Computational Statistics and Data Analysis ••• (••••) ••••••

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

www.elsevier.com/locate/csda

Specification testing for ordinary differential equation models with fixed design and applications to COVID-19 epidemic models

Ran Liu [a,c], Lixing Zhu [b,c,*]

[a] School of Mathematics and Statistics, Beijing Jiaotong University, Beijing, China
[b] Advanced Institute of Natural Sciences, Beijing Normal University at Zhuhai, Zhuhai, Guangdong, China
[c] Department of Mathematics, Hong Kong Baptist University, Hong Kong, China

ARTICLE INFO

Article history:
Received 10 January 2022
Received in revised form 5 September 2022
Accepted 5 September 2022
Available online xxxx

Keywords:
Local smoothing test
Model specification
Ordinary differential equations
SEIR model
U-statistics

ABSTRACT

Checking the models about the ongoing Coronavirus Disease 2019 (COVID-19) pandemic is an important issue. Some famous ordinary differential equation (ODE) models, such as the SIR and SEIR models have been used to describe and predict the epidemic trend. Still, in many cases, only part of the equations can be observed. A test is suggested to check possibly partially observed ODE models with a fixed design sampling scheme. The asymptotic properties of the test under the null, global and local alternative hypotheses are presented. Two new propositions about U-statistics with varying kernels based on independent but non-identical data are derived as essential tools. Some simulation studies are conducted to examine the performances of the test. Based on the available public data, it is found that the SEIR model, for modeling the data of COVID-19 infective cases in certain periods in Japan and Algeria, respectively, maybe not be appropriate by applying the proposed test.

1. Introduction

Since December 2019, a new infectious disease called the Coronavirus Disease 2019 (COVID-19) has spread worldwide, resulting in an ongoing pandemic. To help understand the characteristic of this epidemic and to act suitably to reduce its risk, numerous studies have been done by using several approaches to estimate the sizes of the COVID-19 infected populations in different countries and regions and further predict their trends (e.g., Lin et al. (2020); Altieri et al. (2021); Yang et al. (2020); Tian et al. (2020)). The modeling methods of ordinary differential equations (ODEs) are popularly used by using the mechanisms behind the data, including the Susceptible-Infectious-Removed (SIR) model, the Susceptible-Exposed-Infectious-Removed (SEIR) model, and several other variants. Note that we may only have part of the model equations observed from the public data. This is often the case in practice. Thus, in this paper, we particularly pay attention to such models.

* Corresponding author at: Advanced Institute of Natural Sciences, Beijing Normal University at Zhuhai, Zhuhai, Guangdong, China.
E-mail address: lzhu@bnu.edu.cn (L. Zhu).

https://doi.org/10.1016/j.csda.2022.107616
0167-9473/

Consider a general d-dimensional ODE model as

$$X'(t) = \begin{bmatrix} \frac{dX_1(t)}{dt} \\ \vdots \\ \frac{dX_d(t)}{dt} \end{bmatrix} = \begin{bmatrix} f_1(t, X(t); \theta) \\ \vdots \\ f_d(t, X(t); \theta) \end{bmatrix} = f(t, X(t); \theta), \qquad t \in [t_0, T], \tag{1}$$

with an initial condition $X(t_0) = x_0$. Here $\mathcal{F} = \{f(\cdot; \theta) : \theta \in \Theta \subset R^q\}$ is a given parametric family of functions. The ODE system is commonly observed with noise as

$$Y_{ij} = X_j(t_i) + \epsilon_{ij}, \quad i = 1, \ldots, n, \; j \in \mathcal{O} = \{d_1, \ldots, d_o\},$$

where $d_1, \ldots, d_o$ are chosen from $1, 2, \ldots, d$ and $o \le d$. The measurement error ϵ_i satisfying $E(\epsilon_i) = 0$ has a nonsingular variance-covariance matrix Σ_{ϵ_i}, and is independent with ϵ_j for $j \ne i$. If $o = d$, the ODE system is fully observed, otherwise is a partially observed ODE system, where only a small subset of components can be measured (Dattner (2015)).

Using ODE models to describe dynamic processes can estimate parameters with particular relevance and conduct further statistical analysis. However, an important issue is whether the assumed ODE model rightly describes the measurement data. For instance, we may want to know whether an SEIR model is good enough to model real COVID-19 infective population data for a period of time. If not, we may have to be careful to use the results from the statistical analysis. Thus, a model checking for the assumed ODE model should be accompanied. Given a certain parametric family $\mathcal{F}$, the hypotheses we consider are formally stated as

$$H_0 : X'(t) = f(t, X(t); \theta_0) \in \mathcal{F} \quad \text{versus} \quad H_1 : X'(t) \notin \mathcal{F}, \tag{2}$$

where θ_0 is an unknown parameter vector.

The hypothesis testing problem (2) is a model specification testing problem. Some relevant tests, usually under random design, concerning regression model specification are built by using nonparametric estimations, and thus are called the local smoothing tests. Some literature, such as Fan and Li (2000), points out that local smoothing methods are sensitive to alternative models which are oscillating/highly frequent than global smoothing methods in general. Examples contain Härdle and Mammen (1993), Zheng (1996), Koul and Ni (2004) and Lavergne and Patilea (2012). Some others are based on residual-marked empirical processes and the average over an index set. Then they are called the global smoothing tests. Examples include Stute (1997), Stute et al. (1998a), Stute et al. (1998b), Zhu (2003) and Khmaladze and Koul (2004). These tests can better detect smooth alternative models. However, testing for ODE models specification still receives less attention. For testing the hypotheses in (2), Hooker (2009) proposed a goodness-of-fit test based on estimated forcing functions. It has an exact null distribution under the normality assumption on independent components and the homoscedasticity structure on error terms. It may have difficulty controlling the size when the above assumptions do not hold. As known from the qualitative theory of differential equations, an ODE model often has a periodic non-constant solution, which is called a limit cycle. Thus, the original function $X(t)$ often oscillates. See Hirsch et al. (2013) as a good reference. This phenomenon suggests that local smoothing tests for model checking should be more appropriate than global smoothing tests. Liu et al. (2021) constructed a local smoothing test under more relaxed assumptions on error terms and gave its asymptotic properties with a random design. Their trajectory matching-based test, written as TM_n, takes advantage of the solution trajectory of the ODE system. We note that the good theoretical results of TM_n rely on a random design sampling scheme. The sample (t_i, Y_i)'s are independent and identically distributed (i.i.d.), thus the limiting properties can be derived by using the existing theory of U-statistics. But for ODE systems such as the epidemic model, the variable t usually represents time, and the response is measured at certain time points in practice. Therefore, it is suitable to treat $t_1, \cdots, t_n$ under fixed design rather than i.i.d. random design. The test TM_n that is based on this non-i.i.d. sequence of (t_i, Y_i)'s needs to be investigated to see its usability.

In this paper, we give a modification of TM_n with a fixed design and study the respective properties under the null, local and global alternatives. Without notational confusion, we still write it as TM_n. We then check the SEIR model for two real COVID-19 data sets. Since the sample (t_i, Y_i) are still independent in the case of fixed design, we also consider U-statistics with varying kernels based on independent but not identically distributed random variables by giving two new propositions concerning degenerate and non-degenerate U-statistics.

The paper is organized as follows. In Section 2, we will present some general asymptotic properties of U-statistics. Section 3 will contain the construction of the test TM_n and its asymptotic properties. Section 4 will include some simulation results. In Section 5, we will use the proposed test to check whether a classic SEIR model is tenable for describing the tendency of COVID-19 infective cases from 15 January to 29 February, 2020 in Japan and from 25 February to 18 April, 2020 in Algeria respectively. A summary and a brief discussion will be given in Section 6. The technical proofs will be relegated to Appendix.

2. U-statistics with varying kernels

The theory of U-statistics with i.i.d. random variables has been well developed (see Hoeffding (1948); Lee (1990)). For U-statistics with varying kernels, Powell et al. (1989) established the asymptotic equivalence of a U-statistic and its projection when the U-statistic is non-degenerate. In degenerate case, Hall (1984) proved a central limit theorem, showing that a second order U-statistic has a limiting normal distribution under some regularity conditions. Fan and Li (1996) and Fan and Li (1999) extended Hall's result to more complicated cases. However, the theory of U-statistics with varying kernels based on independent but non-identically distributed data remains unknown. We then study the asymptotic properties of both degenerate and non-degenerate U-statistics in the non-i.i.d. scenarios.

Consider a second order U-statistic U_n:

$$U_n = \binom{n}{2}^{-1} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} h_n\left(z_i, z_j\right),$$

with the kernel $h_n\left(z_i, z_j\right)$ that varies with the sample size n. Recall that $\{z_i, i = 1, \ldots, n\}$ is an independent but non-identical sample. Assuming the expectation of $h_n\left(z_i, z_j\right)$ exists for every z_i and z_j, we define

$$\begin{aligned} r_{nj}(z_i) =& E\left\{h_n(z_i, z_j) \mid z_i\right\}, \quad \theta_{nij} = E\left\{r_{nj}(z_i)\right\} = E\left\{h_n\left(z_i, z_j\right)\right\}, \\ r_n(z_i) =& \binom{n-1}{1}^{-1} \sum_{j \neq i} E\left\{h_n(z_i, z_j) \mid z_i\right\}, \\ \theta_{ni} =& E\left\{r_n(z_i)\right\}, \quad \theta_n = \binom{n}{2}^{-1} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \theta_{nij}, \\ \hat{U}_n =& \theta_n + \frac{2}{n} \sum_{i=1}^{n} \left\{r_n(z_i) - \theta_{ni}\right\}. \end{aligned}$$

$\hat{U}_n$ is the first order projection of U_n. If $var(\hat{U}_n) = 0$, the U-statistic is degenerate, otherwise is non-degenerate. The following proposition is for the non-degenerate case, which is exactly a generalization of Lemma 3.1 of Powell et al. (1989).

Proposition 2.1. *Suppose U_n is non-degenerate. If $E\left\|h_n\left(z_i, z_j\right)\right\|^2 = o(n)$ for every z_i and z_j, then $\sqrt{n}\left(U_n - \hat{U}_n\right) = o_P(1)$.*

Since U_n and $\hat{U}_n$ are asymptotically equivalent, we can use this proposition to establish the asymptotic normality of a non-degenerate U-statistic by using the central limit theorem.

For a degenerate U_n, we assume U_n is centered without loss of generality. Denote $G_{ni}(z_j, z_k) = E\{h(z_i, z_j)h(z_i, z_k) \mid z_j, z_k\}$. We state the following result that is a generalization of Theorem 1 of Hall (1984) to the non-i.i.d. case.

Proposition 2.2. *For every z_i, z_j and n, suppose $r_{nj}(z_i) = E\left\{h_n(z_i, z_j) \mid z_i\right\} = 0$ almost surely and $E\left\{h_n^2(z_i, z_j)\right\} < \infty$, if as $n \to \infty$*

$$\frac{\sum_{2 \leqslant i_1 \leqslant i_2 \leqslant n} \sum_{1 \leqslant j_1 < k_1 \leqslant i_1 - 1} E\left\{G_{ni_1}\left(z_{j_1}, z_{k_1}\right) G_{ni_2}\left(z_{j_1}, z_{k_1}\right)\right\} + n \sum_{i=2}^{n} \sum_{j=1}^{i-1} E\left\{h_n^4\left(z_i, z_j\right)\right\}}{\left[\sum_{i=2}^{n} \sum_{j=1}^{i-1} E\left\{h_n^2\left(z_i, z_j\right)\right\}\right]^2} \to 0,$$

then

$$n^2 \cdot U_n / 2 \left[\sum_{i=2}^{n} \sum_{j=1}^{i-1} E\left\{h_n^2\left(z_i, z_j\right)\right\}\right]^{1/2}$$

is asymptotically standard normal.

3. The test with fixed design

3.1. Test statistic construction

We first define the fixed design sampling scheme.

Condition 1. *The sample points $t_1, \cdots, t_n$ are fixed. There exists a distribution $P(t)$ with the corresponding density function $p(t)$ such that*

$$\sup_t |P_n(t) - P(t)| = O(n^{-1}),$$

where $P_n(t)$ is the empirical distribution of $(t_1, \cdots, t_n)$. The density function $p(t)$ is bounded away from 0 and ∞. The first derivative of $p(t)$ is bounded and continuous.

This condition is used to establish the asymptotic equivalence between the empirical distribution of $t_1, \cdots, t_n$ and a given distribution $P(t)$, which is commonly used in fixed design sampling (Xue et al. (2010)). Hereafter we denote $E_t\{g(t_i)\} = \int g(t)p(t)dt$.

Denote $F(t;\theta)$ as the unique solution trajectory of the ODE system $X'(t) = f(t, X(t);\theta)$. By using the trajectory $F(t;\theta)$, we can convert the hypothesis testing problem (2) to the problem of checking the vector function $X(t) = F(t;\theta_0)$ for some $\theta_0 \in \Theta \subset R^q$.

The following construction is a modification of the test proposed by Liu et al. (2021) that is for the problem with random design. Let $\eta_{ik} = Y_{ik} - F_k(t_i;\theta^*)$ with $\theta^* = \arg\min_\theta E_t[E\{\sum_{k\in\mathcal{O}} |Y_{ik} - F_k(t_i;\theta)|^2\}]$ be the residual of the k-th component. Under H_0, $\eta_{ik} = \epsilon_{ik}$ and $E(\eta_{ik}) = 0$ leads to $E_t[E\{\eta_{ik}E(\eta_{ik})p(t_i)\}] = 0$, while under H_1, $E(\eta_{ik}) = X_k(t_i) - F_k(t_i;\theta) \neq 0$, and $E_t[E\{\eta_{ik}E(\eta_{ik})p(t_i)\}] = E_t[\{E(\eta_{ik})\}^2 p(t_i)] > 0$. At the sample level, we can use $e_{ik} = Y_{ik} - F_k(t_i;\hat{\theta})$ as an estimator of η_{ik} with the nonlinear least-squares estimator $\hat{\theta} = \arg\min_\theta \sum_{i=1}^n \sum_{k\in\mathcal{O}} \{Y_{ik} - F_k(t_i;\theta)\}^2$. Then, for each observed component k, we construct a test statistic that is the sample analogue of $E_t[E\{\eta_{ik}E(\eta_{ik})p(t_i)\}]$ as follows:

$$V_{nk} = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} \frac{1}{h} K\left(\frac{t_i - t_j}{h}\right) e_{ik} e_{jk},$$

where h is the bandwidth parameter and $K(\cdot)$ is the kernel function. V_{nk} is in spirit similar to the test proposed in Zheng (1996). We compose all o tests to form a vector $V_n = \left(V_{nd_1}, \ldots, V_{nd_o}\right)^\top$. The final test statistic is constructed by aggregating all the components of V_n:

$$TM_n = n^2 h V_n^\top (\widehat{\Sigma})^{-1} V_n,$$

with a normalized symmetric matrix $\widehat{\Sigma}$:

$$\widehat{\Sigma} = \frac{2}{n(n-1)} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} \frac{1}{h} K^2\left(\frac{t_i - t_j}{h}\right) (e_i \odot e_j)(e_i \odot e_j)^\top,$$

where $e_i = \left(e_{id_1}, \ldots, e_{id_o}\right)^\top$ and $\odot$ means the element-wise product of two vectors.

3.2. Asymptotic properties

To investigate the asymptotic properties, we first give certain regularity conditions.

1. *$\epsilon_i, i = 1, \ldots, n$ satisfy $E(\epsilon_i) = 0$. They have nonsingular variance-covariance matrix Σ_{ϵ_i}. ϵ_i is independent with ϵ_j for every $j \neq i$.*
2. *All partial derivatives of $X'(t)$ up to 2 with respect to X and t are existent and continuous.*
3. *$K(\cdot)$ is a nonnegative, bounded, continuous and symmetric function with $\int K(u)du = 1$.*
4. *For all $1 \le k \le d$, $E(y_{ik}^4)$ is continuously differentiable and bounded by a measurable function $b(t)$ with $E(b^2(t_i)) < \infty$.*
5. *The parameter space Θ is a convex compact subset of R^q.*
6. *$F_k(t;\theta)$ is a Borel measurable real function on R^p for each θ and is twice continuously differentiable with respect to θ for each t.*
7. *For all $k = d_1, \cdots, d_o$, we have*

$$E_t\left\{\sup_{\theta\in\Theta} F_k^2(t;\theta)\right\} < \infty,$$
$$E_t\left\{\sup_{\theta\in\Theta}\left\| \frac{\partial F_k(t;\theta)}{\partial\theta} \frac{\partial F_k(t;\theta)}{\partial\theta^\top} \right\|\right\} < \infty,$$
$$E_t\left(E\left[sup_{\theta\in\Theta}\left\|\{Y_k - F_k(t;\theta)\}^2 \frac{\partial F_k(t;\theta)}{\partial\theta} \frac{\partial F_k(t;\theta)}{\partial\theta^\top}\right\|\right]\right) < \infty,$$
$$E_t\left(E\left[\sup_{\theta\in\Theta}\left\|\{Y_k - F_k(t;\theta)\}^2 \frac{\partial^2 F_k(t;\theta)}{\partial\theta\partial\theta^\top}\right\|\right]\right) < \infty.$$

8. *$E_t[E\{\sum_{k\in\mathcal{O}}^{d_o} |Y_{ik} - F_k(t_i;\theta)|^2\}]$ takes a unique minimum at $\theta^* \in \Theta$.*
9. *The matrix $E_t\left\{\sum_{k\in\mathcal{O}}^{d_o} \frac{\partial F_k(t_i;\theta_0)}{\partial\theta} \frac{\partial F_k(t_i;\theta_0)}{\partial\theta^\top}\right\}$ is nonsingular.*

Condition (1) is basic. Condition (2) assumes that the ODE system has a unique solution trajectory. The commonly used requirement for the kernel function is put in Condition (3). Conditions (4)-(9) are standard for the nonlinear least squares estimation, which is similar to the conditions given in White (1981), and Xue et al. (2010).

It can be shown that V_n and $\widehat{\Sigma}$ are asymptotic U-statistics with non-i.i.d. observations. Thus we use Propositions 2.1 and 2.2 in Section 2 to derive their asymptotic properties. The limiting null distribution of TM_n and its consistency under the global alternatives are stated in the following theorem.

Theorem 1. *Given Conditions (1)-(9), if $h \to 0$ and $nh \to \infty$, then*

1. *Under the null hypothesis H_0,*

$$TM_n \to \chi_o^2, \text{ in distribution},$$

where $TM_n = n^2 h V_n^\top \widehat{\Sigma}^{-1} V_n$, and χ_o^2 is the chi-square distribution with o degrees of freedom.

2. *Define an $o \times o$ matrix Σ' with the $(k1, k2)$ element, and*

$$\Sigma'_{k1k2} = 2\int \left[\sigma_{d_{k1}d_{k2}}(t) + \left\{X_{d_{k1}}(t) - F_{d_{k1}}\left(t, \theta^*\right)\right\}\left\{X_{d_{k2}}(t) - F_{d_{k2}}\left(t, \theta^*\right)\right\}\right]^2 p^2(t)dt \times \int K^2(u)du,$$

$k1, k2 = 1, \cdots, o$. Then under the global alternative H_1,

$$TM_n/(n^2 h) \to V'^\top \Sigma'^{-1} V', \text{ in probability},$$

where V' is an o-dimensional vector whose k-th component is equal to $V'_k = E\left[\left\{X_{d_k}(t) - F_{d_k}(t;\theta^*)\right\}^2 p(t)\right]$.

We also study the power performance of TM_n under the following sequence of local alternative hypotheses:

$$H_{1n}: X(t) = F(t;\theta_0) + \delta_n L(t),$$

with the bounded multiple response function $L(t)$ and the departure parameter $\delta_n \to 0$. Then we give the following theorem which states the asymptotic power performance of TM_n under H_{1n}.

Theorem 2. *Given Conditions (1)-(9), if $h \to 0$ and $nh \to \infty$, then under H_{1n} with $n^{1/2}h^{1/4}\delta_n \to \infty$,*

$$TM_n/(n^2 h \delta_n^4) \to \mu^\top \Sigma^{-1} \mu, \text{ in probability},$$

where μ is an o-dimensional vector with the i-th component

$$\mu_i = E\left(\left[\left\{L_{d_i}(t) - \frac{\partial F_{d_i}(t;\theta_0)}{\partial \theta^\top} H_{\dot{F}}^{-1} E\left\{\sum_{k\in\mathcal{O}}^{d_o} L_k(t) \frac{\partial F_k(t;\theta_0)}{\partial \theta}\right\}\right\}\right]^2 p(t)\right).$$

Particularly, if $\delta_n = n^{-1/2}h^{-1/4}$,

$$TM_n \to \chi_o^2(\lambda), \text{ in distribution},$$

where $\chi_o^2(\lambda)$ is the noncentral chi-squared distribution where the noncentrality parameter $\lambda = \mu^\top \Sigma^{-1}\mu$ with $\Sigma_{k1k2} = 2\int K^2(u)du \times \int \left\{\sigma_{d_{k1}d_{k2}}(t)\right\}^2 p^2(t)dt$.

The results show that the test has a typical rate $n^{-1/2}h^{-1/4}$ existing local smoothing tests can achieve to detect local alternatives converging to the null model.

4. Simulations

We conduct five simulation examples to evidence the performance of TM_n in finite sample scenarios. In Example 1, the null ODE system is linear. Examples 2 - 4 respectively use three nonlinear ODE models as the null models, which have been often used in economic, neuroscience and ecology. Example 5 reconsiders the above four models but with partially observed ODE systems. In this example, we assume that we only have the data of the second component. Furthermore, we consider both independent and highly correlated components of the error terms for each example.

For the ODE models used in the examples, the variable t can be designed time point taking values $t_1 = 1/n, t_2 = 2/n, \cdots, t_n = 1$. With this fixed design, we apply the test TM_n for each cases. As a comparison, the test proposed by Hooker (2009) is computed as the competitor.

Given the ODE models, the trajectory $F(t;\theta_0)$ is obtained by using the 4-stage Runge-Kutta algorithm such that we have the respective models $Y(t) = X(t) + \epsilon(t)$. The error terms with independent components ϵ_i, $i = 1, \ldots, n$ independently follow

Table 1
Empirical sizes and powers in Example 1.

Model	α	0	0.5	0	0.5	1	0	1
	β	0	0	0.5	0.5	0	1	1
H_{11} with ϵ	TM_n	0.043	1.000	0.819	1.000	1.000	1.000	1.000
	T^H	0.054	1.000	1.000	1.000	1.000	1.000	1.000
H_{12} with ϵ	TM_n	0.053	1.000	0.990	1.000	1.000	1.000	1.000
	T^H	0.059	1.000	1.000	1.000	1.000	1.000	1.000
H_{13} with ϵ	TM_n	0.063	1.000	1.000	1.000	1.000	1.000	1.000
	T^H	0.056	1.000	1.000	1.000	1.000	1.000	1.000
H_{11} with ϵ^*	TM_n	0.051	1.000	0.935	1.000	1.000	1.000	1.000
	T^H	0.071	1.000	1.000	1.000	1.000	1.000	1.000
H_{12} with ϵ^*	TM_n	0.051	1.000	0.998	1.000	1.000	1.000	1.000
	T^H	0.063	1.000	1.000	1.000	1.000	1.000	1.000
H_{13} with ϵ^*	TM_n	0.069	1.000	1.000	1.000	1.000	1.000	1.000
	T^H	0.072	1.000	1.000	1.000	1.000	1.000	1.000

the normal distribution $N(0, \sigma_\epsilon^2 I_2)$, where I_2 is the identity matrix of size 2. Then the error terms with highly correlated components ϵ_i^* are made by $\epsilon_i^* = A\epsilon_i$, where $A = \begin{bmatrix} 0.9 & 0.5 \\ 0.5 & 0.9 \end{bmatrix}$. The initial values of ODE models are assumed to be known. We conduct 1000 experiments for each simulation case. The sample size is 300, and the significance level is 0.05.

To construct the test, we use the Matlab algorithm in OPTI Toolbox (Currie and Wilson (2012)) to get the nonlinear least-squares estimator $\hat{\theta}$. The trajectory $F(t; \hat{\theta})$ is also obtained by utilizing the 4-stage Runge-Kutta algorithm. We choose the Epanechnikov kernel $K(u) = 0.75 \times (1 - u^2)$ in the test construction. As for the choice of bandwidth h, we found a small h is helpful to ensure the asymptotic properties theoretically and makes the test has better performances in the simulations empirically. Thus we recommended a rate of $n^{-2/5}$, which is faster than the optimal rate of $n^{-1/5}$ for the kernel estimation. By the rule of thumb, the coefficient can be chosen as $\nu \times (T - t_0)$, where $(T - t_0)$ is the length of the whole time interval of the ODE model $[t_0, T]$ and $\nu \in [0.05, 0.20]$. In the following examples, the bandwidth is equal to $h = 0.05 \times n^{-2/5}$. To compute Hooker (2009)'s test T^H, the demonstration code given in the supplementary materials of that paper was revised to make it feasible to handle the cases with unknown parameters.

Example 1. Data sets are generated from the following ODE models:

$$H_{11}: \quad X'(t) = \begin{bmatrix} \frac{dX_1}{dt} \\ \frac{dX_2}{dt} \end{bmatrix} = \tau \begin{bmatrix} aX_1 + 0.4\alpha cos(aX_1) \\ aX_1 + bX_2 + 0.08\beta cos(aX_1 + bX_2) \end{bmatrix},$$

$$H_{12}: \quad X'(t) = \begin{bmatrix} \frac{dX_1}{dt} \\ \frac{dX_2}{dt} \end{bmatrix} = \tau \begin{bmatrix} aX_1 + 0.1\alpha (aX_1)^3 \\ aX_1 + bX_2 + 0.0002\beta (aX_1 + bX_2)^3 \end{bmatrix},$$

$$H_{13}: \quad X'(t) = \begin{bmatrix} \frac{dX_1}{dt} \\ \frac{dX_2}{dt} \end{bmatrix} = \tau \begin{bmatrix} aX_1 + 2\alpha exp(aX_1) \\ aX_1 + bX_2 + 0.5\beta exp(aX_1 + bX_2) \end{bmatrix}.$$

Three cases are considered in this example. In each case, the linear null ODE model is added with one kind of disturbance term to form alternative ODE models. From the above ODE models, it can be seen that $\alpha = 0$ and $\beta = 0$ correspond to the null hypothesis, otherwise to the alternative hypotheses. To transform the arbitrary length of sample time interval to 1, we induce a timescale parameter τ. The true parameter is set to be $(a, b) = (-0.06, -0.24)$ and the initial values are $(X_1(0), X_2(0)) = (5, 5)$. We choose $\tau = 10$ and $\sigma_\epsilon = 0.05$. The empirical sizes and powers are reported in Table 1, where ϵ and ϵ^* denote error terms with independent and highly correlated components, respectively. The rows with TM_n and T^H show the values of our test and Hooker's test with a fixed design.

With the two kinds of error terms, the results show that the trajectory matching-based test TM_n maintains the significance level when both $\alpha = 0$ and $\beta = 0$. It also has very good powers with all of the alternative models. The performances of T^H are also good with the independent components of error terms. T^H has good powers with dependent components of error terms, yet its empirical sizes are relatively larger than the significance level consistently under the three null models.

Example 2. Data sets are generated from the following ODE model:

$$H_2: X'(t) = \begin{bmatrix} \frac{dX_1}{dt} \\ \frac{dX_2}{dt} \end{bmatrix} = \tau \begin{bmatrix} aX_1 + 0.04\alpha cos(X_1) \\ bX_1^{2/3} X_2^{1/3} + 0.04\beta cos(X_2) \end{bmatrix}.$$

Table 2
Empirical sizes and powers in Example 2.

Model	α	0	0.5	0	0.5	1	0	1
	β	0	0	0.5	0.5	0	1	1
H_2 with ϵ	TM_n	0.043	0.491	0.776	0.946	1.000	1.000	1.000
	T^H	0.051	1.000	1.000	1.000	1.000	1.000	1.000
H_2 with ϵ^*	TM_n	0.042	0.609	0.898	0.838	1.000	1.000	1.000
	T^H	0.100	1.000	1.000	1.000	1.000	1.000	1.000

Table 3
Empirical sizes and powers in Example 3.

Model	α	0	0.5	0	0.5	1	0	1
	β	0	0	0.5	0.5	0	1	1
H_3 with ϵ	TM_n	0.049	0.844	0.301	1.000	1.000	0.999	1.000
	T^H	0.052	1.000	1.000	1.000	1.000	1.000	1.000
H_3 with ϵ^*	TM_n	0.047	0.645	0.263	0.995	1.000	0.988	1.000
	T^H	0.108	1.000	1.000	1.000	1.000	1.000	1.000

Under the null hypothesis, the model is called the Solow growth model, which is popular to describe long-run economic growth. In this model, X_1 and X_2 represent the labor and capital respectively. We choose the true parameters $(a, b) = (0.1, 0.3)$, $(X_1(0), X_2(0)) = (1, 3)$, $\tau = 10$, $\sigma_\epsilon = 0.05$. We report the empirical sizes and powers in Table 2.

For this nonlinear ODE model, it can be seen that TM_n also has good powers and sizes with a fixed design. T^H has superior performances when the components of the error are independent. It shows larger powers than TM_n. However, in the dependent error components case, it cannot maintain the significance level. This is not surprising because T^H requires the condition of independent components on the derivation of the null distribution.

Example 3. The data sets are generated from the following ODE model:

$$H_3 : X'(t) = \begin{bmatrix} \frac{dX_1}{dt} \\ \frac{dX_2}{dt} \end{bmatrix} = \tau \begin{bmatrix} a(X_1 + X_2 - \frac{X_1^3}{3}) + 0.2\alpha X_1 X_2 \\ -\frac{X_1 + bX_2 - c}{a} + 0.04\beta X_1 X_2 \end{bmatrix}.$$

The null model is the famous FitzHugh-Nagumo ODE model (FitzHugh (1961); Nagumo et al. (1962)) in neuroscience, which describes the behavior of spike potentials in the giant axon of squid neurons. Following Ding and Wu (2014), the true parameter is $(a, b, c) = (3, 0.2, 0.34)$, $\tau = 10$, $\sigma_\epsilon = 0.05$, and the initial values are $(X_1(0), X_2(0)) = (1, -1)$. The empirical sizes and powers are reported in Table 3.

TM_n still works very well in all cases. Similar to Example 2, T^H performs well in the independent error components case, but fails to work in the dependent error components case.

Example 4. Data sets are generated from the following ODE models:

$$H_3 : X'(t) = \begin{bmatrix} \frac{dX_1}{dt} \\ \frac{dX_2}{dt} \end{bmatrix} = \tau \begin{bmatrix} aX_1 + bX_1X_2 + 0.016\alpha X_2 \\ cX_2 + dX_1X_2 + 0.02\beta X_1 \end{bmatrix}.$$

The null ODE is called the Lotka-Volterra model in ecology, which is designed for modeling the evolution of prey-predator populations (Lotka (1910); Volterra (1928); Goel et al. (1971)). Following Brunel (2008), we set the true parameters $(a, b, c, d) = (1, -1.5, -1.5, 2)$, $(X_1(0), X_2(0)) = (1, 2)$, $\tau = 10$ and $\sigma_\epsilon = 0.05$. The empirical sizes and powers are summarized in Table 4. This model favors T^H which can have much higher power than TM_n, where both of them can hold the significance level well.

Example 5. The models are the same as in the previous three examples, except that the data of the second component $Y_{i2} = X_2(t) + \epsilon_{i2}$ are measured.

The empirical sizes and powers are presented in Table 5. Since we only use the second component data, the independent and dependent error components cases are similar. The results show that TM_n can well control the size under all the null hypotheses. It also shows good powers under most of the alternatives. The test still has powers when $\beta = 0$ and $\alpha \neq 0$, meaning that as the unobserved component is correlated with the observed component, the test TM_n could still be sensitive in a certain extent. For H_{11}, H_{12} and H_2, the test does not show powers when $\beta = 0$ and $\alpha \neq 0$ because the magnitude of

Table 4
Empirical sizes and powers in Example 4.

Hypothesis	α	0	0.5	0	0.5	1	0	1
	β	0	0	0.5	0.5	0	1	1
H_4 with ϵ	TM_n	0.042	0.194	0.908	1.000	0.951	1.000	1.000
	T^H	0.040	0.998	1.000	1.000	1.000	1.000	1.000
H_4 with ϵ^*	TM_n	0.051	0.248	0.869	1.000	0.944	1.000	1.000
	T^H	0.060	1.000	1.000	1.000	1.000	1.000	1.000

Table 5
Empirical sizes and powers in Example 5.

Hypothesis	α	0	0.5	0	0.5	1	0	1
	β	0	0	0.5	0.5	0	1	1
H_{11} with ϵ	TM_n	0.042	0.045	0.443	0.395	0.054	1.000	1.000
	T^H	0.057	0.041	1.000	0.994	0.276	1.000	1.000
H_{12} with ϵ	TM_n	0.045	0.059	0.896	0.913	0.053	1.000	1.000
	T^H	0.046	0.088	1.000	1.000	0.339	1.000	1.000
H_{13} with ϵ	TM_n	0.050	0.078	0.543	0.074	0.214	0.999	0.196
	T^H	0.047	0.693	1.000	0.689	0.989	1.000	0.989
H_2 with ϵ	TM_n	0.048	0.059	0.592	0.645	0.066	1.000	1.000
	T^H	0.055	0.204	1.000	1.000	0.646	1.000	1.000
H_3 with ϵ	TM_n	0.062	0.290	0.173	0.808	0.993	0.965	1.000
	T^H	0.038	1.000	1.000	1.000	1.000	1.000	1.000
H_4 with ϵ	TM_n	0.068	0.136	0.807	0.997	0.940	1.000	1.000
	T^H	0.047	0.994	1.000	1.000	1.000	1.000	1.000
H_{11} with ϵ^*	TM_n	0.064	0.061	0.382	0.338	0.051	1.000	1.000
	T^H	0.050	0.054	0.999	0.995	0.253	1.000	1.000
H_{12} with ϵ^*	TM_n	0.051	0.059	0.855	0.886	0.045	1.000	1.000
	T^H	0.053	0.093	1.000	1.000	0.305	1.000	1.000
H_{13} with ϵ^*	TM_n	0.054	0.072	0.506	0.060	0.196	0.998	0.181
	T^H	0.046	0.666	1.000	0.676	0.983	1.000	0.990
H_2 with ϵ^*	TM_n	0.044	0.054	0.526	0.808	0.072	1.000	1.000
	T^H	0.047	0.200	1.000	1.000	0.616	1.000	1.000
H_3 with ϵ^*	TM_n	0.052	0.238	0.176	0.779	0.993	0.946	1.000
	T^H	0.052	1.000	1.000	1.000	1.000	1.000	1.000
H_4 with ϵ^*	TM_n	0.054	0.136	0.756	0.999	0.918	1.000	1.000
	T^H	0.049	0.988	1.000	1.000	1.000	1.000	1.000

disturbance is not sufficiently large. In fixed design we set $(\alpha, \beta) = (4, 0)$, the powers rise up to 0.508 and 0.101 for H_{11} and H_{12} with the independent components error respectively. T^H also performs well and it has larger powers than TM_n.

We conclude that TM_n has well-controlled sizes and good powers for ODE systems with fixed design. It is suitable for fully observed as well as partially observed ODE systems. T^H has good performances with independent components error terms, showing higher powers than TM_n in some cases. However, as seen in Examples 2 and 3, T^H cannot well maintain the significance level for some ODE models with dependent components error terms. Some unreported results also show its failure to maintain the significance level in heteroscedastic cases, while TM_n still works well in this aspect. This would be because T^H needs more restrictive prerequisites to use than TM_n needs. Thus, TM_n is more robust against model settings than T^H. Another advantage of TM_n is its computing stability. It avoids computing the determinants and eigenvalues of certain matrices the algorithm of computing T^H needs, which could cause computational errors. This has also been shown in some unreported numerical studies. This advantage also means lower computing costs. For example, to complete 1000 experiments in case H_{11} with ϵ and $(\alpha, \beta) = (0, 0)$ using MATLAB R2021a, computing T^H spends 126.87 s, while computing TM_n only needs 17.10 s. Overall, these two tests could be complements to each other, and TM_n is of robustness property against model settings.

Remark 1. We have further comment on the test T^H. The observation about the significance level maintenance problem of T^H motivates using an adjusted value $1 + c_n$ in the denominator where $c_n \to 0$ as $n \to \infty$ such that $T^H/(1 + c_n)$ is asymptotically equivalent to T^H, and can bring down the size in finite sample cases. This is a frequently used approach in practice. However, we have tried this method and found that this strategy seems not to work for T^H. In some cases with dependent components error terms or heteroscedasticity, we need a large c_n to help maintain the significance level. In contrast, in some other cases, a large c_n causes the test to be too conservative. Also, it is understandable that a large denominator also brings the test power down. It is hard to find a suitable denominator to use in all cases.

5. The application to COVID-19 epidemic models

The classic SEIR ODE model has been used for respectively describing the infected population in Japan and Algeria (see, Kuniya (2020), and Bentout et al. (2020)). We apply our test, because of its robustness against different models, to check whether the model is plausible.

The form of the mechanical ODE system is as follows:

$$\begin{aligned} \frac{d}{dt}X_1 &= \tau\left\{-a_\beta X_1 X_3\right\} \\ \frac{d}{dt}X_2 &= \tau\left\{a_\beta X_1 X_3 - a_\varepsilon X_2\right\} \\ \frac{d}{dt}X_3 &= \tau\left\{a_\varepsilon X_2 - a_\gamma X_3\right\} \\ \frac{d}{dt}X_4 &= \tau\left\{a_\gamma X_3\right\}. \end{aligned} \tag{3}$$

Here $X_1, \cdots, X_4$ represent the susceptible, exposed, infective and removed populations respectively. We also add a timescale parameter τ to normalize the sample time interval. The ODE model (3) has been applied to fit reported infective cases data during a period starting from the date that the first case was reported. Specifically, we use the data from 15 January to 29 February in Japan and from 25 February to 18 April, 2020 in Algeria. Note that this is a partially observed ODE system since we only observed infective cases corresponding to X_3. Following the instruction in Kuniya (2020) and Bentout et al. (2020), we fix $a_\varepsilon = 0.2$ and $a_\gamma = 0.1$. Further, let $X_1 + X_2 + X_3 + X_4 = 1$, thus each population is actually the proportion to the total population. Also transfer the number of reported infective cases N_I to the corresponding proportion Y_3 by using $Y_3 = N_I/(p_{fr} \times N_p)$, where N_p is the total number of people and p_{fr} is the fraction of infective individuals that are identified by diagnosis. We then use the nonlinear least squares method to estimate a_β. According to Kuniya (2020), p_{fr} almost does not affect the estimation. Here we suppose $p_{fr} = 0.1$. As for N_p, we have $N_p = 1.26 * 10^8$ for Japan and $N_p = 43,411,571$ for Algeria. The initial values are set to be $X(0) = (1 - 1/(p_{fr} \times N_p), 0, 1/(p_{fr} * N_p), 0)$. We use the Epanechnikov kernel $K(u) = 0.75 \times (1 - u^2)$ and the bandwidth $h = 0.15 \times n^{-2/5}$ to ensure it is larger than the interval between two consecutive time points.

For the Japan epidemic data, the value of TM_n is 18.65 and the corresponding p-value is about 0. For the Algeria epidemic data, these values are 16.49 and 0 respectively. These results suggest that the SEIR model, which is the ODE model under the null, may not be plausible (Figs. 1 and 2).

6. Conclusion

This paper studies model checking for ordinary differential equations model with fixed design. The asymptotic properties of a test TM_n, under the null, global, and local alternative hypothesis, are presented. We give two new propositions about U-statistics with varying kernels for independent but non-identical data to derive these results. These propositions are useful to handle many test statistics that are U-statistics with non-i.i.d. data. As an important application, we use TM_n to check the SEIR models by using openly accessible data, finding that it may not be suitable for describing the COVID-19 infected size in Japan and Algeria in the period we described in the application.

As for checking epidemic ODE models, some critical issues are still unsolved. For instance, the parameters may be estimated by a Bayesian method instead of the frequentist method. This is because we may want to estimate them more accurately using the past data. This deserves a further study.

Acknowledgements

The research was supported by an NSFC grant (NSFC12131006), a grant from the University Grants Committee of Hong Kong (HKBU12303921), and the Fundamental Research Funds for the Central Universities (2022JBMC040). The first draft was done when the first author was at Hong Kong Baptist University, and the current version was finished when the first author worked at Beijing Jiaotong University.

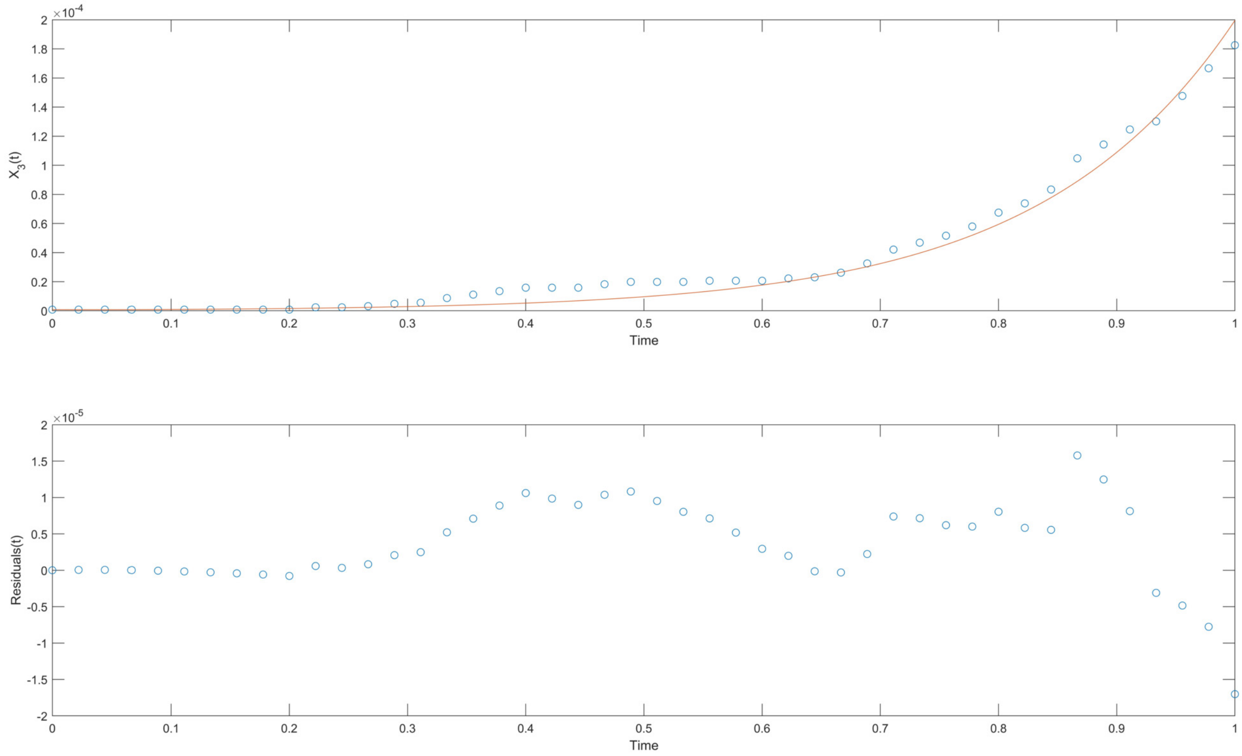

Fig. 1. The data of Japan infective cases: time course of responses and residuals.

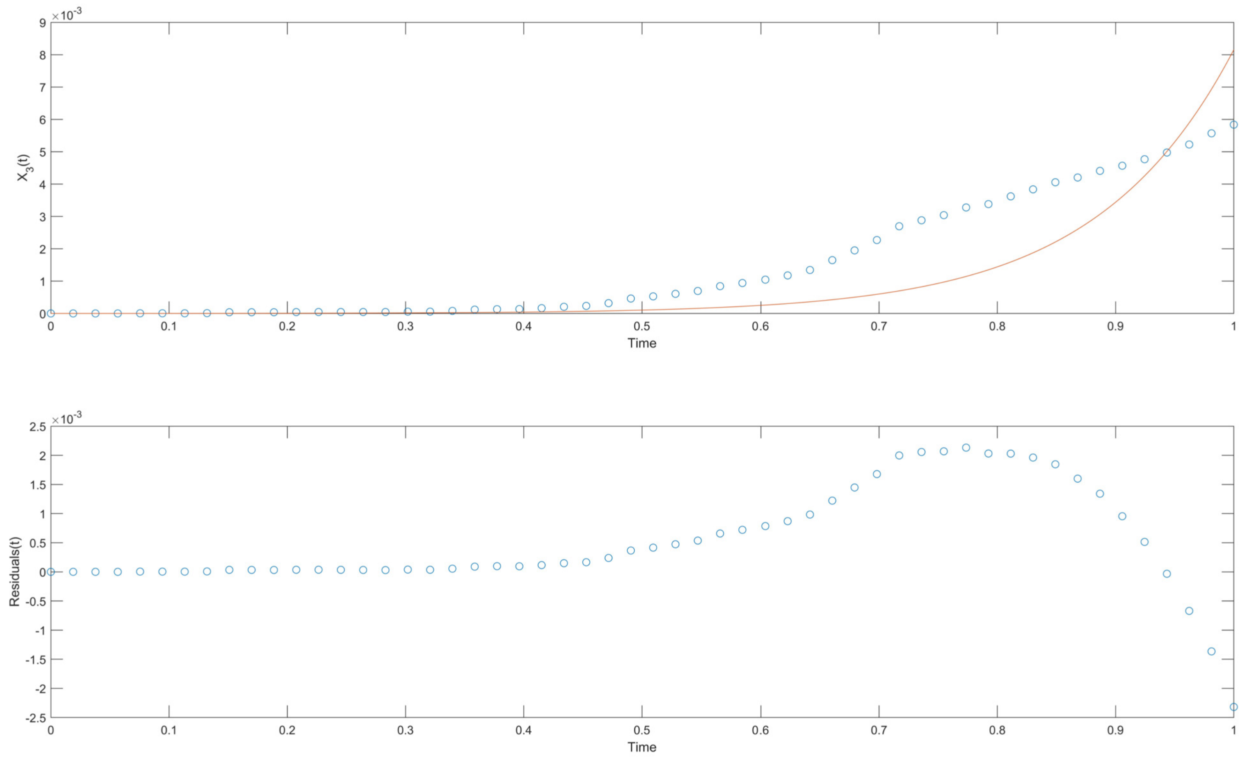

Fig. 2. The data of Algeria infective cases: time course of responses and residuals.

Appendix A. Proof of Proposition 2.1

Proof. This proof is an extension of the proof of Lemma 3.1 of Powell et al. (1989). To prove $\sqrt{n}(U_n - \hat{U}_n) = o_P(1)$, it is sufficient to show $nE\left\{\left\|U_n - \hat{U}_n\right\|^2\right\} = o(1)$. Recall $r_n(z_i) = \binom{n-1}{1}^{-1} \sum_{j\neq i} E\left(h_n(z_i, z_j) \mid z_i\right)$, $\theta_{ni} = E\{r_n(z_i)\}$ and $\theta_{nij} = E\left\{h_n\left(z_i, z_j\right)\right\}$. Define

$$q_n\left(z_i, z_j\right) = \left\{h_n\left(z_i, z_j\right) - r_n(z_i) - r_n\left(z_j\right) + \theta_{ni} + \theta_{nj} - \theta_{nij}\right\},$$

so that

$$U_n - \hat{U}_n = \binom{n}{2}^{-1} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} q_n\left(z_i, z_j\right).$$

The expectation of the squared length of the vector $U_n - \hat{U}_n$ is

$$E\left\{\left\|U_n - \hat{U}_n\right\|^2\right\} = \binom{n}{2}^{-2} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sum_{l=1}^{n-1} \sum_{m=l+1}^{n} E\left\{q_n\left(z_i, z_j\right)' q_n(z_l, z_m)\right\}.$$

If $i \neq l$ and $j \neq m$, we have $E\left\{q_n\left(z_i, z_j\right)' q_n(z_l, z_m)\right\} = 0$ since $z_i, i = 1, \cdots, n$ are independent random vectors. At next we consider the case that there is only one same terms in $q_n\left(z_i, z_j\right)$ and $q_n(z_l, z_m)$. Without loss of generality, we assume $i = m$. Then we have

$$\begin{aligned}
& E\left\{q_n\left(z_i, z_j\right)' q_n(z_i, z_l)\right\} \\
=& E\left\{h_n\left(z_i, z_j\right) - r_n(z_i) - r_n\left(z_j\right) + \theta_{ni} + \theta_{nj} - \theta_{nij}\right\}' \\
& \times \{h_n(z_i, z_l) - r_n(z_i) - r_n(z_l) + \theta_{ni} + \theta_{nl} - \theta_{nil}\} \\
=& E\left\{h_n\left(z_i, z_j\right)\right\}' \{h_n(z_i, z_l)\} - E\left\{h_n\left(z_i, z_j\right)\right\}' \{r_n(z_i)\} - E\left\{h_n\left(z_i, z_j\right)\right\}' \{r_n(z_l)\} \\
& + E\left\{h_n\left(z_i, z_j\right)\right\}' \{\theta_{ni} + \theta_{nl} - \theta_{nil}\} - E\{r_n(z_i)\}' \{h_n(z_i, z_l)\} + E\{r_n(z_i)\}' \{r_n(z_i)\} \\
& + E\{r_n(z_i)\}' \{r_n(z_l)\} - E\{r_n(z_i)\}' \{\theta_{ni} + \theta_{nl} - \theta_{nil}\} - E\left\{r_n\left(z_j\right)\right\}' \{h_n(z_i, z_l)\} \\
& + E\left\{r_n\left(z_j\right)\right\}' \{r_n(z_i)\} + E\left\{r_n\left(z_j\right)\right\}' \{r_n(z_l)\} - E\left\{r_n\left(z_j\right)\right\}' \{\theta_{ni} + \theta_{nl} - \theta_{nil}\} \\
& + E\left\{\theta_{ni} + \theta_{nj} - \theta_{nij}\right\}' \{h_n(z_i, z_l)\} - E\left\{\theta_{ni} + \theta_{nj} - \theta_{nij}\right\}' \{r_n(z_i)\} \\
& - E\left\{\theta_{ni} + \theta_{nj} - \theta_{nij}\right\}' \{r_n(z_l)\} + E\left\{\theta_{ni} + \theta_{nj} - \theta_{nij}\right\}' \{\theta_{ni} + \theta_{nl} - \theta_{nil}\} \\
=& E\left\{r_{nj}(z_i)\right\}' \{r_{nl}(z_i)\} - E\left\{r_{nj}(z_i)\right\}' \{r_n(z_i)\} - \theta_{nij}'\theta_{nl} + \theta_{nij}' \{\theta_{ni} + \theta_{nl} - \theta_{nil}\} \\
& - E\{r_n(z_i)\}' \{r_{nl}(z_i)\} + E\{r_n(z_i)\}' \{r_n(z_i)\} + \theta_{ni}'\theta_{nl} - \theta_{ni}' \{\theta_{ni} + \theta_{nl} - \theta_{nil}\} \\
& - \theta_{nj}'\theta_{nil} + \theta_{nj}'\theta_{ni} + \theta_{nj}'\theta_{nl} - \theta_{nj}' \{\theta_{ni} + \theta_{nl} - \theta_{nil}\} + \left\{\theta_{ni} + \theta_{nj} - \theta_{nij}\right\}' \theta_{nil} \\
& - \left\{\theta_{ni} + \theta_{nj} - \theta_{nij}\right\}' \theta_{ni} - \left\{\theta_{ni} + \theta_{nj} - \theta_{nij}\right\}' \theta_{nl} + \left\{\theta_{ni} + \theta_{nj} - \theta_{nij}\right\}' \{\theta_{ni} + \theta_{nl} - \theta_{nil}\} \\
=& E\left\{r_{nj}(z_i)\right\}' \{r_{nl}(z_i)\} - E\left\{r_{nj}(z_i)\right\}' \{r_n(z_i)\} + \theta_{nij}'\theta_{ni} - E\{r_n(z_i)\}' \{r_{nl}(z_i)\} \\
& + E\{r_n(z_i)\}' \{r_n(z_i)\} + \theta_{ni}'\theta_{nil} - \theta_{nij}'\theta_{nil} - \theta_{ni}'\theta_{ni}.
\end{aligned}$$

Note $\theta_{ni} = \frac{1}{n-1} \sum_{j\neq i} \theta_{nij}$ and $r_n(z_i) = \frac{1}{n-1} \sum_{j\neq i} r_{nj}(z_i)$. We have

$$\begin{aligned}
& \theta_{nij}'\theta_{ni} + \theta_{ni}'\theta_{nil} - \theta_{nij}'\theta_{nil} - \theta_{ni}'\theta_{ni} \\
=& \frac{1}{n-1} \sum_{k\neq i} \theta_{nij}'\theta_{nik} + \frac{1}{n-1} \sum_{k\neq i} \theta_{nil}'\theta_{nik} - \frac{1}{(n-1)^2} \sum_{k_1\neq i} \sum_{k_2\neq i} \theta_{nik_1}'\theta_{nik_2} - \theta_{nij}'\theta_{nil},
\end{aligned}$$

$$\begin{aligned}
& E\left\{r_{nj}(z_i)\right\}' \{r_{nl}(z_i)\} - E\left\{r_{nj}(z_i)\right\}' \{r_n(z_i)\} - E\{r_n(z_i)\}' \{r_{nl}(z_i)\} \\
& + E\{r_n(z_i)\}' \{r_n(z_i)\} \\
=& E\left\{r_{nj}(z_i)\right\}' \{r_{nl}(z_i)\} - E\left\{r_{nj}(z_i)\right\}' \left\{\frac{1}{n-1} \sum_{j\neq i} r_{nj}(z_i)\right\}
\end{aligned}$$

$$- E\left\{\frac{1}{n-1}\sum_{j\neq i} r_{nj}(z_i)\right\}'\{r_{nl}(z_i)\}$$
$$+ E\left\{\frac{1}{n-1}\sum_{j\neq i} r_{nj}(z_i)\right\}'\left\{\frac{1}{n-1}\sum_{j\neq i} r_{nj}(z_i)\right\}.$$

Thus it can be shown that

$$\begin{aligned}
&\frac{1}{n^2(n-1)^2}\sum_{i=1}^{n}\sum_{j\neq i}\sum_{\substack{l\neq i\\ l\neq j}} E\left\{q_n\left(z_i,z_j\right)' q_n\left(z_l,z_i\right)\right\}\\
=&\frac{1}{n^2(n-1)^2}\sum_{i=1}^{n}\sum_{j\neq i}\sum_{l\neq i}\Bigg[\frac{1}{n-1}\sum_{k\neq i}\theta_{nij}'\theta_{nik}+\frac{1}{n-1}\sum_{k\neq i}\theta_{nil}'\theta_{nik}\\
&-\frac{1}{(n-1)^2}\sum_{k_1\neq i}\sum_{k_2\neq i}\theta_{nik_1}'\theta_{nik_2}-\theta_{nij}'\theta_{nil}+E\left\{r_{nj}(z_i)\right\}'\{r_{nl}(z_i)\}\\
&-E\left\{r_{nj}(z_i)\right\}'\left\{\frac{1}{n-1}\sum_{j\neq i} r_{nj}(z_i)\right\}-E\left\{\frac{1}{n-1}\sum_{j\neq i} r_{nj}(z_i)\right\}'\{r_{nl}(z_i)\}\\
&+E\left\{\frac{1}{n-1}\sum_{j\neq i} r_{nj}(z_i)\right\}'\left\{\frac{1}{n-1}\sum_{j\neq i} r_{nj}(z_i)\right\}\Bigg]\\
&-\frac{1}{n^2(n-1)^2}\sum_{i=1}^{n}\sum_{j\neq i} E\left\{q_n\left(z_i,z_j\right)' q_n\left(z_i,z_j\right)\right\}\\
=&\frac{2}{n^2(n-1)^2}\sum_{i=1}^{n}\sum_{j\neq i}\sum_{k\neq i}[\theta_{nij}'\theta_{nik}+E\left\{r_{nj}(z_i)\right\}'\{r_{nk}(z_i)\}]\\
&-\frac{1}{n^2(n-1)^2}\sum_{i=1}^{n}\sum_{k_1\neq i}\sum_{k_2\neq i}[\theta_{nik_1}'\theta_{nik_2}+E\left\{r_{nk_1}(z_i)\right\}'\left\{r_{nk_2}(z_i)\right\}]\\
&-\frac{1}{n^2(n-1)^2}\sum_{i=1}^{n}\sum_{j\neq i}\sum_{l\neq i}[\theta_{nij}'\theta_{nil}+E\left\{r_{nj}(z_i)\right\}'\{r_{nl}(z_i)\}]-o(\frac{1}{n})\\
=&o(\frac{1}{n}).
\end{aligned}$$

Then we have

$$E\left\|U_n-\hat{U}_n\right\|^2=\binom{n}{2}^{-2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} E\left\|q_n\left(z_i,z_j\right)\right\|^2+o(\frac{1}{n}).$$

Since each $O\left(E\left\|q_n\left(z_i,z_j\right)\right\|^2\right)=O\left(E\left\|h_n\left(z_i,z_j\right)\right\|^2\right)=o(n)$. We conclude

$$\begin{aligned}
nE\left\|U_n-\hat{U}_n\right\|^2&=n\binom{n}{2}^{-2}O\left(n^2\right)o(n)+o(1)\\
&=o(1). \quad \square
\end{aligned}$$

Appendix B. Proof of Proposition 2.2

Proof. This proof is an extension of the proof of Theorem 1 of Hall (1984). We apply Brown's Martingale central limit theorem (Brown (1971); Hall and Heyde (2014)) and only need to check the following two conditions:

$$s_n^{-2}\sum_{i=2}^{n}E\left\{Y_{ni}^2 I\left(|Y_{ni}|>\eta s_n\right)\right\}\to 0, \text{ for each } \eta>0, \tag{B.1}$$

$$s_n^{-2}V_n^2\to 1, \text{ in probability}, \tag{B.2}$$

where $Y_{ni}=\sum_{j=1}^{i-1}h_n\left(z_i,z_j\right)$ and $s_n^2=E\left[\{\sum\sum_{1\leqslant i<j\leqslant n}h_n\left(z_i,z_j\right)\}^2\right]$.

For every i and j, we have $E\{h_n(z_i,z_j|z_i)\}=0$. Then

$$E\left(Y_{ni}^2\right)=\sum_{j=1}^{i-1}\sum_{k=1}^{i-1}E\left\{h_n\left(z_i,z_j\right)h_n\left(z_i,z_k\right)\right\}$$
$$=\sum_{j=1}^{i-1}E\left\{h_n^2\left(z_i,z_j\right)\right\},$$

$$s_n^2=E(\sum_{i=2}^{n}Y_{ni})^2=\sum_{i=2}^{n}E\left(Y_{ni}^2\right)=\sum_{i=2}^{n}\sum_{j=1}^{i-1}E\left\{h_n^2\left(z_i,z_j\right)\right\}.$$

Furthermore, we have

$$E\left\{h_n\left(z_1,z_2\right)h_n\left(z_1,z_3\right)h_n\left(z_1,z_4\right)h_n\left(z_1,z_5\right)\right\}=E\left\{h_n\left(z_1,z_2\right)h_n^3\left(z_1,z_3\right)\right\}=0.$$

Thus

$$E\left(Y_{ni}^4\right)=\sum_{j=1}^{i-1}E\left\{h_n^4\left(z_i,z_j\right)\right\}+3\sum_{1\leqslant j,k\leqslant i-1;\, j\neq k}E\left\{h_n^2\left(z_i,z_j\right)h_n^2\left(z_i,z_k\right)\right\}$$

whence

$$\begin{aligned}
&\sum_{i=2}^{n}E\left(Y_{ni}^4\right)\\
&=\sum_{i=2}^{n}\sum_{j=1}^{i-1}E\left\{h_n^4\left(z_i,z_j\right)\right\}+3\sum_{i=2}^{n}\sum_{1\leqslant j,k\leqslant i-1;\, j\neq k}E\left\{h_n^2\left(z_i,z_j\right)h_n^2\left(z_i,z_k\right)\right\}\\
&\leq\sum_{i=2}^{n}\sum_{j=1}^{i-1}E\left\{h_n^4\left(z_i,z_j\right)\right\}+\frac{3}{2}\sum_{i=2}^{n}\sum_{1\leqslant j,k\leqslant i-1;\, j\neq k}\{h_n^4(z_i,z_j)+h_n^4(z_i,z_k)\}\\
&\leq C_1\sum_{i=2}^{n}\sum_{j=1}^{i-1}(i-1)E\left\{h_n^4\left(z_i,z_j\right)\right\}\\
&\leq C_2 n\sum_{i=2}^{n}\sum_{j=1}^{i-1}E\left\{h_n^4\left(z_i,z_j\right)\right\},
\end{aligned} \tag{B.3}$$

where C_1 and C_2 are constants.

From equation (B.3) and the condition

$$\frac{\sum_{2\leqslant i_1\leqslant i_2\leqslant n}\sum_{1\leqslant j_1<k_1\leqslant i_1-1}E\left\{G_{ni_1}\left(z_{j_1},z_{k_1}\right)G_{ni_2}\left(z_{j_1},z_{k_1}\right)\right\}+n\sum_{i=2}^{n}\sum_{j=1}^{i-1}E\left\{h_n^4\left(z_i,z_j\right)\right\}}{\left[\sum_{i=2}^{n}\sum_{j=1}^{i-1}E\left\{h_n^2\left(z_i,z_j\right)\right\}\right]^2}$$
$$\to 0,$$

it can be shown that

$$s_n^{-4}\sum_{i=2}^{n}E\left(Y_{ni}^4\right)\leq\frac{Cn\sum_{i=2}^{n}\sum_{j=1}^{i-1}E\left\{h_n^4\left(z_i,z_j\right)\right\}}{(\sum_{i=2}^{n}\sum_{j=1}^{i-1}E\left\{h_n^2\left(z_i,z_j\right)\right\})^2}\to 0,$$

which imply the first condition (B.1). On the other hand,

$$v_{ni} = E\left(Y_{ni}^2 \mid z_1, \ldots, z_{i-1}\right) = \sum_{j=1}^{i-1}\sum_{k=1}^{i-1} G_{ni}\left(z_j, z_k\right)$$
$$= 2 \sum_{1\leqslant j<k\leqslant i-1} G_{ni}\left(z_j, z_k\right) + \sum_{j=1}^{i-1} G_{ni}\left(z_j, z_j\right).$$

If $j_1 \leq k_1$ and $j_2 \leq k_2$, then

$$\begin{aligned}
& E\left\{G_{ni_1}\left(z_{j_1}, z_{k_1}\right) G_{ni_2}\left(z_{j_2}, z_{k_2}\right)\right\} \\
=& E\left\{G_{ni_1}\left(z_{j_1}, z_{j_1}\right) G_{ni_2}\left(z_{j_1}, z_{j_1}\right)\right\} \quad \text{if} \quad j_1 = k_1 = j_2 = k_2 \\
=& \left[E\left\{G_{ni_1}\left(z_{j_1}, z_{j_1}\right)\right\}\right]\left[E\left\{G_{ni_2}\left(z_{j_2}, z_{j_2}\right)\right\}\right] \quad \text{if} \quad j_1 = k_1 \neq j_2 = k_2 \\
=& E\left\{G_{ni_1}\left(z_{j_1}, z_{k_1}\right) G_{ni_2}\left(z_{j_1}, z_{k_1}\right)\right\} \quad \text{if} \quad j_1 = j_2, k_1 = k_2, j_1 < k_1 \\
=& 0 \quad \text{otherwise.}
\end{aligned}$$

Hence if $i_1 \leq i_2$,

$$\begin{aligned}
& E\left(v_{ni_1} v_{ni_2}\right) \\
=& 4 \sum_{1\leqslant j_1<k_1\leqslant i_1-1} E\left\{G_{ni_1}\left(z_{j_1}, z_{k_1}\right) G_{ni_2}\left(z_{j_1}, z_{k_1}\right)\right\} \\
& + \sum_{j_1=1}^{i_1-1}\sum_{j_2=1}^{i_2-1} E\left\{G_{ni_1}\left(z_{j_1}, z_{j_1}\right)\right\} E\left\{G_{ni_2}\left(z_{j_2}, z_{j_2}\right)\right\} \\
& + \sum_{j=1}^{i_1-1}\left(E\left\{G_{ni_1}\left(z_{j_1}, z_{j_1}\right) G_{ni_2}\left(z_{j_1}, z_{j_1}\right)\right\} - E\left\{G_{ni_1}\left(z_{j_1}, z_{j_1}\right)\right\} E\left\{G_{ni_2}\left(z_{j_1}, z_{j_1}\right)\right\}\right) \\
=& 4 \sum_{1\leqslant j_1<k_1\leqslant i_1-1} E\left\{G_{ni_1}\left(z_{j_1}, z_{k_1}\right) G_{ni_2}\left(z_{j_1}, z_{k_1}\right)\right\} \\
& + \sum_{j_1=1}^{i_1-1}\sum_{j_2=1}^{i_2-1} E\left\{G_{ni_1}\left(z_{j_1}, z_{j_1}\right)\right\} E\left\{G_{ni_2}\left(z_{j_2}, z_{j_2}\right)\right\} \\
& + \sum_{j_1=1}^{i_1-1} cov\left\{G_{ni_1}\left(z_{j_1}, z_{j_1}\right), G_{ni_2}\left(z_{j_1}, z_{j_1}\right)\right\}.
\end{aligned}$$

Then we have

$$\begin{aligned}
E\left(V_n^4\right) =& 2 \sum_{2\leqslant i_1<i_2\leqslant n} E\left(v_{ni_1} v_{ni_2}\right) + \sum_{i=2}^{n} E\left(v_{ni}^2\right) \\
=& 8 \sum_{2\leqslant i_1<i_2\leqslant n}\ \sum_{1\leqslant j_1<k_1\leqslant i_1-1} E\left\{G_{ni_1}\left(z_{j_1}, z_{k_1}\right) G_{ni_2}\left(z_{j_1}, z_{k_1}\right)\right\} \\
& + 2 \sum_{2\leqslant i_1<i_2\leqslant n}\ \sum_{j_1=1}^{i_1-1}\sum_{j_2=1}^{i_2-1} E\left\{G_{ni_1}\left(z_{j_1}, z_{j_1}\right)\right\} E\left\{G_{ni_2}\left(z_{j_2}, z_{j_2}\right)\right\} \\
& + 2 \sum_{2\leqslant i_1<i_2\leqslant n}\ \sum_{j_1=1}^{i_1-1} cov\left\{G_{ni_1}\left(z_{j_1}, z_{j_1}\right), G_{ni_2}\left(z_{j_1}, z_{j_1}\right)\right\} \\
& + 4 \sum_{2\leqslant i\leqslant n}\ \sum_{1\leqslant j_1<k_1\leqslant i-1} E\left\{G_{ni}\left(z_{j_1}, z_{k_1}\right) G_{ni}\left(z_{j_1}, z_{k_1}\right)\right\} \\
& + \sum_{2\leqslant i\leqslant n}\ \sum_{j_1=1}^{i-1}\sum_{j_2=1}^{i-1} E\left\{G_{ni}\left(z_{j_1}, z_{j_1}\right)\right\} E\left\{G_{ni}\left(z_{j_2}, z_{j_2}\right)\right\}
\end{aligned}$$

$$+\sum_{2\leqslant i\leqslant n}\sum_{j_1=1}^{i-1} cov\left\{G_{ni}\left(z_{j_1}, z_{j_1}\right), G_{ni}\left(z_{j_1}, z_{j_1}\right)\right\}$$

whence

$$E\left(V_n^2-s_n^2\right)^2=E\left(V_n^4\right)-s_n^4$$
$$\leqslant C\left[\sum_{2\leqslant i_1\leqslant i_2\leqslant n}\sum_{1\leqslant j_1<k_1\leqslant i_1-1} E\left\{G_{ni_1}\left(z_{j_1}, z_{k_1}\right) G_{ni_2}\left(z_{j_1}, z_{k_1}\right)\right\}+n\sum_{i=2}^{n}\sum_{j=1}^{i-1} E\left\{h_n^4\left(z_i, z_j\right)\right\}\right], \tag{B.4}$$

with C a constant. According to the condition

$$\frac{\sum_{2\leqslant i_1\leqslant i_2\leqslant n}\sum_{1\leqslant j_1<k_1\leqslant i_1-1} E\left\{G_{ni_1}\left(z_{j_1}, z_{k_1}\right) G_{ni_2}\left(z_{j_1}, z_{k_1}\right)\right\}+n\sum_{i=2}^{n}\sum_{j=1}^{i-1} E\left\{h_n^4\left(z_i, z_j\right)\right\}}{\left(\sum_{i=2}^{n}\sum_{j=1}^{i-1} E\left\{h_n^2\left(z_i, z_j\right)\right\}\right)^2}$$
$$\rightarrow 0,$$

the above equation (B.4) proves the second condition (B.2) by showing that $s_n^{-4}E\left(V_n^2-s_n^2\right)^2\rightarrow 0$. □

Appendix C. Proof of Theorem 1 (under the null hypothesis and global alternatives)

Proof. The proof is very similar to the proof for Theorem 1 of Liu et al. (2021), except that we use Propositions 1-2 instead of the existing theories of U-statistics for i.i.d. data. We then only give a sketch of the proof and focus on the main different steps. Without loss of generality, we assume, the ODE system is all observed hereafter.

1. Under the null hypothesis, we first study the limiting property of V_n.

For every component k, we decompose V_{nk} into three terms:

$$\begin{aligned}
V_{nk}=&\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\frac{1}{h}K\left(\frac{t_i-t_j}{h}\right)e_{ik}e_{jk}\\
=&\left[\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\frac{1}{h}K\left(\frac{t_i-t_j}{h}\right)\eta_{ik}\eta_{jk}\right]\\
&-2\left[\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\frac{1}{h}K\left(\frac{t_i-t_j}{h}\right)\eta_{ik}\left\{F_k\left(t_j,\hat{\theta}\right)-F_k\left(t_j,\theta_0\right)\right\}\right]\\
&+\left[\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\frac{1}{h}K\left(\frac{t_i-t_j}{h}\right)\right.\\
&\left.\times\left\{F_k\left(t_j,\hat{\theta}\right)-F_k\left(t_j,\theta_0\right)\right\}\left\{F_k\left(t_j,\hat{\theta}\right)-F_k\left(t_j,\theta_0\right)\right\}\right]\\
\equiv& V_{1nk}-2V_{2nk}+V_{3nk}.
\end{aligned}\tag{C.1}$$

Then we can write the vector V_n as $V_{1n}-2V_{2n}+V_{3n}$. We now show that $nh^{1/2}V_{1n}\rightarrow N(0,\Sigma)$ where Σ is a symmetric matrix with the entries: for any pair (k_1,k_2) with $1\leq k_1,k_2\leq d$,

$$\Sigma_{k_1k_2}=2\int K^2(u)du\times\int\{\sigma_{k1k2}(t)\}^2p^2(t)dt.$$

According to the Cramér-Wald device, we can prove this result by verifying that for every $\lambda\in R^p$, $nh^{1/2}\lambda^{\top}V_{1n}\rightarrow N(0,\lambda^{\top}\Sigma\lambda)$ in distribution. We then write $\lambda^{\top}V_{1n}$ in a U-statistic form with the kernel:

$$\tilde{H}_n\left(z_i,z_j\right)=\sum_{k=1}^{d}\frac{\lambda_k}{h}K\left(\frac{t_i-t_j}{h}\right)\eta_{ik}\eta_{jk}$$

where $z_i = (t_i, \eta_i)$. Since it is a one-dimensional degenerate U-statistic, we apply Proposition 2 to obtain its asymptotic distribution. To this end, we verity the conditions of this proposition. For some $(z_i, z_j) = (z_1, z_2)$, we have

$$\begin{aligned}
&E\left\{\tilde{H}_n^2(z_1, z_2)\right\} \\
&= E\left[E\left\{\tilde{H}_n^2(z_1, z_2) \mid t_1, t_2\right\}\right] \\
&= \frac{1}{h^2}\sum_{k_1=1}^{d}\sum_{k_2=1}^{d}\lambda_{k_1}\lambda_{k_2}K^2\left(\frac{t_1-t_2}{h}\right)\sigma_{k_1k_2}(t_1)\sigma_{k_1k_2}(t_2).
\end{aligned}$$

Then

$$\begin{aligned}
&\sum_{i=2}^{n}\sum_{j=1}^{i-1}E\left\{\tilde{H}_n^2(z_i, z_j)\right\} \\
=&\sum_{i=2}^{n}\sum_{j=1}^{i-1}\frac{1}{h^2}\sum_{k_1=1}^{d}\sum_{k_2=1}^{d}\lambda_{k_1}\lambda_{k_2}K^2\left(\frac{t_1-t_2}{h}\right)\sigma_{k_1k_2}(t_1)\sigma_{k_1k_2}(t_2) \\
=&\frac{n(n-1)}{2h^2}\sum_{k_1=1}^{d}\sum_{k_2=1}^{d}\sum_{i=2}^{n}\sum_{j=1}^{i-1}\frac{2}{n(n-1)}\lambda_{k_1}\lambda_{k_2}K^2\left(\frac{t_i-t_j}{h}\right)\sigma_{k_1k_2}(t_i)\sigma_{k_1k_2}(t_j) \\
=&\frac{n(n-1)}{2h^2}\sum_{k_1=1}^{d}\sum_{k_2=1}^{d}\int\lambda_{k_1}\lambda_{k_2}K^2\left(\frac{t_1-t_2}{h}\right)\sigma_{k_1k_2}(t_1)\sigma_{k_1k_2}(t_2)\,dP_n(t_1)dP_n(t_2) \\
=&\frac{n(n-1)}{2h^2}\sum_{k_1=1}^{d}\sum_{k_2=1}^{d}\lambda_{k_1}\lambda_{k_2}\int K^2\left(\frac{t_1-t_2}{h}\right)\sigma_{k_1k_2}(t_1)\sigma_{k_1k_2}(t_2)\,dP(t_1)dP(t_2)+O(1/h^2) \\
=&\frac{n(n-1)}{2h^2}\sum_{k_1=1}^{d}\sum_{k_2=1}^{d}\lambda_{k_1}\lambda_{k_2}\int K^2\left(\frac{t_1-t_2}{h}\right)\sigma_{k_1k_2}(t_1)\sigma_{k_1k_2}(t_2)\,p(t_1)p(t_2)dt_1dt_2+O(1/h^2) \\
=&\frac{n(n-1)}{2h}\sum_{k_1=1}^{d}\sum_{k_2=1}^{d}\lambda_{k_1}\lambda_{k_2}\int K^2(u)\sigma_{k_1k_2}(t_1)\sigma_{k_1k_2}(t_1-hu)p(t)p(t-hu)hdtdu+O(1/h^2) \\
=&\frac{n(n-1)}{2h}\sum_{k_1=1}^{d}\sum_{k_2=1}^{d}\lambda_{k_1}\lambda_{k_2}\int K^2(u)du\int\left\{\sigma_{k_1k_2}(t)\right\}^2p^2(t)dt+\mathrm{o}\left(n^2/h\right) \\
=&\mathrm{O}\left(n^2/h\right).
\end{aligned}$$

The fourth step in the above equation is based on the fixed design sampling condition. Similarly,

$$\begin{aligned}
&E\left\{\tilde{H}_n^4(z_1, z_2)\right\} \\
=&\frac{1}{h^4}K^4\left(\frac{t_1-t_2}{h}\right)\sum_{k_1=1}^{d}\sum_{k_2=1}^{d}\sum_{k_3=1}^{d}\sum_{k_4=1}^{d}\lambda_{k_1}\lambda_{k_2}\lambda_{k_3}\lambda_{k_4}\left[E\left\{\eta_{1k_1}\eta_{1k_2}\eta_{1k_3}\eta_{1k_4}\right\}\right. \\
&\left.\times E\left\{\eta_{2k_1}\eta_{2k_2}\eta_{2k_3}\eta_{2k_4}\right\}\right] \\
=&\frac{1}{h^4}\left[\sum_{k_1=1}^{d}\sum_{k_2=1}^{d}\sum_{k_3=1}^{d}\sum_{k_4=1}^{d}\lambda_{k_1}\lambda_{k_2}\lambda_{k_3}\lambda_{k_4}K^4\left(\frac{t_1-t_2}{h}\right)\{\sigma_{k_1k_2k_3k_4}(t_1)\sigma_{k_1k_2k_3k_4}(t_2)\}\right].
\end{aligned}$$

$$\begin{aligned}
&\sum_{i=2}^{n}\sum_{j=1}^{i-1}E\left\{\tilde{H}_n^4(z_i, z_j)\right\} \\
=&\sum_{i=2}^{n}\sum_{j=1}^{i-1}\frac{1}{h^4}\left[\sum_{k_1=1}^{d}\sum_{k_2=1}^{d}\sum_{k_3=1}^{d}\sum_{k_4=1}^{d}\lambda_{k_1}\lambda_{k_2}\lambda_{k_3}\lambda_{k_4}K^4\left(\frac{t_1-t_2}{h}\right)\{\sigma_{k_1k_2k_3k_4}(t_1)\sigma_{k_1k_2k_3k_4}(t_2)\}\right]
\end{aligned}$$

$$
\begin{aligned}
=&\frac{n(n-1)}{2h^4}\left[\sum_{k_1=1}^{d}\sum_{k_2=1}^{d}\sum_{k_3=1}^{d}\sum_{k_4=1}^{d}\lambda_{k_1}\lambda_{k_2}\lambda_{k_3}\lambda_{k_4}\int K^4(\frac{t_1-t_2}{h})\{\sigma_{k_1k_2k_3k_4}(t_1)\sigma_{k_1k_2k_3k_4}(t_2)\}dt_1dt_2\right]\\
&+O(1/h^4)\\
=&\frac{n(n-1)}{2h^4}\left[\sum_{k_1=1}^{d}\sum_{k_2=1}^{d}\sum_{k_3=1}^{d}\sum_{k_4=1}^{d}\lambda_{k_1}\lambda_{k_2}\lambda_{k_3}\lambda_{k_4}\int K^4(u)\{\sigma_{k_1k_2k_3k_4}(t_1)\right.\\
&\left.\times\sigma_{k_1k_2k_3k_4}(t_1-hu)\}p(t_1)p(t_1-hu)hdt_1du\right]+O(1/h^4)\\
=&O\left(n^2/h^3\right).
\end{aligned}
$$

Also, by writing $(j_1,j_2,i_1,i_2)=(1,2,3,4)$ for simplicity, we have

$$
\begin{aligned}
&E\left\{G_{ni_1}\left(z_{j_1},z_{j_2}\right)G_{ni_2}\left(z_{j_1},z_{j_2}\right)\right\}\\
=&E\left[E\left\{\tilde{H}_n\left(z_3,z_1\right)\tilde{H}_n\left(z_3,z_2\right)\mid z_1,z_2\right\}E\left\{\tilde{H}_n\left(z_4,z_1\right)\tilde{H}_n\left(z_4,z_2\right)\mid z_1,z_2\right\}\right]\\
=&\frac{1}{h^4}E\left(\left[\sum_{k_1=1}^{d}\sum_{k_2=1}^{d}\lambda_{k_1}\lambda_{k_2}\eta_{1k_1}\eta_{2k_2}K\left(\frac{t_3-t_1}{h}\right)K\left(\frac{t_3-t_2}{h}\right)\sigma_{k_1k_2}\left(t_3\right)\right]\right.\\
&\left.\times\left[\sum_{k_1=1}^{d}\sum_{k_2=1}^{d}\lambda_{k_1}\lambda_{k_2}\eta_{1k_1}\eta_{2k_2}K\left(\frac{t_4-t_1}{h}\right)K\left(\frac{t_4-t_2}{h}\right)\sigma_{k_1k_2}\left(t_4\right)\right]\right)\\
=&\frac{1}{h^4}\sum_{k_1=1}^{d}\sum_{k_2=1}^{d}\sum_{k_3=1}^{d}\sum_{k_4=1}^{d}\lambda_{k_1}\lambda_{k_2}\lambda_{k_3}\lambda_{k_4}\sigma_{k_1k_3}\left(t_1\right)\sigma_{k_2k_4}\left(t_2\right)K\left(\frac{t_3-t_1}{h}\right)K\left(\frac{t_3-t_2}{h}\right)\\
&\times K\left(\frac{t_4-t_1}{h}\right)K\left(\frac{t_4-t_2}{h}\right)\sigma_{k_1k_2}\left(t_3\right)\sigma_{k_3k_4}\left(t_4\right).\\
&\sum_{2\leqslant i_1\leqslant i_2\leqslant n}\ \sum_{1\leqslant j_1<j_2\leqslant i_1-1}E\left\{G_{ni_1}\left(z_{j_1},z_{j_2}\right)G_{ni_2}\left(z_{j_1},z_{j_2}\right)\right\}\\
\leq&\frac{Cn^4}{h^4}\sum_{k_1=1}^{d}\sum_{k_2=1}^{d}\sum_{k_3=1}^{d}\sum_{k_4=1}^{d}\lambda_{k_1}\lambda_{k_2}\lambda_{k_3}\lambda_{k_4}\int\sigma_{k_1k_3}\left(t_1\right)\sigma_{k_2k_4}\left(t_2\right)K\left(\frac{t_3-t_1}{h}\right)K\left(\frac{t_3-t_2}{h}\right)\\
&\times K\left(\frac{t_4-t_1}{h}\right)K\left(\frac{t_4-t_2}{h}\right)\sigma_{k_1k_2}\left(t_3\right)\sigma_{k_3k_4}\left(t_4\right)p(t_1)p(t_2)p(t_3)p(t_4)dt_1dt_2dt_3dt_4\\
&+O(1/h^4)\\
=&\frac{Cn^4}{h^2}\sum_{k_1=1}^{d}\sum_{k_2=1}^{d}\sum_{k_3=1}^{d}\sum_{k_4=1}^{d}\lambda_{k_1}\lambda_{k_2}\lambda_{k_3}\lambda_{k_4}\int\sigma_{k_1k_3}\left(t_1\right)\sigma_{k_2k_4}\left(t_2\right)K\left(u_1+\frac{t_1-t_2}{h}\right)K\left(u_2+\frac{t_1-t_2}{h}\right)\\
&\times K\left(u_1\right)K\left(u_2\right)\sigma_{k_1k_2}\left(t_1+hu_1\right)\sigma_{k_3k_4}\left(t_1+hu_2\right)p(t_1)p(t_2)p(t_1+hu_1)p(t_1+hu_2)dt_1dt_2du_1du_2\\
&+O(1/h^4)\\
=&\frac{Cn^4}{h}\sum_{k_1=1}^{d}\sum_{k_2=1}^{d}\sum_{k_3=1}^{d}\sum_{k_4=1}^{d}\lambda_{k_1}\lambda_{k_2}\lambda_{k_3}\lambda_{k_4}\int\sigma_{k_1k_3}\left(t_1\right)\sigma_{k_2k_4}\left(t_1-hu_3\right)K\left(u_1\right)K\left(u_1+u_3\right)K\left(u_2+u_3\right)\\
&\times K\left(u_2\right)\sigma_{k_1k_2}\left(t_1+hu_1\right)\sigma_{k_3k_4}\left(t_1+hu_2\right)p(t_1)p(t_1-hu_3)p(t_1+hu_1)p(t_1+hu_2)dt_1du_1du_2du_3\\
&+O(1/h^4)\\
=&\frac{Cn^4}{h}\sum_{k_1=1}^{d}\sum_{k_2=1}^{d}\sum_{k_3=1}^{d}\sum_{k_4=1}^{d}\lambda_{k_1}\lambda_{k_2}\lambda_{k_3}\lambda_{k_4}\int\sigma_{k_1k_3}\left(t_1\right)\sigma_{k_2k_4}\left(t_1\right)K\left(u_1\right)K\left(u_1+u_3\right)K\left(u_2+u_3\right)\\
&\times K\left(u_2\right)\sigma_{k_1k_2}\left(t_1\right)\sigma_{k_3k_4}\left(t_1\right)p^4(t_1)dt_1du_1du_2du_3+o(n^4/h)+O(1/h^4)\\
=&O(n^4/h),
\end{aligned}
$$

where C is a constant.

From these equations, we have

$$\begin{aligned}
&\frac{\sum_{2\leqslant i_1\leqslant i_2\leqslant n}\sum_{1\leqslant j_1<k_1\leqslant i_1-1} E\left\{G_{ni_1}\left(z_{j_1},z_{k_1}\right)G_{ni_2}\left(z_{j_1},z_{k_1}\right)\right\}+n\sum_{i=2}^{n}\sum_{j=1}^{i-1}E\left\{\tilde{H}_n^4\left(z_i,z_j\right)\right\}}{\left[\sum_{i=2}^{n}\sum_{j=1}^{i-1}E\left\{\tilde{H}_n^2\left(z_i,z_j\right)\right\}\right]^2}\\
&=\frac{\mathrm{O}\left(n^4/h\right)+n\mathrm{O}\left(n^2/h^3\right)}{\mathrm{O}\left(n^4/h^2\right)}\\
&=\mathrm{O}\left(h\right)+\mathrm{O}\left(1/(nh)\right)\to 0.
\end{aligned}$$

Since the conditions in Proposition 1 are verified, we then have

$$n^2\lambda^\top V_{1n}/2\left[\sum_{i=2}^{n}\sum_{j=1}^{i-1}E\left\{\tilde{H}_n^2\left(z_i,z_j\right)\right\}\right]^{1/2}\to N(0,1), \text{ in distribution.}$$

This implies that

$$nh^{1/2}\lambda^\top V_{1n}\to N\left(0,2\sum_{k_1=1}^{d}\sum_{k_2=1}^{d}\lambda_{k_1}\lambda_{k_2}\int K^2(u)du\times\int\left\{\sigma_{k1k2}(t)\right\}^2p^2(t)dt\right), \quad \text{in distribution.}$$

The asymptotic variance is actually $\lambda^\top\Sigma\lambda$. The asymptotic normality of V_{1n} is derived.

Following the proof of Lemma 1 in Liu et al. (2021), we can easily show that $nh^{1/2}V_{2n}$ and $nh^{1/2}V_{3n}$ converge to zero in probability. Then we have $nh^{1/2}V_n\to N(0,\Sigma)$ in distribution.

At next, we show that $\hat{\Sigma}$ is a consistent estimator of Σ. This can be proven by using similar statements in proof of Lemma 2 in Liu et al. (2021).

Having the limiting properties of V_n and $\hat{\Sigma}$, the final result is easily derived by using Slutsky's theorem and continuous mapping theorem.

2. Under the global alternatives, V_n can be decomposed as

$$\begin{aligned}
V_n&=\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\frac{1}{h}K\left(\frac{t_i-t_j}{h}\right)e_i\odot e_j\\
&=\left\{\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\frac{1}{h}K\left(\frac{t_i-t_j}{h}\right)\eta_i\odot\eta_j\right\}+o_P(1)\\
&\equiv S_n+o_P(1),
\end{aligned}\tag{C.2}$$

where $\odot$ denotes the element-wise product of two vectors. S_n is a U-statistic with the kernel:

$$H_n\left(z_i,z_j\right)=\frac{1}{h}K\left(\frac{t_i-t_j}{h}\right)\eta_i\odot\eta_j.$$

Since it is non-degenerate, we then apply Proposition 1. Note

$$\begin{aligned}
r_{nj}\left(z_i\right)=&E\left\{H_n(z_i,z_j)\mid z_i\right\}=\frac{1}{h}K\left(\frac{t_i-t_j}{h}\right)\eta_i\odot\left\{X\left(t_j\right)-F\left(t_j,\theta_1^*\right)\right\},\\
\theta_{nij}=&E\left\{r_{nj}\left(z_i\right)\right\}=E\left\{H_n\left(z_i,z_j\right)\right\}\\
=&\frac{1}{h}K\left(\frac{t_i-t_j}{h}\right)\left\{X\left(t_i\right)-F\left(t_i,\theta_1^*\right)\right\}\odot\left\{X\left(t_j\right)-F\left(t_j,\theta_1^*\right)\right\},\\
r_n\left(z_i\right)=&\binom{n-1}{1}^{-1}\sum_{j\neq i}r_{nj}(z_i),\\
\theta_{ni}=&E\left\{r_n\left(z_i\right)\right\},
\end{aligned}$$

$$\begin{aligned}
\theta_n =&\binom{n}{2}^{-1}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\theta_{nij}\\
=&\frac{1}{h}\int K\left(\frac{t_i-t_j}{h}\right)\left\{X(t_i)-F\left(t_i,\theta_1^*\right)\right\}\odot\left\{X\left(t_j\right)-F\left(t_j,\theta_1^*\right)\right\}p(t_i)\,p\left(t_j\right)dt_idt_j\\
=&\frac{1}{h}\int K(u)\left\{X(t_i)-F\left(t_i,\theta_1^*\right)\right\}\odot\left\{X(t_i-hu)-F\left(t_i-hu,\theta_1^*\right)\right\}\\
&\times p(t_i)\,p(t_i-hu)\,dt_ihdu\\
=&\int\left\{X(t)-F\left(t,\theta_1^*\right)\right\}^2\odot p^2(t)dt+o(1)\\
=&E\left[\left\{X(t_i)-F\left(t_i,\theta_1^*\right)\right\}^2\odot p(t_i)\right]+o(1),\\
\hat{U}_n =&\theta_n+\frac{2}{n}\sum_{i=1}^{n}\left\{r_n(z_i)-\theta_{ni}\right\},\\
E(\hat{U}) =&\theta_n.
\end{aligned}$$

The conditions in Proposition 1 can be easily verified, thus we have $\sqrt{n}\left(S_n-\hat{U}_n\right)=o_P(1)$. Since the projection $\hat{U}_n$ is a sample average, by applying the law of large numbers, we have

$$S_n\to E\left[\left\{X(t_i)-F\left(t_i,\theta_1^*\right)\right\}^2\odot p(t_i)\right],\text{ in probability.}$$

V_n also converges to $E\left[\left\{X(t_i)-F\left(t_i,\theta_1^*\right)\right\}^2\odot p(t_i)\right]$ in probability by considering equation (C.2). Similarly we can prove $\hat{\Sigma}$ is a consistent estimator of Σ' under the global alternatives.

The final result is an easily derived consequence of the limiting results of V_n and $\hat{\Sigma}$. □

Appendix D. Proof of Theorem 2 (under local alternatives)

Proof. The proof is similar to the proof for Theorem 2 of Liu et al. (2021). Thus we omit the details. □

References

Altieri, N., Barter, R.L., Duncan, J., Dwivedi, R., Kumbier, K., Li, X., Netzorg, R., Park, B., Singh, C., Tan, Y.S., Tang, T., Wang, Y., Zhang, C., Yu, B., 2021. Special issue 1 - covid-19: unprecedented challenges and chances. In: Harvard Data Science Review.

Bentout, S., Chekroun, A., Kuniya, T., 2020. Parameter estimation and prediction for coronavirus disease outbreak 2019 (covid-19) in Algeria. AIMS Public Health 7, 306–318. https://doi.org/10.3934/publichealth.2020026.

Brown, B.M., 1971. Martingale central limit theorems. Ann. Math. Stat. 42, 59–66. https://doi.org/10.1214/aoms/1177693494.

Brunel, N.J.B., 2008. Parameter estimation of ode's via nonparametric estimators. Electron. J. Stat. 2, 1242–1267. https://doi.org/10.1214/07-EJS132.

Currie, J., Wilson, D.I., 2012. OPTI: lowering the barrier between open source optimizers and the industrial MATLAB user. In: Foundations of Computer-Aided Process Operations. Savannah, Georgia, USA.

Dattner, I., 2015. A model-based initial guess for estimating parameters in systems of ordinary differential equations. Biometrics 71, 1176–1184. https://doi.org/10.1111/biom.12348.

Ding, A.A., Wu, H., 2014. Estimation of ordinary differential equation parameters using constrained local polynomial regression. Stat. Sin. 24, 1613–1631. https://doi.org/10.5705/ss.2012.304.

Fan, Y., Li, Q., 1996. Consistent model specification tests: omitted variables and semiparametric functional forms. Econometrica 64, 865–890.

Fan, Y., Li, Q., 1999. Central limit theorem for degenerate u-statistics of absolutely regular processes with applications to model specification testing. J. Nonparametr. Stat. 10, 245–271.

Fan, Y., Li, Q., 2000. Consistent model specification tests: Kernel-based tests versus bierens' icm tests. Econom. Theory, 1016–1041.

FitzHugh, R., 1961. Impulses and physiological states in theoretical models of nerve membrane. Biophys. J. 1, 445–466. https://doi.org/10.1016/S0006-3495(61)86902-6.

Goel, N.S., Maitra, S.C., Montroll, E.W., 1971. On the Volterra and other nonlinear models of interacting populations. Rev. Mod. Phys. 43, 231–276. https://doi.org/10.1103/RevModPhys.43.231.

Hall, P., 1984. Central limit theorem for integrated square error of multivariate nonparametric density estimators. J. Multivar. Anal. 14, 1–16. https://doi.org/10.1016/0047-259X(84)90044-7.

Hall, P., Heyde, Christopher C., 2014. Martingale Limit Theory and Its Application. Academic Press.

Härdle, W., Mammen, E., 1993. Comparing nonparametric versus parametric regression fits. Ann. Stat. 21, 1926–1947. https://doi.org/10.1214/aos/1176349403.

Hirsch, M.W., Smale, S., Devaney, R.L., 2013. Differential Equations, Dynamical Systems, and an Introduction to Chaos, 3rd ed. Academic Press, Waltham, MA.

Hoeffding, W., 1948. A class of statistics with asymptotically normal distribution. Ann. Inst. Stat. Math. 19, 293–325.

Hooker, G., 2009. Forcing function diagnostics for nonlinear dynamics. Biometrics 65, 928–936. https://doi.org/10.1111/j.1541-0420.2008.01172.x.

Khmaladze, E.V., Koul, H.L., 2004. Martingale transforms goodness-of-fit tests in regression models. Ann. Stat. 32, 995–1034. https://doi.org/10.1214/009053604000000274.

Koul, H.L., Ni, P., 2004. Minimum distance regression model checking. J. Stat. Plan. Inference 119, 109–141. https://doi.org/10.1016/S0378-3758(02)00415-9.

Kuniya, T., 2020. Prediction of the epidemic peak of coronavirus disease in Japan. J. Clin. Med. 9, 789. https://doi.org/10.3390/jcm9030789.

Lavergne, P., Patilea, V., 2012. One for all and all for one: regression checks with many regressors. J. Bus. Econ. Stat. 30, 41–52. https://doi.org/10.1198/jbes.2011.07152.

Lee, A.J., 1990. U-Statistics: Theory and Practice. M. Dekker, New York.

Lin, Q.S., Hu, T.J., Zhou, X.H., 2020. Estimating the daily trend in the size of the covid-19 infected population in Wuhan. Infect. Dis. Poverty 9, 69. https://doi.org/10.1186/s40249-020-00693-4.

Liu, R., Fang, Y., Zhu, L., 2021. Model checking for parametric ordinary differential equations systems. Stat. Sin. https://doi.org/10.5705/ss.202020.0429, in press.

Lotka, A.J., 1910. Contribution to the theory of periodic reactions. J. Phys. Chem. 14, 271–274. https://doi.org/10.1021/j150111a004.

Nagumo, J., Arimoto, S., Yoshizawa, S., 1962. An active pulse transmission line simulating nerve axon. Proc. IRE 50, 2061–2070. https://doi.org/10.1109/JRPROC.1962.288235.

Powell, J.L., Stock, J.H., Stoker, T.M., 1989. Semiparametric estimation of index coefficients. Econometrica 57, 1403. https://doi.org/10.2307/1913713.

Stute, W., 1997. Nonparametric model checks for regression. Ann. Stat. 25, 613–641. https://doi.org/10.1214/aos/1031833666.

Stute, W., Manteiga, W.G., Quindimil, M.P., 1998a. Bootstrap approximations in model checks for regression. J. Am. Stat. Assoc. 93, 141–149. https://doi.org/10.1080/01621459.1998.10474096.

Stute, W., Thies, S., Zhu, L.X., 1998b. Model checks for regression: an innovation process approach. Ann. Stat. 26, 1916–1934. https://doi.org/10.1214/aos/1024691363.

Tian, H., Liu, Y., Li, Y., Wu, C.H., Chen, B., Kraemer, M.U.G., Li, B., Cai, J., Xu, B., Yang, Q., Wang, B., Yang, P., Cui, Y., Song, Y., Zheng, P., Wang, Q., Bjornstad, O.N., Yang, R., Grenfell, B.T., Pybus, O.G., Dye, C., 2020. An Investigation of Transmission Control Measures During the First 50 Days of the Covid-19 Epidemic in China, vol. 368. Science, New York, N.Y., pp. 638–642.

Volterra, V., 1928. Variations and fluctuations of the number of individuals in animal species living together. ICES J. Mar. Sci. 3, 3–51. https://doi.org/10.1093/icesjms/3.1.3.

White, H., 1981. Consequences and detection of misspecified nonlinear regression models. J. Am. Stat. Assoc. 76, 419. https://doi.org/10.2307/2287845.

Xue, H., Miao, H., Wu, H., 2010. Sieve estimation of constant and time-varying coefficients in nonlinear ordinary differential equation models by considering both numerical error and measurement error. Ann. Stat. 38, 2351–2387. https://doi.org/10.1214/09-AOS784.

Yang, Z., Zeng, Z., Wang, K., Wong, S.S., Liang, W., Zanin, M., Liu, P., Cao, X., Gao, Z., Mai, Z., Liang, J., Liu, X., Li, S., Li, Y., Ye, F., Guan, W., Yang, Y., Li, F., Luo, S., Xie, Y., Liu, B., Wang, Z., Zhang, S., Wang, Y., Zhong, N., He, J., 2020. Modified SEIR and AI prediction of the epidemics trend of covid-19 in China under public health interventions. J. Thorac. Dis. 12, 165–174. https://doi.org/10.21037/jtd.2020.02.64.

Zheng, J.X., 1996. A consistent test of functional form via nonparametric estimation techniques. J. Econom. 75, 263–289. https://doi.org/10.1016/0304-4076(95)01760-7.

Zhu, L.X., 2003. Model checking of dimension-reduction type for regression. Stat. Sin. 13, 283–296.