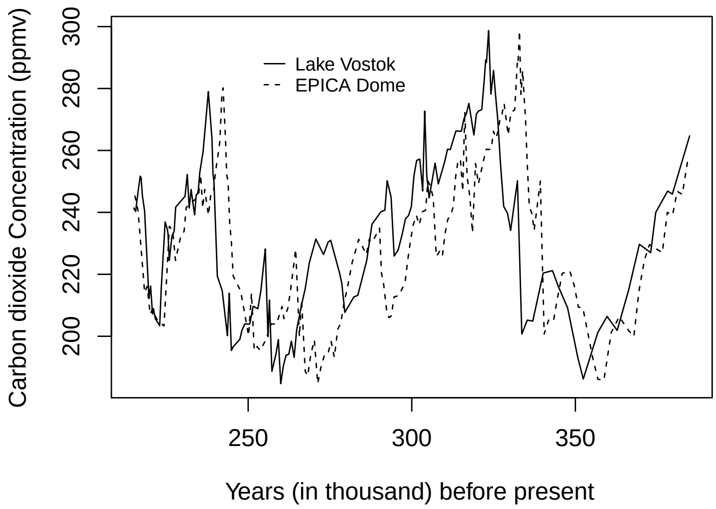

Figure 5: Ice core datasets

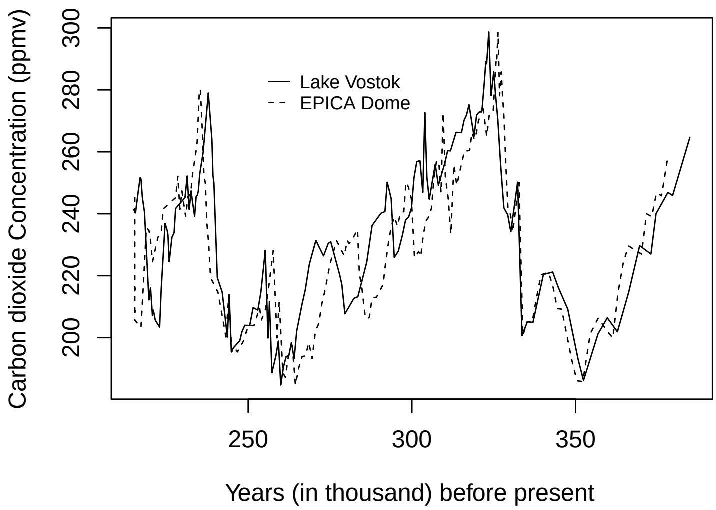

Figure 6: Ice core datasets after alignment by the method of Ramsay and Li

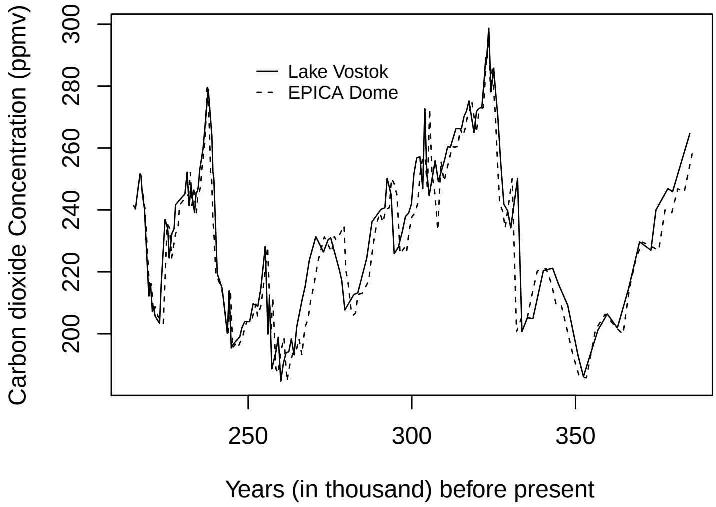

Figure 7: Ice core datasets after alignment by the proposed method

which produced the alignment shown in Figure 7, was 96.99. The proposed method is found to produce better alignment.

Appendix

A Proof of theoretical results

When Assumption 1 holds, we denote by M_f a positive upper bound for the densities $f_{\epsilon 1}$, $f_{\epsilon 2}$, f_1 and f_2. Let the positive real numbers c, M_K, M_K' be such that $0 < c \le K_i(x) \le M_K$ and $|K_i'(x)| \le M_K'$ for $i = 1, 2$ whenever Assumptions 2 and 2A hold.

Proof of Theorem 3.1 For a given time transformation $g \in \mathbb{G}$, from (2), we have

$$C_n(g) = \frac{N_n(g)}{D_n(g)}, \tag{5}$$

where

$$N_n(g) = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \frac{1}{h_1} K_1\left(\frac{t_i - g(s_j)}{h_1}\right) \frac{1}{h_2} K_2\left(\frac{Y_{t_i} - Y'_{s_j}}{h_2}\right), \tag{6}$$

$$D_n(g) = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \frac{1}{h_1} K_1\left(\frac{t_i - g(s_j)}{h_1}\right). \tag{7}$$

We first establish that

$$N_n(g) \xrightarrow{P} \int_{-\infty}^{\infty}\int_0^{\infty} f_1(g(y))f_2(y)f_{\epsilon 1}(v - m(g(y)) + m(g_0(y)))f_{\epsilon 2}(v)dydv \qquad (8)$$

$$D_n(g) \xrightarrow{P} \int_0^{\infty} f_1(g(y))f_2(y)dy. \qquad (9)$$

The proof is then completed by using the continuous mapping theorem of convergence in probability (see Billingsley, 1985).

Note that, from (6), we have

$$E[N_n(g)] = \frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} E\left[\frac{1}{h_1}K_1\left(\frac{t_i - g(s_j)}{h_1}\right)\frac{1}{h_2}K_2\left(\frac{Y_{t_i} - Y'_{s_j}}{h_2}\right)\right].$$

From the description of the model (1), we have

$$E[N_n(g)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_0^{\infty}\int_0^{\infty}\frac{1}{h_1}K_1\left(\frac{x - g(y)}{h_1}\right)\frac{1}{h_2}K_2\left(\frac{m(x) - m(g_0(y)) + u - v}{h_2}\right)$$
$$\times f_1(x)f_2(y)f_{\epsilon 1}(u)f_{\epsilon 2}(v)dxdydudv. \qquad (10)$$

By making the transformations $z_1 = \frac{x-g(y)}{h_1}$ and $z_2 = \frac{m(x)-m(g_0(y))+u-v}{h_2}$, we have

$$E[N_n(g)] = \int_{-\infty}^{\infty}\int_0^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S_n(z_1, z_2, y, v)dz_1dz_2dydv \qquad (11)$$

where

$$S_n(z_1, z_2, y, v) = I_{(-g(y)/h_1,\infty)}(z_1)K_1(z_1)K_2(z_2)f_1(g(y) + z_1h_1)$$
$$\times f_2(y)f_{\epsilon 1}(v - m(g(y) + z_1h_1) + m(g_0(y)) + z_2h_2)f_{\epsilon 2}(v). \qquad (12)$$

As g is a positive and increasing function, any given real z_1 is contained in $(-g(y)/h_1,\infty)$ for sufficiently small h_1. Assumption 4, for any given z_1 we have By using Assumptions 1, 3 and 4 and the fact that m is a continuous function, for any fixed (z_1, z_2, y, v), we have

$$\lim_{n\to\infty} S_n(z_1, z_2, y, v) = K_1(z_1)K_2(z_2)f_1(g(y))f_2(y)f_{\epsilon 1}(v - m(g(y)) + m(g_0(y)))f_{\epsilon 2}(v). \qquad (13)$$

Note that, from Assumption 1, we have

$$0 \le S_n(z_1, z_2, y, v) \le M_f^2 K_1(z_1)K_2(z_2)f_2(y)f_{\epsilon 2}(v). \qquad (14)$$

Assumption 2 ensures that the bounding function on the right hand side of (14) is integrable. Then, by using the Dominated Convergence Theorem (DCT), we have

$$\lim_{n\to\infty} E[N_n(g)] = \int_{-\infty}^{\infty}\int_0^{\infty} f_1(g(y))f_2(y)f_{\epsilon 1}(v - m(g(y)) + m(g_0(y)))f_{\epsilon 2}(v)dydv. \qquad (15)$$

From Lemma A.1, we have $\lim_{n\to\infty} Var[N_n(g)] = 0$. This establishes (8).

We now turn to $D_n(g)$. By using (7) and making the transformation $z = \frac{x-g(y)}{h_1}$, we have

$$E[D_n(g)] = \int_0^\infty \int_{-\infty}^\infty R_n(z,y)dzdy, \tag{16}$$

where

$$R_n(z,y) = I_{\left(\frac{-g(y)}{h_1},\infty\right)}(z)K_1(z)f_1(g(y)+zh_1)f_2(y). \tag{17}$$

From (17) and Assumptions 1, 3 and 4, we have

$$\lim_{n\to\infty} R_n(z,y) = K_1(z)f_1(g(y))f_2(y).$$

By Assumption 1, we have the dominance

$$0 \le R_n(z,y) \le M_f K_1(z)f_2(y). \tag{18}$$

Assumption 1 and 2 ensures that the bounding function, on the right hand side of (18), is integrable. Thus, by applying DCT we have,

$$\lim_{n\to\infty} E[D_n(g)] = \int_0^\infty f_1(g(y))f_2(y)dy. \tag{19}$$

Now from Lemma A.1, we have $\lim_{n\to} Var[D_n(g)] = 0$. This establishes (9) and completes the proof. $\square$

Lemma A.1 *Under assumptions 1, 2, 3 and 4, for any $g \in \mathbb{G}$, we have*

$$\lim_{n\to\infty} Var[N_n(g)] = 0, \tag{20}$$
$$\lim_{n\to\infty} Var[D_n(g)] = 0, \tag{21}$$

where $N_n(g)$ and $D_n(g)$ are defined in (6) and (7) respectively.

Proof of Lemma A.1 From (6), we have

$$\begin{aligned}
&Var(N_n(g)) \\
&= \frac{1}{(n_1n_2h_1h_2)^2}\sum_{i=1}^{n1}\sum_{j=1}^{n2}\sum_{i'=1}^{n1}\sum_{j'=1}^{n2} Cov\Bigg\{K_1\left(\frac{t_i-g(s_j)}{h_1}\right)K_2\left(\frac{Y_{t_i}-Y_{s_j}}{h_2}\right), \\
&\qquad K_1\left(\frac{t_{i'}-g(s_{j'})}{h_1}\right)K_2\left(\frac{Y_{t_{i'}}-Y_{s_{j'}}}{h_2}\right)\Bigg\} \\
&= V_1+V_2+V_3+V_4,
\end{aligned} \tag{22}$$

where

$$V_1 = \frac{1}{(n_1n_2h_1h_2)^2}\sum_{i=1}^{n1}\sum_{j=1}^{n2} Var\Big\{K_1\left(\frac{t_i - g(s_j)}{h_1}\right)K_2\left(\frac{Y_{t_i} - Y_{s_j}}{h_2}\right)\Big\}, \tag{23}$$

$$V_2 = \frac{1}{(n_1n_2h_1h_2)^2}\sum_{i=1}^{n1}\sum_{j=1}^{n2}\sum_{\substack{i'=1\\(\neq i)}}^{n1} Cov\Big\{K_1\left(\frac{t_i - g(s_j)}{h_1}\right)K_2\left(\frac{Y_{t_i} - Y_{s_j}}{h_2}\right) \tag{24}$$

$$K_1\left(\frac{t_{i'} - g(s_j)}{h_1}\right)K_2\left(\frac{Y_{t_{i'}} - Y_{s_j}}{h_2}\right)\Big\}, \tag{25}$$

$$V_3 = \frac{1}{(n_1n_2h_1h_2)^2}\sum_{i=1}^{n1}\sum_{j=1}^{n2}\sum_{\substack{j'=1\\(\neq j)}}^{n2} Cov\Big\{K_1\left(\frac{t_i - g(s_j)}{h_1}\right)K_2\left(\frac{Y_{t_i} - Y_{s_j}}{h_2}\right),$$

$$K_1\left(\frac{t_i - g(s_{j'})}{h_1}\right)K_2\left(\frac{Y_{t_i} - Y_{s_{j'}}}{h_2}\right)\Big\}, \tag{26}$$

$$V_4 = \frac{1}{(n_1n_2h_1h_2)^2}\sum_{i=1}^{n1}\sum_{j=1}^{n2}\sum_{\substack{i'=1\\(\neq i)}}^{n1}\sum_{\substack{j'=1\\(\neq j)}}^{n2} Cov\Big\{K_1\left(\frac{t_i - g(s_j)}{h_1}\right)K_2\left(\frac{Y_{t_i} - Y_{s_j}}{h_2}\right),$$

$$K_1\left(\frac{t_{i'} - g(s_{j'})}{h_1}\right)K_2\left(\frac{Y_{t_{i'}} - Y_{s_{j'}}}{h_2}\right)\Big\}. \tag{27}$$

We consider the convergence of each term of right hand side of (22) separately. Consider the term V_1, then by using the model specifications (1) and (23), we have

$$V_1 = \frac{1}{n_1n_2h_1h_2}V_{11} - \frac{1}{n_1n_2}(E[N_n(g)])^2, \tag{28}$$

where

$$V_{11} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{0}^{\infty}\int_{0}^{\infty}\frac{1}{h_1}K_1^2\left(\frac{x-g(y)}{h_1}\right)\frac{1}{h_2}K_2^2\left(\frac{m(x)-m(g_0(y))+u-v}{h_2}\right)$$
$$\times f_1(x)f_2(y)f_{\epsilon 1}(u)f_{\epsilon 2}(v)dxdydudv. \tag{29}$$

By making the transformations $z_1 = \frac{x-g(y)}{h_1}$ and $z_2 = \frac{m(x)-m(g_0(y))+u-v}{h_2}$, we have

$$V_{11} = \int_{-\infty}^{\infty}\int_{0}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I_{(\frac{-g(y)}{h_1},\infty)}(z_1)K_1^2(z_1)K_2^2(z_2)f_1(g(y)+z_1h1)$$
$$\times f_2(y)f_{\epsilon 1}(v - m(g(y)+z_1h_1) + m(g_0(y)) + z_2h_2)f_{\epsilon 2}(v)dz_1dz_2dydv. \tag{30}$$

From Assumptions 1, 3, and 4 and the continuity of m, for any fixed real (z_1, z_2, y, v), a similar argument as in the proof of (13), shows that the integrand function on the right hand side of (30), as $n \to \infty$, converges to

$$K_1^2(z_1)K_2^2(z_2)f_1(g(y))f_2(y)f_{\epsilon 1}(v - m(g(y)) + m(g_0(y)))f_{\epsilon 2}(v).$$

We have the dominance of the integrand function on the right hand side of (30) by the integrable function

$$M_f^2 \left\{K_1^2(z_1)\right\} \left\{K_2^2(z_2)\right\} f_2(y) f_{\epsilon 2}(v).$$

By using DCT and convergence of the integrand on the right hand side of (30), we have

$$\begin{aligned}\lim_{n\to\infty} V_{11} &= \int_{-\infty}^{\infty} K_1^2(z_1)dz_1 \int_{-\infty}^{\infty} K_2^2(z_2)dz_2 \\ &\quad \times \int_{-\infty}^{\infty}\int_0^{\infty} f_1(g(y)) f_2(y) f_{\epsilon 1}(v - m(g(y)) + m(g_0(y))) f_{\epsilon 2}(v) dy dv.\end{aligned}$$

Now, from (15), the second term on the right hand side of (28), we have

$$\frac{1}{n_1 n_2}(E[N_n(g)])^2 = O\left(\frac{1}{n_1 n_2}\right).$$

Thus, we have

$$V_1 = O\left(\frac{1}{n_1 n_2 h_1 h_2}\right) + O\left(\frac{1}{n_1 n_2}\right) = O\left(\frac{1}{n_1 n_2 h_1 h_2}\right). \tag{31}$$

We now consider the term V_2. From (25) and by using model specifications (1), we have

$$V_2 = \frac{n_1 - 1}{n_1 n_2} V_{21} - \frac{n_1 - 1}{n_1 n_2}(E[N_n(g)])^2, \tag{32}$$

Where

$$\begin{aligned} V_{21} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_0^{\infty}\int_0^{\infty}\int_0^{\infty} \frac{1}{h_1}K_1\left(\frac{x - g(y)}{h_1}\right)\frac{1}{h_2}K_2\left(\frac{m(x) - m(g_0(y)) + u - v}{h_2}\right) \\ &\quad \times \frac{1}{h_1}K_1\left(\frac{x' - g(y))}{h_1}\right)\frac{1}{h_2}K_2\left(\frac{m(x') - m(g_0(y)) + u' - v}{h_2}\right) \\ &\quad \times f_1(x) f_1(x') f_2(y) f_{\epsilon 1}(u) f_{\epsilon 1}(u') f_{\epsilon 2}(v) dx dx' dy du du' dv. \end{aligned} \tag{33}$$

By making the transformations $z_1 = \frac{x - g(y)}{h_1}$, $z_2 = \frac{x' - g(y)}{h_1}$, $z_3 = \frac{m(x) - m(g_0(y)) + u - v}{h_2}$, and $z_4 = \frac{m(x') - m(g_0(y)) + u' - v}{h_2}$, the integrand on the right hand side of (33) is

$$\begin{aligned} &I_{(-g(y)/h_1,\infty)}(z_1)K_1(z_1)I_{(-g(y)/h_1,\infty)}(z_2)K_1(z_2)K_2(z_3)K_2(z_4)f_1(g(y) + z_1h_1)f_1(g(y) + z_2h_1)f_2(y) \\ &\times f_{\epsilon 1}(v - m(g(y) + z_1h_1) + m(g_0(y)) + h_2z_3)f_{\epsilon 1}(v - m(g(y) + z_2h_1) + m(g_0(y)) + h_2z_4)f_{\epsilon 2}(v). \end{aligned} \tag{34}$$

From Assumptions 1, 3, and 4 and the continuity of m it follows via a similar argument as in (13) that for any z_1, z_2, z_3, z_4, y and v the above function converges, as $n \to \infty$, to

$$K_1(z_1)K_1(z_2)K_2(z_3)K_2(z_4)[f_1(g(y))]^2 f_2(y)[f_{\epsilon 1}(v - m(g(y)) + m(g_0(y)))]^2 f_{\epsilon 2}(v).$$

The integrand function on the right hand side of (34) is dominated by the integrable function

$$M_f^4 K_1(z_1)K_1(z_2)K_2(z_3)K_2(z_4)f_2(y)f_{\epsilon 2}(v).$$

By using Assumption 2, and the convergence of (34) and applying DCT on the right hand side of (33), we have

$$\lim_{n\to\infty} V_{21} \quad = \quad \int_0^\infty \int_{-\infty}^\infty [f_1(g(y))]^2 f_2(y)[f_{\epsilon 1}(v - m(g(y)) + m(g_0(y)))]^2 f_{\epsilon 2}(v)dvdy.$$

Now, from (15), the second term on the right hand side of (32) turns out to be

$$\frac{n_1 - 1}{n_1 n_2}(E[N_n(g)])^2 = O\left(\frac{1}{n_2}\right).$$

$$n_2 V_2 \to \int_0^\infty \int_{-\infty}^\infty [f_1(g(y))]^2 f_2(y)[f_{\epsilon 1}(v - m(g(y)) + m(g_0(y)))]^2 f_{\epsilon 2}(v)dvdy - [E(N_n(g))]^2$$

i.e.,

$$V_2 = O\left(\frac{1}{n_2}\right). \tag{35}$$

By using a similar argument as for the term V_2, we have

$$n_2 V_2 \to \int_0^\infty \int_{-\infty}^\infty \frac{[f_2(y)]^2}{g'(y)} f_1(g(y))f_{\epsilon 1}(v - m(g(y)) + m(g_0(y)))[f_{\epsilon 2}(v)]^2 dvdy - [E(N_n(g))]^2$$

i.e.,

$$V_3 = O\left(\frac{1}{n_1}\right). \tag{36}$$

Finally, we consider the term V_4. By using the model specification (1), we have

$$V_4 = 0. \tag{37}$$

The proof of (20) is completed from (31), (35), (36), (37) and by using Assumption 3 and 4.

We now compute the $Var[D_n(g)]$. Note that, from (7), we have

$$Var[D_n(g)] \quad = \quad T_1 + T_2 + T_3 + T_4,$$

Where

$$T_1 = \frac{1}{(n_1n_2h_1)^2}\sum_{i=1}^{n1}\sum_{j=1}^{n2} Var\left\{K_1\left(\frac{t_i - g(s_j)}{h_1}\right)\right\}, \tag{38}$$

$$T_2 = \frac{1}{(n_1n_2h_1)^2}\sum_{i=1}^{n1}\sum_{j=1}^{n2}\sum_{\substack{i'=1\\(\neq i)}}^{n1} Cov\left\{K_1\left(\frac{t_i - g(s_j)}{h_1}\right), K_1\left(\frac{t_{i'} - g(s_j)}{h_1}\right)\right\}, \tag{39}$$

$$T_3 = \frac{1}{(n_1n_2h_1)^2}\sum_{i=1}^{n1}\sum_{j=1}^{n2}\sum_{\substack{j'=1\\(\neq j)}}^{n2} Cov\left\{K_1\left(\frac{t_i - g(s_j)}{h_1}\right), K_1\left(\frac{t_i - g(s_{j'})}{h_1}\right)\right\}, \tag{40}$$

$$T_4 = \frac{1}{(n_1n_2h_1)^2}\sum_{i=1}^{n1}\sum_{j=1}^{n2}\sum_{\substack{i'=1\\(\neq i)}}^{n1}\sum_{\substack{j'=1\\(\neq j)}}^{n2} Cov\left\{K_1\left(\frac{t_i - g(s_j)}{h_1}\right), K_1\left(\frac{t_{i'} - g(s_{j'})}{h_1}\right)\right\}. \tag{41}$$

We consider the convergence for the terms T_1, T_2, T_3 and T_4 separately. Consider the term T_1. By making the transformation $z = \frac{x-g(y)}{h_1}$ and by using the model specification, we have

$$T_1 = \frac{1}{n_1n_2h_1}T_{11} - \frac{1}{n_1n_2}E^2[D_n(g)], \tag{42}$$

where

$$T_{11} = \int_0^\infty \int_{-\infty}^\infty I_{(-g(y)/h_1,\infty)}(z)K_1^2(z)f_1(g(y)+h_1z)f_2(y)dzdy. \tag{43}$$

Note that the integrand on the right hand side of (43) is bounded by the integrable function $M_fK_1^2(z)f_2(y)$. Further, a similar argument, as used for the $Var(N_n)$, shows that, for any given y and z, the integrand function, as $n\to\infty$, converges to $K_1^2(z)f_1(g(y))f_2(y)$. Thus, by applying DCT and Assumption 3 and 4, we have

$$\lim_{n\to\infty} T_{11} = \int_{-\infty}^\infty K_1^2(z)dz\int_0^\infty f_1(g(y))f_2(y)dy. \tag{44}$$

Now, from (19), the second term on the right hand side of (42) turns out to be

$$\frac{1}{n_1n_2}E^2[D_n(g)] = O\left(\frac{1}{n_1n_2}\right).$$

Thus, we have

$$T_1 = O\left(\frac{1}{n_1n_2h_1}\right) + O\left(\frac{1}{n_1n_2}\right) = O\left(\frac{1}{n_1n_2h_1}\right).$$

We now consider the term T_2. From (39) and the model specifications (1) and by making the transformations $z = \frac{x-g(y)}{h_1}$ and $z' = \frac{x'-g(y)}{h_1}$, we have

$$T_2 = \frac{n_1-1}{n_1n_2}T_{21} - \frac{n_1-1}{n_1n_2}E^2[D_n(g)], \tag{45}$$

where

$$T_{21} = \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty I_{(-g(y)/h_1,\infty)}(z) K_1(z) I_{(-g(y)/h_1,\infty)}(z') K_1(z') \times f_1(g(y)+zh_1) f_1(g(y)+z'h_1) f_2(y) dz dz' dy \qquad (46)$$

Note that, the integrand on the right hand side of (46) is bounded by the integrable function $M_f^2 K_1(z)K_1(z')f_2(y)$. A similar argument as used for the convergence of the term T_{11} shows that, for any given z, z' and y, the integrand function converges, as $n \to \infty$, to $K_1(z)K_1(z')f_1^2(g(y))f_2(y)$. Thus, by applying DCT and Assumption 2, 3 and 4, we have

$$\lim_{n\to\infty} T_{21} = \int_0^\infty f_1^2(g(y))f_2(y)dy. \qquad (47)$$

Now, from (19), the second term on the right hand side of (45) turns out to be

$$\frac{n_1 - 1}{n_1 n_2} E^2[D_n(g)] = O\left(\frac{1}{n_2}\right).$$

Thus, we have

$$n_2 T_2 \to \int_0^\infty f_1^2(g(y))f_2(y)dy - E^2[D_n(g)].$$

i.e.,

$$T_2 = O\left(\frac{1}{n_2}\right).$$

We now consider the term T_3. A similar argument, as used for the convergence of the term T_2, shows that

$$n_1 T_3 \to \int_0^\infty \frac{f_1(g(y))[f_2(y)]^2}{g'(y)} dy - E^2[D_n(g)].$$

i.e.,

$$T_3 = O\left(\frac{1}{n_1}\right).$$

Finally, by using the model specification, we have $T_4 = 0$. This completes the proof of (21). $\square$

Proof of Theorem 3.2 We first obtain a stochastic upper bound of the oscillation of $C_n(g)$. From (5), for any given $\tilde{g}, g \in \mathbb{G}$, we have

$$\begin{aligned} |C_n(\tilde{g}) - C_n(g)| &= \left| \frac{N_n(\tilde{g})}{D_n(\tilde{g})} - \frac{N_n(g)}{D_n(g)} \right|, \\ &= \frac{|N_n(\tilde{g})D_n(g) - N_n(g)D_n(\tilde{g})|}{D_n(\tilde{g})D_n(g)}. \end{aligned}$$

Observe that by using (7) and Assumption 2A, we have $D_n(g) \geq \frac{c}{h_1}$. Therefore,

$$\begin{aligned} |C_n(\tilde{g}) - C_n(g)| &\leq \frac{h_1^2}{c^2} \left| N_n(\tilde{g})D_n(g) - N_n(g)D_n(\tilde{g}) \right|, \\ &\leq \frac{h_1^2}{c^2} \left\{ D_n(\tilde{g}) \left| N_n(\tilde{g}) - N_n(g) \right| + N_n(\tilde{g}) \left| D_n(\tilde{g}) - D_n(g) \right| \right\}. \qquad (48) \end{aligned}$$

We now compute the upper bounds for both the terms on the right hand side of (48). Note that, from (6), we have

$$\begin{aligned}
|N_n(\tilde g) - N_n(g)| &= \left| \frac{1}{n_1 n_2} \sum_{i=1}^{n1} \sum_{j=1}^{n2} \frac{1}{h_1} \left\{ K_1\left(\frac{t_i - \tilde g(s_j)}{h_1}\right) - K_1\left(\frac{t_i - g(s_j)}{h_1}\right) \right\} \frac{1}{h_2} K_2\left(\frac{Y_{t_i} - Y_{s_j}}{h_2}\right) \right|, \\
&\le \frac{1}{n_1 n_2 h_1 h_2} \sum_{i=1}^{n1} \sum_{j=1}^{n2} K_2\left(\frac{Y_{t_i} - Y_{s_j}}{h_2}\right) \left| K_1\left(\frac{t_i - \tilde g(s_j)}{h_1}\right) - K_1\left(\frac{t_i - g(s_j)}{h_1}\right) \right|.
\end{aligned}$$

By using the mean value theorem, we have

$$|N_n(\tilde g) - N_n(g)| \le \frac{1}{n_1 n_2 h_1 h_2} \sum_{i=1}^{n1} \sum_{j=1}^{n2} K_2\left(\frac{Y_{t_i} - Y_{s_j}}{h_2}\right) \left| K_1'(x_0(t_i, s_j, \tilde g, g)) \right| \left| \frac{g(s_j) - \tilde g(s_j)}{h_1} \right|, \tag{49}$$

where $x_0(t_i, s_j, \tilde g, g) \in \left(\min\left(\frac{t_i - \tilde g(s_j)}{h_1}, \frac{t_i - g(s_j)}{h_1}\right), \max\left(\frac{t_i - \tilde g(s_j)}{h_1}, \frac{t_i - g(s_j)}{h_1}\right)\right)$. Now, from (49) and Assumption 2A, we have

$$|N_n(\tilde g) - N_n(g)| \le \frac{M_K'}{h_1^2} \cdot \|g - \tilde g\| \cdot U_n, \tag{50}$$

where

$$U_n = \frac{1}{n_1 n_2 h_2} \sum_{i=1}^{n1} \sum_{j=1}^{n2} K_2\left(\frac{Y_{t_i} - Y_{s_j}}{h_2}\right). \tag{51}$$

We now turn to the second term on the right hand side of (48). From (7), we have

$$|D_n(\tilde g) - D_n(g)| \le \frac{1}{n_1 n_2 h_1} \sum_{i=1}^{n1} \sum_{j=1}^{n2} \left| K_1\left(\frac{t_i - \tilde g(s_j)}{h_1}\right) - K_1\left(\frac{t_i - g(s_j)}{h_1}\right) \right|.$$

From Assumption 2A and the mean value theorem, we have

$$\begin{aligned}
|D_n(\tilde g) - D_n(g)| &\le \frac{1}{n_1 n_2 h_1} \sum_{i=1}^{n1} \sum_{j=1}^{n2} \left| K_1'(x_0(t_i, s_j, \tilde g, g)) \right| \left| \frac{g(s_j) - \tilde g(s_j)}{h_1} \right|, \\
&\le \frac{M_K'}{h_1^2} \cdot \|g - \tilde g\|, \qquad (52)
\end{aligned}$$

where $x_0(\cdot)$ is as in (49). Now, by using (50), (52) and (48), we have

$$|C_n(\tilde g) - C_n(g)| \le B_n(\tilde g) \cdot \|g - \tilde g\|, \tag{53}$$

where

$$B_n(\tilde g) = \frac{M_K'}{c^2} \left\{ N_n(\tilde g) + U_n \cdot D_n(\tilde g) \right\}.$$

The expression on the right hand side of (53) gives an upper bound on the oscillation of the functional $C_n(\cdot)$ over the time transformation functions in $\mathbb{G}$.

Note that

$$|C_n(g) - C(g)| \leq |C_n(g) - C_n(\tilde{g})| + |C(\tilde{g}) - C(g)| + |C_n(\tilde{g}) - C(\tilde{g})|, \tag{54}$$

where C is defined as in (4).

Set $\epsilon > 0$.

Lemma A.3 implies that there exists $\delta_\epsilon > 0$ such that

$$\|g - \tilde{g}\| < \delta_\epsilon \text{ implies } |C(g) - C(\tilde{g})| < \frac{\epsilon}{3}. \tag{55}$$

Theorem 3.1 and Lemma A.2 implies that for all $\tilde{g}$ there exists $M_{\tilde{g}}$ such that $B_n(\tilde{g}) \xrightarrow{P} M_{\tilde{g}}$, which ensures

$$P\left(B_n(\tilde{g}) > \max\left\{\frac{\epsilon}{3\delta_\epsilon}, 2M_{\tilde{g}}\right\}\right) \to 0. \tag{56}$$

Define $\mathcal{N}_\eta(\tilde{g}) = \{g : \|g - \tilde{g}\| < \eta\}$. For given $\tilde{g}$, let

$$\delta(\tilde{g}, \epsilon) = \begin{cases} \min\left\{\frac{\epsilon}{6M_{\tilde{g}}}, \delta_\epsilon\right\} & \text{if } \ M_{\tilde{g}} > 0 \\ \delta_\epsilon & \text{if } \ M_{\tilde{g}} = 0 \end{cases} \tag{57}$$

For g in $\mathcal{N}_{\delta(\tilde{g},\epsilon)}(\tilde{g})$, from (53) we have,

$$|C_n(\tilde{g}) - C_n(g)| < \delta(\tilde{g}, \epsilon) \cdot B_n(\tilde{g}). \tag{58}$$

Note that $\left\{\mathcal{N}_{\delta(\tilde{g},\epsilon)}(\tilde{g}) : \tilde{g} \in \mathbb{G}\right\}$ is an open cover of $\mathbb{G}$. By Assumption 5 there exists a finite subcover say $\left\{\mathcal{N}_{\delta(\tilde{g}_j,\epsilon)}(\tilde{g}_j)\right\}_{j=1\ldots k_\epsilon}$ i.e. $\mathbb{G} \subset \cup_{j=1}^{k_\epsilon} \mathcal{N}_{\delta(\tilde{g}_j,\epsilon)}(\tilde{g}_j)$ for some finite k_ϵ. From (54), (55), and (58) we have,

$$\begin{aligned}
&\sup_{g \in \mathbb{G}} |C_n(g) - C(g)| \\
&\quad \leq \max_{j=1,\ldots,K_\epsilon} \sup_{g \in \mathcal{N}_{\delta(\tilde{g}_j,\epsilon)}(\tilde{g}_j)} |C_n(g) - C(g)|, \\
&\quad \leq \max_{j=1,\ldots,k_\epsilon} \left\{ \sup_{g \in \mathcal{N}_{\delta(\tilde{g}_j,\epsilon)}(\tilde{g}_j)} |C_n(g) - C(g)| + \sup_{g \in \mathcal{N}_{\delta(\tilde{g}_j,\epsilon)}(\tilde{g}_j)} |C(\tilde{g}_j) - C(g)| + \sup_{g \in \mathcal{N}_{\delta(\tilde{g}_j,\epsilon)}(\tilde{g}_j)} |C_n(\tilde{g}_j) - C(\tilde{g}_j)| \right\}, \\
&\quad \leq \max_{1,\ldots,k_\epsilon} \left\{ \delta(\tilde{g}_j, \epsilon) B_n(\tilde{g}_j) + \frac{\epsilon}{3} + |C_n(\tilde{g}_j) - C(\tilde{g}_j)| \right\} \\
&\quad \leq \max_{1,\ldots,k_\epsilon} \delta(\tilde{g}_j, \epsilon) B_n(\tilde{g}_j) + \frac{\epsilon}{3} + \sum_{j=1}^{k_\epsilon} |C_n(\tilde{g}_j) - C(\tilde{g}_j)|.
\end{aligned} \tag{59}$$

From (57) and (59) we have,

$$
\begin{aligned}
& P\left\{\sup_{g\in\mathbb{G}}|C_n(g)-C(g)|>\epsilon\right\} && (60)\\
\leq\ & P\left\{\max_{j=1}^{k_\epsilon}\delta(\tilde{g}_j,\epsilon)B_n(\tilde{g}_j)>\frac{\epsilon}{3}\right\}+P\left\{\sum_{j=1}^{k_\epsilon}|C_n(\tilde{g}_j)-C(\tilde{g}_j)|>\frac{\epsilon}{3}\right\}\\
=\ & \sum_{j=1}^{k_\epsilon}P\left\{B_n(\tilde{g}_j)>\frac{\epsilon}{3\delta(\tilde{g}_j,\epsilon)}\right\}+P\left\{\sum_{j=1}^{k_\epsilon}|C_n(\tilde{g}_j)-C(\tilde{g}_j)|>\frac{\epsilon}{3}\right\}. && (61)
\end{aligned}
$$

Each summand of the first term on the right hand side of (60) goes to zero by (56), while the second term goes to zero by Theorem 3.1. This completes the proof. □

Lemma A.2 *Let*

$$U_n=\frac{1}{n_1n_2h_2}\sum_{i=1}^{n1}\sum_{j=1}^{n2}K_2\left(\frac{Y_{t_i}-Y_{s_j}}{h_2}\right). \quad (62)$$

Then, under assumption 1, 2, 3 and 4,

$$U_n\xrightarrow{P}\int_{-\infty}^{\infty}\int_0^{\infty}\int_0^{\infty}f_1(x)f_2(y)f_{\epsilon_1}(v-m(x)+m(g_0(y)))f_{\epsilon_2}(v)dxdydv. \quad (63)$$

Proof: From (62) and the model specification, we have

$$E(U_n)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_0^{\infty}\int_0^{\infty}\frac{1}{h_2}K_2\left(\frac{m(x)-m(g_0(y))+u-v}{h_2}\right)f_1(x)f_2(y)f_{\epsilon_1}(u)f_{\epsilon_2}(v)dxdydudv. \quad (64)$$

By making the transformation $w=\frac{m(x)-m(g_0(y))+u-v}{h_2}$, we have

$$E(U_n)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_0^{\infty}\int_0^{\infty}K_2(w)f_1(x)f_2(y)f_{\epsilon_1}(v-m(x)+m(g(y))+wh_2)f_{\epsilon_2}(v)dxdydwdv. \quad (65)$$

From Assumption 1, 3 and 4, the integrand on the right hand side of (65), as $n\to\infty$, converges to

$$K_2(w)f_1(x)f_2(y)f_{\epsilon_1}(v-m(x)+m(g(y)))f_{\epsilon_2}(v),$$

and is bounded by the integrable function

$$M_fK_2(w)f_1(x)f_2(y)f_{\epsilon_2}(v).$$

By applying DCT, we have

$$\lim_{n\to\infty}E(U_n)=\int_{-\infty}^{\infty}\int_0^{\infty}\int_0^{\infty}f_1(x)f_2(y)f_{\epsilon_1}(v-m(x)+m(g(y)))f_{\epsilon_2}(v)dxdydv. \quad (66)$$

We now turn to the variance of Un. From (62), we have

$$Var(U_n) \quad = \quad V_1 + V_2 + V_3 + V_4, \tag{67}$$

where

$$V_1 \quad = \quad \frac{1}{(n_1n_2h_2)^2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} Var\left\{K_2\left(\frac{Y_{t_i}-Y_{s_j}}{h_2}\right)\right\}, \tag{68}$$

$$V_2 \quad = \quad \frac{1}{(n_1n_2h_2)^2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{\substack{i'=1\\(\neq i)}}^{n_1} Cov\left\{K_2\left(\frac{Y_{t_i}-Y_{s_j}}{h_2}\right), K_2\left(\frac{Y_{t_{i'}}-Y_{s_j}}{h_2}\right)\right\}, \tag{69}$$

$$V_3 \quad = \quad \frac{1}{(n_1n_2h_2)^2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{\substack{j'=1\\(\neq j)}}^{n_2} Cov\left\{K_2\left(\frac{Y_{t_i}-Y_{s_j}}{h_2}\right), K_2\left(\frac{Y_{t_i}-Y_{s_{j'}}}{h_2}\right)\right\}, \tag{70}$$

$$V_4 \quad = \quad \frac{1}{(n_1n_2h_2)^2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{\substack{i'=1\\(\neq i)}}^{n_1}\sum_{\substack{j'=1\\(\neq j)}}^{n_2} Cov\left\{K_2\left(\frac{Y_{t_i}-Y_{s_j}}{h_2}\right), K_2\left(\frac{Y_{t_{i'}}-Y_{s_{j'}}}{h_2}\right)\right\}. \tag{71}$$

We show that the terms V_i, for $i = 1, \ldots, 4$, converges to zero as $n \to \infty$. Consider the term V_1, from (68) and model specifications and by making the transformation $w = \frac{m(x)-m(g_0(y))+u-v}{h_2}$, we have

$$V_1 \quad = \quad \frac{1}{n_1n_2h_2}V_{11} - \frac{1}{n_1n_2}[E(U_n)]^2, \tag{72}$$

where

$$V_{11}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{0}^{\infty}\int_{0}^{\infty} K_2^2(w)f_1(x)f_2(y)f_{\epsilon_1}(v-m(x)+m(g_0(y))+h_2w)f_{\epsilon_2}(v)dxdydwdv. \tag{73}$$

From Assumption 1, 3 and 4, the integrand on the right hand side of (73), as $n \to \infty$, converges to

$$K_2^2(w)f_1(x)f_2(y)f_{\epsilon_1}(v - m(x) + m(g_0(y)))f_{\epsilon_2}(v)$$

and is dominated by the integrable function

$$M_f K_2^2(w)f_1(x)f_2(y)f_{\epsilon_2}(v).$$

Thus, by applying DCT, we have

$$\lim_{n\to\infty} V_{11}=\int_{-\infty}^{\infty} K_2^2(w)dw\int_{-\infty}^{\infty}\int_{0}^{\infty}\int_{0}^{\infty} f_1(x)f_2(y)f_{\epsilon_1}(v-m(x)+m(g_0(y)))f_{\epsilon_2}(v)dxdydv. \tag{74}$$

Now, from (66), the second term on the right hand side of (72) turns out to be

$$\frac{1}{n_1n_2}(E(U_n))^2 = O\left(\frac{1}{n_1n_2}\right).$$

Thus, we have

$$V_1 = O\left(\frac{1}{n_1 n_2 h_2}\right) + O\left(\frac{1}{n_1 n_2}\right) = O\left(\frac{1}{n_1 n_2 h_2}\right).$$

From (69) and the model specification and by making the transformation $w = \frac{m(x)-m(g_0(y))+u-v}{h_2}$, we have

$$V_2 = \frac{n_1 - 1}{n_1 n_2} V_{21} - \frac{n_1 - 1}{n_1 n_2} [E(U_n)]^2, \tag{75}$$

where

$$\begin{aligned} V_{21} & \\ = & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{0}^{\infty}\int_{0}^{\infty}\int_{0}^{\infty} K_2(w)K_2(w')f_1(x)f_2(y)f_1(x')f_{\epsilon_1}(v - m(x) + m(g_0(y)) + h_2 w) \\ & \times f_{\epsilon_1}(v - m(x') + m(g_0(y)) + h_2 w')f_{\epsilon_2}(v)dxdydx'dwdw'dv. \end{aligned} \tag{76}$$

From Assumption 1, 3 and 4, the integrand function on the right hand side of (76), as $n \to \infty$, converges to

$$K_2(w)K_2(w')f_1(x)f_2(y)f_1(x')f_{\epsilon_1}(v - m(x) + m(g_0(y)))f_{\epsilon_1}(v - m(x') + m(g_0(y)))f_{\epsilon_2}(v),$$

and is dominated by the integrable function

$$M_f^2 K_2(w)K_2(w')f_1(x)f_2(y)f_1(x')f_{\epsilon_2}(v).$$

Thus, by using DCT, we have

$$\lim_{n\to\infty} V_{21} = \int_{-\infty}^{\infty}\int_{0}^{\infty} f_2(y)f_{\epsilon_2}(v)\left[\int_0^{\infty} f_1(x)f_{\epsilon_1}(v - m(x) + m(g_0(y)))dx\right]^2 dvdy. \tag{77}$$

Now, from (66), the second term on the right hand side of (75) turns out to be

$$\frac{n_1 - 1}{n_1 n_2}[E(U_n)]^2 = O\left(\frac{1}{n_2}\right).$$

Thus, we have

$$\begin{aligned} n_2 V_2 \ \to \ & \int_{-\infty}^{\infty}\int_{0}^{\infty} f_2(y)f_{\epsilon_2}(v)\left[\int_0^{\infty} f_1(x)f_{\epsilon_1}(v - m(x) + m(g_0(y)))dx\right]^2 dvdy \\ & - \left[\int_{-\infty}^{\infty}\int_{0}^{\infty}\int_{0}^{\infty} f_1(x)f_2(y)f_{\epsilon_1}(v - m(x) + m(g(y)))f_{\epsilon_2}(v)dxdydv\right]^2, \end{aligned}$$

i.e.,

$$V_2 = O\left(\frac{1}{n_2}\right).$$

A similar argument, as for the term V_2, shows that

$$\begin{aligned} n_1 V_3 \ \rightarrow \ & \int_{-\infty}^{\infty}\int_0^{\infty} f_1(x) f_{\epsilon_1}(v)\left[\int_0^{\infty} f_2(y) f_{\epsilon_2}(v+m(x)-m(g_0(y)))dy\right]^2 dvdx \\ & -\left[\int_{-\infty}^{\infty}\int_0^{\infty}\int_0^{\infty} f_1(x)f_2(y)f_{\epsilon_1}(v-m(x)+m(g(y)))f_{\epsilon_2}(v)dxdydv\right]^2, \end{aligned}$$

i.e.,

$$V_3 = O\left(\frac{1}{n_1}\right).$$

The term $V_4 = 0$ from the model specification. This completes the proof. □

Lemma A.3 *Let the functional C be defined on $\mathbb{G}$ as*

$$C(g) = \frac{\int_{-\infty}^{\infty}\int_0^{\infty} f_1(g(y))f_2(y)f_{\epsilon 1}(v-m(g(y))+m(g_0(y)))f_{\epsilon 2}(v)dydv}{\int_0^{\infty} f_1(g(y))f_2(y)dy}.$$

Then, under Assumptions 1 and 5, C is uniformly continuous on $\mathbb{G}$.

Proof: Let

$$N(g) = \int_{-\infty}^{\infty}\int_0^{\infty} f_1(g(y))f_2(y)f_{\epsilon 1}(v-m(g(y))+m(g_0(y)))f_{\epsilon 2}(v)dydv, \quad (78)$$

$$D(g) = \int_0^{\infty} f_1(g(y))f_2(y)dy. \quad (79)$$

Then, $C(g) = \frac{N(g)}{D(g)}$.

Let $g \in \mathbb{G}$ and $\{g_k \in \mathbb{G}; k = 1, 2, \ldots\}$ be such that $\lim_{k\to\infty} \sup|g_k - g| = 0$. Now,

$$\lim_{k\to\infty} N(g_k) = \lim_{k\to\infty}\int_{-\infty}^{\infty}\int_0^{\infty} f_1(g_k(y))f_2(y)f_{\epsilon 1}(v-m(g_k(y))+m(g_0(y)))f_{\epsilon 2}(v)dydv. \quad (80)$$

Note that the integrand on the right hand side of (80) is bounded by the integrable function $M_f^2 f_2(y)f_{\epsilon 2}(v)$. Thus, applying DCT, we have

$$\lim_{k\to\infty} N(g_k) = \int_{-\infty}^{\infty}\int_0^{\infty}\left\{\lim_{k\to\infty} f_1(g_k(y))\right\} f_2(y)\left\{\lim_{k\to\infty} f_{\epsilon 1}(v-m(g_k(y))+m(g_0(y)))\right\} f_{\epsilon 2}(v)dydv. \quad (81)$$

Note that $g_k \to g$ as $k \to \infty$ pointwise. Now using the fact from the model specification that functions m is continuous and Assumption 1, we have

$$\begin{aligned} \lim_{k\to\infty} f_1(g_k(y)) &= f_1(g(y)), \\ \lim_{k\to\infty} f_{\epsilon 1}(v-m(g_k(y))+m(g_0(y))) &= f_{\epsilon 1}(v-m(g(y))+m(g_0(y))). \end{aligned}$$

Thus, from (81), we have

$$\lim_{k\to\infty} N(g_k) = \int_{-\infty}^{\infty}\int_{0}^{\infty} f_1(g(y))f_2(y)f_{\epsilon 1}(v - m(g(y)) + m(g_0(y)))f_{\epsilon 2}(v)dydv$$
$$= N(g).$$

This shows that the functional $N(\cdot)$ is continuous on $\mathbb{G}$.

A similar argument, as used for the continuity of $N(\cdot)$, shows that $D(\cdot)$, as defined in (79), is continuous. Further, note that, from Assumption 1 and (79), $D(g) > 0$ for any $g \in \mathbb{G}$. This establishes that C is continuous on $\mathbb{G}$. From Assumption 5, C is uniformly continuous on $\mathbb{G}$. This completes the proof. $\square$

Proof of Theorem 3.3: For any given $\epsilon > 0$, we have

$$\begin{aligned} &P\{|C_n(\hat{g}_n) - C(g_0)| > \epsilon\} \\ \leq\ & P\{|C_n(\hat{g}_n) - C(g_0)| > \epsilon, |C_n(g_0) - C(g_0)| \leq \epsilon\} + P\{|C_n(g_0) - C(g_0)| > \epsilon\}, \end{aligned} \tag{82}$$

where $\hat{g}_n$ is as in (3) which is given by

$$\hat{g}_n = \arg\max_{g\in\mathbb{G}} C_n(g).$$

From (82), we have

$$\begin{aligned} &P\{|C_n(\hat{g}_n) - C(g_0)| > \epsilon\} \\ &\quad\leq\ P\{|C_n(\hat{g}_n) - C(g_0)| > \epsilon, |C_n(g_0) - C(g_0)| \leq \epsilon, |C_n(\hat{g}_n) - C(\hat{g}_n)| \leq \epsilon\} \\ &\qquad +P\{|C_n(\hat{g}_n) - C(\hat{g}_n)| > \epsilon\} + P\{|C_n(g_0) - C(g_0)| > \epsilon\}. \end{aligned} \tag{83}$$

We will complete the proof by establishing that all the three terms on the right hand side of (83) are arbitrarily small.

We begin with the first term on the right hand side of (83). Note that, from (3), we have

$$C_n(g_0) \leq C_n(\hat{g}_n).$$

Therefore, from (3), we have

$$\text{if } |C_n(g_0) - C(g_0)| \leq \epsilon \quad \text{then} \quad C_n(\hat{g}_n) \geq C(g_0) - \epsilon. \tag{84}$$

We now turn to computing an upper bound for $C_n(\hat{g}_n)$ in terms of $C(g_0)$. Note that from (4) for all $g \in \mathbb{G}$,

$$C(g) \;=\; \frac{\int_0^\infty f_1(g(y))f_2(y)\left\{\int_{-\infty}^\infty f_{\epsilon 1}(v - m(g(y)) + m(g_0(y)))f_{\epsilon 2}(v)dv\right\}dy}{\int_0^\infty f_1(g(y))f_2(y)dy}. \tag{85}$$

The inner integral of the numerator of (85) can be viewed as the convolution of the densities f_{ϵ_1} and f_{ϵ_2}. Therefore, from Assumption 1, we have

$$\int_{-\infty}^{\infty} f_{\epsilon 1}(v - m(g(y)) + m(g_0(y))) f_{\epsilon 2}(v) dv \leq \int_{-\infty}^{\infty} f_{\epsilon 1}(v) f_{\epsilon 2}(v) dv. \tag{86}$$

Now from (85) and (86), we have

$$C(g) \quad \leq \quad \int_{-\infty}^{\infty} f_{\epsilon 1}(v) f_{\epsilon 2}(v) dv = C(g_0) \quad \text{for all} \quad g \in \mathbb{G}. \tag{87}$$

From (3) and (87) we have

$$C(\hat{g}_n) \leq C(g_0). \tag{88}$$

Now, from (88), we have

$$\text{if } |C_n(\hat{g}_n) - C(\hat{g}_n)| \leq \epsilon \quad \text{then} \quad C_n(\hat{g}_n) \leq C(g_0) + \epsilon. \tag{89}$$

From (84) and (89), we have

$$\text{if } |C_n(g_0) - C(g_0)| \leq \epsilon \quad \text{and} \quad |C_n(\hat{g}_n) - C(\hat{g}_n)| \leq \epsilon \quad \text{then} \quad |C_n(\hat{g}_n) - C(g_0)| \leq \epsilon. \tag{90}$$

Thus, from (90),

$$P\{|C_n(\hat{g}_n) - C(g_0)| > \epsilon, |C_n(g_0) - C(g_0)| \leq \epsilon, |C_n(\hat{g}_n) - C(\hat{g}_n)| \leq \epsilon\} = 0,$$

which is the first term on the right hand side of (83).

We now consider the second term on the right hand side of (83). Observe that,

$$|C_n(\hat{g}_n) - C(\hat{g}_n)| \leq \sup_{g \in \mathbb{G}} |C_n(g) - C(g)|. \tag{91}$$

From (91) and Theorem 3.2 we have,

$$C_n(\hat{g}_n) - C(\hat{g}_n) \xrightarrow{P} 0.$$

This ensures that the second term on the right hand side of (83)goes to zero as $n \to \infty$. Further, Theorem (3.1) ensures that the last term on the right hand side of (83) goes to zero too. This completes the proof. □

Proof of Theorem 3.4: If $\hat{g}_n \overset{P}{\nrightarrow} g_0$, then there exists an $\epsilon > 0$ and a $\delta > 0$ such that

$$P\{\sup |\hat{g}_n - g_0| \geq \epsilon\} > \delta \text{ infinitely often.} \tag{92}$$

Note that, $\mathcal{N}_\epsilon(g_0) = \{g : \sup |g - g_0| \geq \epsilon,\ g \in \mathbb{G}\}$ is a closed subset of $\mathbb{G}$. From Assumption 5 and Lemma A.3, there exists a $\tilde{g} \in \mathcal{N}_\epsilon(g_0)$ such that $\tilde{g} = \text{argmax}_{g \in \mathcal{N}_\epsilon(g_0)} C(g)$. From Assumption 6, when $\hat{g}_n \in \mathcal{N}_\epsilon(g_0)$, implies

$$|C(g_0) - C(\hat{g}_n)| = C(g_0) - C(\hat{g}_n) \geq C(g_0) - C(\tilde{g}) > 0. \tag{93}$$

Denote $\eta = C(g_0) - C(\tilde{g})$. By using the triangular inequality, we have

$$|C(g_0) - C_n(\hat{g}_n)| + |C_n(\hat{g}_n) - C(\hat{g}_n)| \quad \geq \quad |C(g_0) - C(\hat{g}_n)|. \tag{94}$$

From (93) and (94), $\hat{g}_n \in \mathcal{N}_\epsilon(g_0)$, implies

$$|C(g_0) - C_n(\hat{g}_n)| \quad \geq \quad \eta - |C_n(\hat{g}_n) - C(\hat{g}_n)|. \tag{95}$$

Now from (95),

$$\text{if } \hat{g}_n \in \mathcal{N}_\epsilon(g_0) \text{ then } \sup_{g \in \mathbb{G}} |C_n(g) - C(g)| < \frac{\eta}{2} \text{ then } |C_n(\hat{g}_n) - C(g_0)| > \frac{\eta}{2}. \tag{96}$$

Therefore, from (96), we have

$$\begin{aligned} P\{|C_n(\hat{g}_n) - C(g_0)| > \frac{\eta}{2}\} \quad &\geq \quad P\{\hat{g}_n \in \mathcal{N}_\epsilon(g_0) \text{ and } \sup_{g \in \mathbb{G}} |C_n(g) - C(g)| < \frac{\eta}{2}\}, \\ &\geq \quad P\{\sup |\hat{g}_n - g_0| \geq \epsilon\} + P\{\sup_{g \in \mathbb{G}} |C_n(g) - C(g)| < \frac{\eta}{2}\} - 1. \end{aligned} \tag{97}$$

From (92), the first term on the right hand side of (97) is greater δ infinitely often. From Theorem 3.2, the second term on the right hand side of (97) is greater than $1 - \frac{\delta}{2}$ for all but finitely many n. Therefore,

$$P\{|C_n(\hat{g}_n) - C(g_0)| > \eta/2\} \quad > \quad \frac{\delta}{2} \quad \text{infinitely often}. \tag{98}$$

This contradicts Theorem 3.3 and completes the proof. $\square$

References

[1] Jérémie Bigot. Landmark-based registration of 1D curves and functional analysis of variance with wavelets. Technical Report 0333, Interuniversity Attraction Pole Statistics Network, 2003.

[2] Jérémie Bigot. A scale-space approach with wavelets to singularity estimation. *ESAIM: Probability and Statistics*, 9:143–164, Apr. 2005.

[3] Patrick Billingsley. *Probability and Measure*. John Wiley, 1985.

[4] Fred L. Bookstein. *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press, Cambridge, England, 1991.

[5] Lyndia C. Brumback and Mary J. Lindstrom. Self modeling with flexible, random time transformations. *Biometrics*, 60(2):461–470, Jun. 2004.

[6] Céline Bugli, Philippe Lambert, and Jérémie Bigot. Curve registration using fractional polynomials with application to electroencephalograms analysis. Discussion Paper 0513, IAP Statistics Network, 2005.

[7] C. K. Chu, I. K. Glad, F. Godtliebsen, and J. S. Marron. Edge-preserving smoothers for image processing. *Journal of the American Statistical Association*, 93(442):526–541, Jun. 1998.

[8] Theo Gasser and Alois Kneip. Searching for structure in curve samples. *Journal of the American Statistical Association*, 90(432):1179–1188, Dec. 1995.

[9] Daniel Gervini and Theo Gasser. Self-modelling warping functions. *Journal of the Royal Statistical Society, Series B*, 66(4):959–971, 2004.

[10] Peter Hall and D. M. Titterington. Edge-preserving and peak-preserving smoothing. *Technometrics*, 34(4):429–440, Nov. 1992.

[11] Gareth M. James. Curve alignment by moments. *The Annals of Applied Statistics*, 1(2):480–501, Dec. 2007.

[12] A. Kneip, X. Li, K. B. MacGibbon, and J. O. Ramsay. Curve registration by local regression. *The Canadian Journal of Statistics*, 28(1):19–29, Mar. 2000.

[13] Alois Kneip and Theo Gasser. Convergence and consistency results for self-modeling nonlinear regression. *The Annals of Statistics*, 16(1):82–112, Mar. 1988.

[14] Alois Kneip and Theo Gasser. Statistical tools to analyze data representing a sample of curves. *The Annals of Statistics*, 20(3):1266–1305, Sep. 1992.

[15] W. H. Lawton, E. A. Sylvestre, and M. S. Maggio. Self modeling nonlinear regression. *Technometrics*, 14(3):513–532, Aug. 1972.

[16] Xueli Liu and Hans-Georg Müller. Functional convex averaging and synchronization for time-warped random curves. *Journal of the American Statistical Association*, 99(467):687–699, Sep. 2004.

[17] John Alan McDonald and Art B. Owen. Smoothing with split linear fits. *Technometrics*, 28(3):195–208, Aug. 1986.

[18] Manfred Mudelsee. *Climate Time Series Analysis Classical Statistical and Bootstrap Methods*, volume 42. Springer, New York, first edition edition, 2010.

[19] Hans-Georg Müller. Change-points in nonparametric regression analysis. *Ann. Stat.*, 20(2):737–761, Jun. 1992.

[20] Peihua Qiu, Chooichiro Asano, and Xianping Li. Estimation of jump regression function. *Bulletin of Informatics and Cybernetics*, 24(3 4):197–212, 1991.

[21] J. O. Ramsay. When the data are functions. *Psychometrika*, 47(4):379–396, Dec. 1982.

[22] J. O. Ramsay. Estimating smooth monotone functions. *Journal of the Royal Statistical Society, Series B*, 60(2):365–375, 1998.

[23] J. O. Ramsay, Giles Hooker, and Spencer Graves. *Functional Data Analysis with R and MATLAB*. Springer, 2009.

[24] J. O. Ramsay and Xiaochun Li. Curve registration. *Journal of the Royal Statistical Society, Series B*, 60(2):351–363, 1998.

[25] Jim Ramsay and Bernard W. Silverman. *Applied Functional Data Analysis: Methods and Case Studies*. Springer, New York, first edition edition, Jun. 2002.

[26] Jim Ramsay and Bernard W. Silverman. *Functional Data Analysis*. Springer, New York, second edition edition, Jun. 2005.

[27] Birgitte B. Rønn. Nonparametric maximum likelihood estimation for shifted curves. *Journal of the Royal Statistical Society, Series B*, 63(2):243–259, 2001.

[28] Hiroaki Sakoe and Seibi Chiba. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-26(1):43–49, Feb. 1978.

[29] B. W. Silverman. Incorporating parametric effects into functional principal components analysis. *Journal of the Royal Statistical Society, Series B*, 57(4):673–689, 1995.

[30] Donatello Telesca and Lurdes Y. T. Inoue. Bayesian hierarchical curve registration. *Journal of the American Statistical Association*, 103(481):328–339, Mar. 2008.

[31] Kongming Wang and Theo Gasser. Alignment of curves by dynamic time warping. *The Annals of Statistics*, 25(3):1251–1276, Jun. 1997.

[32] Kongming Wang and Theo Gasser. Synchronizing sample curves nonparametrically. *The Annals of Statistics*, 27(2):439–460, Apr. 1999.

[33] J. S. Wu and C. K. Chu. Nonparametric function estimation and bandwidth selection for discontinuous regression functions. *Statistical Sinica*, 3:557–576, 1993.