Supplement to "Ordinary Differential Equation Models for a Collection of Discretized Functions"

Lingxuan Shao

Department of Statistics and Data Science, School of Management, Fudan University, Shanghai, China

E-mail: shao_lingxuan@fudan.edu.cn

Fang Yao

Department of Probability and Statistics, School of Mathematical Sciences, Peking University, Beijing, China

E-mail: fyao@math.pku.edu.cn

Summary. This Supplementary Material provides the additional simulations and technical proofs of Theorems 1, 2 and 3 in the main paper.

S1. Additional Simulation and Real Data Results

S1.1. More pairs of (n,m)

We extend our analysis of the proposed method with additional simulations involving larger m (i.e., denser sampling) as a supplement to Section 5. Specifically, we have incorporated the following pairs (n, m) where $n \in \{30, 100, 200\}$ and $m \in \{15, 25, 35, 45\}$. The results of parameter estimation and recovery, presented in Tables S1 and S2, alongside Tables 1 and 3 in the main article, illustrate that our proposed method consistently delivers accurate estimates and recoveries across a diverse spectrum of ODE models, distribution types, and sampling frequencies. The empirical errors $\mathbb{E}\,|\widehat{\beta} - \beta|^2/|\beta|^2$ and $\frac{1}{n}\sum_{i=1}^{n}\int_0^1 \left(\widehat{X}_i(t) - X_i(t)\right)^2 dt$ diminish as either the sample size n or the sampling frequency m increases, transitioning from a sparse design ($m = 1$) to a dense design (where

m is on the same scale as n).

S1.2. Varying sampling frequency m

We conducted additional simulations of our proposed method with varying sampling frequencies m as an extension of Section 5. Specifically, we considered $n \in \{100, 200\}$ and random sampling frequencies m satisfying (a) Uniform$(5, 10)$, (b) Uniform$(10, 15)$, (c) Poisson(4), and (d) Poisson(8). The outcomes of parameter estimation and recovery are presented in Tables S3 and S4, respectively. These results indicate that our method maintains satisfactory performance under varying m. The empirical errors $\mathbb{E}\,|\widehat{\beta}-\beta|^2/|\beta|^2$ and $\frac{1}{n}\sum_{i=1}^{n}\int_0^1 \big(\widehat{X}_i(t) - X_i(t)\big)^2 dt$ decrease as either the sample size n or the sampling frequency m increases, i.e., from (a) Uniform$(5, 10)$ to (b) Uniform$(10, 15)$, or from (c) Poisson(4) to (d) Poisson(8).

S1.3. Examining tuning parameter κ

This section presents the selected κ using K-fold cross-validation in our recovery procedure, along with the first and second terms in Equation (10) across all realizations in Section 3.2, which are

$$\frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\big(\widehat{X}_i(T_{ij}) - X_{ij}\big)^2 \quad \text{and} \quad \frac{\kappa}{n|\mathcal{T}|}\sum_{i=1}^{n}\int_{\mathcal{T}}\big(F_{\widehat{\beta}}(\widehat{X}_i)\big)^2 dt,$$

respectively. The ratio of these terms, specifically the second term divided by the first term, is expressed as

$$\text{ratio} = \frac{\frac{\kappa}{n|\mathcal{T}|}\sum_{i=1}^{n}\int_{\mathcal{T}}\big(F_{\widehat{\beta}}(\widehat{X}_i)\big)^2 dt}{\frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\big(\widehat{X}_i(T_{ij}) - X_{ij}\big)^2}.$$

Table S5 presents the results in Model (ii) as an illustrative example. The table includes the mean square error $\frac{1}{n}\sum_{i=1}^{n}\int_0^1 \big(\widehat{X}_i(t) - X_i(t)\big)^2 dt$, the selected κ, the first term, the second term, and their ratio for $n \in \{100, 200\}$ and $m \in \{5, 10, 15, 20\}$. The results indicate that the selected κ achieves a balance between the first and second terms, and that the ratio decreases as the sampling frequency m increases. This trend is anticipated, as a larger m implies that the data fitting term already provides a well-performed curve estimate, thus reducing the necessity for the influence from the ODE conformality term.

S1.4. Recovered trajectories for Covid-19 data and AIDS data

This section presents the plots of recovered trajectories in Sections 6.1 and 6.2.

Figure S1 illustrates the recovered infection counts of Covid-19 data across all $n = 33$ provinces, complementing Section 6.1. These visualizations reveal that the recovered trajectories closely approximate the original data, suggesting that our proposed methodology successfully recovers all trajectories even under sparse sampling designs. Consequently, our proposed methodology exhibits commendable performance under conditions of sparse sampling frequencies, which is attributed to the proposed Functional Moment Method that can effectively aggregate measurements from various functions. Considering the resource-intensive nature of Covid-19 data collection and the associated risks of reinfection during nucleic acid testing, the implementation of a more sparse sampling frequency with our methodology can conserve resources and mitigate reinfection risks while still providing reliable estimates.

Figure S2 illustrates the recovered AIDS infection counts across all $n = 31$ provinces, complementing the analysis in Section 6.2. These recovered trajectories fill the gaps for the remaining $366 - 12 = 354$ days. The trajectories display a clear upward trend consistent with the SIRD ODE model and closely align with the sparsely observed data points. These findings suggest that our proposed method is effective in reconstructing the spread process over the entire time interval.

Table S1. Relative mean square errors $\mathbb{E}\,|\widehat{\beta}-\beta|^2/|\beta|^2$, assessed by 100 Mento Carlo runs with standard errors in the parentheses.

Model (i)	n	$m=15$	$m=25$	$m=35$	$m=45$
Uniform	30	8.1e-03 (1.9e-03)	5.7e-03 (1.1e-03)	3.1e-03 (6.4e-04)	2.5e-03 (5.6e-04)
	100	3.1e-03 (7.0e-04)	1.2e-03 (2.4e-04)	8.1e-04 (2.1e-04)	6.7e-04 (1.4e-04)
	200	1.3e-03 (2.5e-04)	5.7e-04 (1.3e-04)	3.9e-04 (5.5e-05)	5.3e-04 (1.2e-04)
Normal	30	2.7e-03 (5.7e-04)	1.6e-03 (4.2e-04)	1.0e-03 (2.0e-04)	9.4e-04 (2.8e-04)
	100	6.6e-04 (1.4e-04)	3.5e-04 (5.5e-05)	3.3e-04 (5.5e-05)	2.6e-04 (6.1e-05)
	200	3.5e-04 (1.1e-04)	2.0e-04 (4.1e-05)	1.4e-04 (1.5e-05)	1.7e-04 (3.0e-05)
LogNormal	30	1.3e-03 (2.6e-04)	9.8e-04 (2.3e-04)	7.0e-04 (1.4e-04)	7.8e-04 (1.3e-04)
	100	4.3e-04 (6.4e-05)	4.7e-04 (9.7e-05)	3.6e-04 (1.1e-04)	2.1e-04 (5.2e-05)
	200	2.6e-04 (4.4e-05)	1.8e-04 (3.9e-05)	1.8e-04 (4.8e-05)	1.1e-04 (1.7e-05)
Model (ii)	n	$m=15$	$m=25$	$m=35$	$m=45$
Uniform	30	5.9e-03 (8.2e-04)	3.0e-03 (6.0e-04)	1.2e-03 (1.9e-04)	1.3e-03 (1.9e-04)
	100	1.8e-03 (3.0e-04)	1.1e-03 (2.0e-04)	6.2e-04 (1.0e-04)	3.9e-04 (5.7e-05)
	200	8.9e-04 (1.7e-04)	5.5e-04 (9.2e-05)	3.6e-04 (7.0e-05)	1.3e-04 (2.6e-05)
Normal	30	8.7e-04 (1.6e-04)	5.8e-04 (1.5e-04)	2.2e-04 (3.2e-05)	2.2e-04 (3.6e-05)
	100	3.1e-04 (6.7e-05)	1.3e-04 (2.1e-05)	1.2e-04 (1.8e-05)	6.0e-05 (9.3e-06)
	200	2.0e-04 (3.3e-05)	7.1e-05 (1.0e-05)	4.8e-05 (7.2e-06)	3.0e-05 (4.4e-06)
LogNormal	30	1.8e-03 (3.3e-04)	9.7e-04 (1.6e-04)	4.9e-04 (1.1e-04)	4.4e-04 (8.1e-05)
	100	7.8e-04 (1.4e-04)	2.9e-04 (5.1e-05)	2.1e-04 (3.7e-05)	1.1e-04 (2.0e-05)
	200	2.9e-04 (4.8e-05)	1.6e-04 (3.0e-05)	1.4e-04 (2.6e-05)	5.0e-05 (6.9e-06)
Model (iii)	n	$m=15$	$m=25$	$m=35$	$m=45$
Uniform	30	1.2e-01 (1.8e-02)	6.7e-02 (1.2e-02)	5.8e-02 (9.8e-03)	4.3e-02 (8.9e-03)
	100	4.5e-02 (7.2e-03)	2.4e-02 (4.2e-03)	1.4e-02 (2.5e-03)	8.6e-03 (1.1e-03)
	200	1.9e-02 (3.2e-03)	1.7e-02 (2.6e-03)	5.7e-03 (6.9e-04)	8.4e-03 (1.1e-03)
Normal	30	5.7e-02 (1.1e-02)	2.3e-02 (3.4e-03)	1.8e-02 (2.3e-03)	1.7e-02 (2.5e-03)
	100	1.7e-02 (3.4e-03)	8.4e-03 (1.4e-03)	6.3e-03 (1.0e-03)	7.6e-03 (1.3e-03)
	200	5.6e-03 (7.9e-04)	3.5e-03 (5.7e-04)	3.7e-03 (6.2e-04)	2.1e-03 (3.3e-04)
LogNormal	30	3.8e-02 (4.6e-03)	2.6e-02 (3.9e-03)	1.3e-02 (1.8e-03)	1.4e-02 (2.6e-03)
	100	1.1e-02 (1.7e-03)	5.6e-03 (1.0e-03)	5.3e-03 (8.3e-04)	3.6e-03 (5.4e-04)
	200	5.2e-03 (9.3e-04)	3.1e-03 (4.1e-04)	3.1e-03 (5.1e-04)	1.6e-03 (2.8e-04)

Table S2. Mean square errors $\frac{1}{n}\sum_{i=1}^{n}\int_0^1 \left(\widehat{X}_i(t) - X_i(t)\right)^2 dt$, assessed by 100 Mento Carlo runs with standard errors in the parentheses.

Model (i)	n	$m = 15$	$m = 25$	$m = 35$	$m = 45$
	30	5.3e-03 (4.2e-04)	3.0e-03 (2.1e-04)	2.2e-03 (1.2e-04)	2.1e-03 (1.4e-04)
Uniform	100	2.5e-03 (1.1e-04)	1.5e-03 (7.7e-05)	1.3e-03 (6.6e-05)	1.0e-03 (5.2e-05)
	200	1.9e-03 (8.5e-05)	1.2e-03 (4.6e-05)	1.0e-03 (5.3e-05)	8.9e-04 (4.1e-05)
	30	1.5e-03 (6.0e-05)	9.8e-04 (4.3e-05)	7.5e-04 (3.4e-05)	6.5e-04 (3.3e-05)
Normal	100	1.1e-03 (3.2e-05)	6.5e-04 (1.9e-05)	5.1e-04 (1.7e-05)	4.5e-04 (2.0e-05)
	200	9.0e-04 (1.6e-05)	5.8e-04 (1.3e-05)	4.4e-04 (9.9e-06)	3.5e-04 (9.0e-06)
	30	3.0e-03 (2.1e-04)	1.9e-03 (1.0e-04)	1.6e-03 (9.1e-05)	1.3e-03 (8.9e-05)
LogNormal	100	1.6e-03 (6.9e-05)	1.2e-03 (6.7e-05)	9.4e-04 (5.2e-05)	7.3e-04 (4.3e-05)
	200	1.3e-03 (4.4e-05)	9.1e-04 (4.0e-05)	7.1e-04 (3.4e-05)	5.6e-04 (2.5e-05)
Model (ii)	n	$m = 15$	$m = 25$	$m = 35$	$m = 45$
	30	1.5e-02 (4.8e-04)	1.2e-02 (2.9e-04)	1.3e-02 (3.1e-04)	1.2e-02 (3.1e-04)
Uniform	100	1.2e-02 (2.4e-04)	1.2e-02 (1.9e-04)	1.2e-02 (1.9e-04)	1.1e-02 (2.0e-04)
	200	1.1e-02 (1.4e-04)	1.1e-02 (1.7e-04)	1.1e-02 (1.3e-04)	1.1e-02 (1.1e-04)
	30	1.0e-02 (3.9e-04)	9.7e-03 (3.2e-04)	9.5e-03 (3.5e-04)	9.4e-03 (2.7e-04)
Normal	100	9.6e-03 (2.7e-04)	9.3e-03 (2.3e-04)	9.2e-03 (2.1e-04)	9.1e-03 (1.9e-04)
	200	9.5e-03 (2.9e-04)	9.2e-03 (2.5e-04)	9.0e-03 (1.8e-04)	9.0e-03 (1.5e-04)
	30	1.3e-02 (4.9e-04)	1.2e-02 (4.4e-04)	1.3e-02 (4.6e-04)	1.2e-02 (4.0e-04)
LogNormal	100	1.2e-02 (3.1e-04)	1.2e-02 (2.7e-04)	1.2e-02 (2.6e-04)	1.1e-02 (2.5e-04)
	200	1.2e-02 (2.8e-04)	1.2e-02 (2.0e-04)	1.1e-02 (2.5e-04)	1.1e-02 (2.0e-04)
Model (iii)	n	$m = 15$	$m = 25$	$m = 35$	$m = 45$
	30	3.0e-03 (1.3e-04)	1.9e-03 (9.6e-05)	1.3e-03 (5.5e-05)	9.8e-04 (4.3e-05)
Uniform	100	1.4e-03 (3.6e-05)	8.5e-04 (2.3e-05)	6.4e-04 (2.0e-05)	4.7e-04 (1.2e-05)
	200	1.1e-03 (2.4e-05)	6.6e-04 (1.3e-05)	4.7e-04 (8.4e-06)	3.9e-04 (6.2e-06)
	30	1.2e-03 (4.0e-05)	7.2e-04 (2.1e-05)	5.1e-04 (1.8e-05)	3.9e-04 (1.3e-05)
Normal	100	8.4e-04 (1.4e-05)	5.0e-04 (9.0e-06)	3.7e-04 (6.3e-06)	2.8e-04 (4.9e-06)
	200	7.7e-04 (8.2e-06)	4.7e-04 (5.1e-06)	3.3e-04 (3.8e-06)	2.6e-04 (2.7e-06)
	30	1.9e-03 (8.0e-05)	1.1e-03 (5.1e-05)	8.8e-04 (4.9e-05)	6.7e-04 (3.3e-05)
LogNormal	100	1.1e-03 (2.7e-05)	6.5e-04 (1.7e-05)	4.8e-04 (1.0e-05)	3.7e-04 (1.0e-05)
	200	8.8e-04 (1.4e-05)	5.5e-04 (8.9e-06)	3.9e-04 (7.7e-06)	3.1e-04 (4.9e-06)

Table S3. Relative mean square errors $\mathbb{E}\,|\widehat{\beta}-\beta|^2/|\beta|^2$ for varying sampling frequency, assessed by 100 Mento Carlo runs with standard errors in the parentheses.

Model (i)	n	Uniform(5, 10)	Uniform(10, 15)	Poisson(4)	Poisson(8)
Uniform	100	4.3e-03 (7.3e-04)	4.5e-03 (1.0e-03)	6.4e-03 (1.7e-03)	6.4e-03 (1.3e-03)
	200	2.2e-03 (3.4e-04)	1.1e-03 (2.2e-04)	4.0e-03 (1.0e-03)	1.9e-03 (3.1e-04)
Normal	100	2.5e-03 (4.8e-04)	7.4e-04 (1.1e-04)	2.5e-03 (4.7e-04)	1.5e-03 (2.7e-04)
	200	6.8e-04 (1.4e-04)	5.2e-04 (8.8e-05)	1.3e-03 (3.1e-04)	7.3e-04 (1.5e-04)
LogNormal	100	1.9e-03 (4.8e-04)	7.4e-04 (1.3e-04)	1.2e-03 (2.2e-04)	1.1e-03 (1.4e-04)
	200	7.4e-04 (1.7e-04)	4.0e-04 (8.0e-05)	6.9e-04 (1.2e-04)	6.1e-04 (1.4e-04)
Model (ii)	n	Uniform(5, 15)	Uniform(15, 25)	Poisson(5)	Poisson(10)
Uniform	100	8.6e-03 (1.4e-03)	2.4e-03 (3.8e-04)	2.2e-02 (4.2e-03)	8.2e-03 (1.4e-03)
	200	5.7e-03 (1.4e-03)	1.1e-03 (2.0e-04)	1.1e-02 (2.6e-03)	3.9e-03 (7.2e-04)
Normal	100	1.0e-03 (1.5e-04)	7.3e-04 (2.2e-04)	2.9e-03 (5.2e-04)	1.1e-03 (2.2e-04)
	200	9.4e-04 (2.1e-04)	3.1e-04 (6.1e-05)	1.0e-03 (2.0e-04)	4.0e-04 (8.0e-05)
LogNormal	100	3.3e-03 (5.0e-04)	9.5e-04 (2.1e-04)	9.1e-03 (2.1e-03)	4.0e-03 (1.2e-03)
	200	1.2e-03 (2.7e-04)	6.9e-04 (1.3e-04)	4.4e-03 (1.0e-03)	9.1e-04 (1.8e-04)
Model (iii)	n	Uniform(5, 15)	Uniform(15, 25)	Poisson(5)	Poisson(10)
Uniform	100	6.9e-02 (1.1e-02)	4.7e-02 (5.3e-03)	1.4e-01 (2.3e-02)	6.5e-02 (9.4e-03)
	200	5.0e-02 (1.0e-02)	2.2e-02 (4.4e-03)	9.5e-02 (1.8e-02)	3.9e-02 (8.4e-03)
Normal	100	4.1e-02 (6.6e-03)	2.2e-02 (3.4e-03)	5.7e-02 (1.0e-02)	3.2e-02 (6.0e-03)
	200	1.9e-02 (3.8e-03)	9.3e-03 (1.6e-03)	1.8e-02 (3.0e-03)	1.7e-02 (2.3e-03)
LogNormal	100	3.2e-02 (6.7e-03)	1.3e-02 (2.2e-03)	3.3e-02 (5.3e-03)	2.1e-02 (3.5e-03)
	200	1.5e-02 (2.1e-03)	6.6e-03 (9.1e-04)	2.5e-02 (5.0e-03)	1.0e-02 (1.5e-03)

Table S4. Mean square errors $\frac{1}{n}\sum_{i=1}^{n}\int_0^1 \left(\widehat{X}_i(t) - X_i(t)\right)^2 dt$ for varying sampling frequency, assessed by 100 Mento Carlo runs with standard errors in the parentheses.

Model (i)	n	Uniform(5, 10)	Uniform(10, 15)	Poisson(4)	Poisson(8)
Uniform	100	5.4e-03 (2.5e-04)	2.7e-03 (9.6e-05)	7.6e-03 (3.0e-04)	4.2e-03 (1.7e-04)
	200	3.9e-03 (1.7e-04)	2.3e-03 (8.7e-05)	5.9e-03 (2.0e-04)	3.3e-03 (1.2e-04)
Normal	100	2.4e-03 (5.6e-05)	1.4e-03 (3.5e-05)	4.0e-03 (1.0e-04)	2.1e-03 (5.4e-05)
	200	2.2e-03 (4.7e-05)	1.1e-03 (1.6e-05)	3.4e-03 (6.9e-05)	1.9e-03 (3.3e-05)
LogNormal	100	3.6e-03 (1.4e-04)	1.9e-03 (6.8e-05)	5.3e-03 (2.0e-04)	3.0e-03 (1.3e-04)
	200	3.0e-03 (9.9e-05)	1.6e-03 (5.0e-05)	4.7e-03 (1.6e-04)	2.5e-03 (9.0e-05)
Model (ii)	n	Uniform(5, 15)	Uniform(15, 25)	Poisson(5)	Poisson(10)
Uniform	100	1.7e-02 (7.5e-04)	1.3e-02 (5.1e-04)	1.9e-02 (7.0e-04)	1.6e-02 (6.3e-04)
	200	1.6e-02 (9.1e-04)	1.3e-02 (2.9e-04)	2.0e-02 (2.8e-03)	1.5e-02 (6.4e-04)
Normal	100	1.1e-02 (5.6e-04)	9.9e-03 (4.0e-04)	1.3e-02 (8.9e-04)	1.1e-02 (4.2e-04)
	200	1.2e-02 (6.0e-04)	1.0e-02 (3.5e-04)	1.2e-02 (5.5e-04)	1.1e-02 (3.9e-04)
LogNormal	100	1.5e-02 (7.9e-04)	1.3e-02 (4.6e-04)	2.1e-02 (3.3e-03)	1.5e-02 (5.6e-04)
	200	1.4e-02 (5.6e-04)	1.3e-02 (3.4e-04)	1.6e-02 (5.4e-04)	1.4e-02 (6.8e-04)
Model (iii)	n	Uniform(5, 15)	Uniform(15, 25)	Poisson(5)	Poisson(10)
Uniform	100	3.7e-03 (1.3e-04)	1.8e-03 (5.3e-05)	6.1e-03 (2.2e-04)	2.8e-03 (8.3e-05)
	200	2.6e-03 (5.2e-05)	1.4e-03 (2.6e-05)	4.3e-03 (1.1e-04)	2.3e-03 (4.2e-05)
Normal	100	2.1e-03 (3.5e-05)	1.1e-03 (2.1e-05)	3.4e-03 (8.6e-05)	1.8e-03 (3.2e-05)
	200	1.9e-03 (2.3e-05)	1.0e-03 (1.3e-05)	3.0e-03 (4.1e-05)	1.6e-03 (2.0e-05)
LogNormal	100	2.8e-03 (9.3e-05)	1.5e-03 (4.1e-05)	4.3e-03 (1.3e-04)	2.4e-03 (9.1e-05)
	200	2.3e-03 (4.5e-05)	1.2e-03 (2.2e-05)	3.6e-03 (7.2e-05)	1.9e-03 (3.3e-05)

Table S5. The mean square error $\frac{1}{n}\sum_{i=1}^{n}\int_0^1\left(\widehat{X}_i(t)-X_i(t)\right)^2dt$, the selected κ, the first term $\frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\left(\widehat{X}_i(T_{ij})-X_{ij}\right)^2$, the second term $\frac{\kappa}{n|\mathcal{T}|}\sum_{i=1}^{n}\int_{\mathcal{T}}\left(F_{\widehat{\beta}}(\widehat{X}_i)\right)^2dt$, and the ratio under Model (ii), evaluated over 100 Monte Carlo runs with standard errors in parentheses.

Uniform	MSE	selected κ	first item	second item	ratio
(100, 5)	1.6e-02 (5.5e-04)	2.2e-03 (3.9e-04)	2.1e-02 (4.9e-04)	5.4e-02 (1.9e-02)	2.42 (0.69)
(100, 10)	1.4e-02 (6.3e-04)	4.4e-03 (4.7e-04)	2.2e-02 (6.3e-04)	4.0e-02 (8.5e-03)	1.82 (0.37)
(100, 15)	1.3e-02 (5.4e-04)	4.5e-03 (4.8e-04)	2.1e-02 (5.4e-04)	2.5e-02 (3.7e-03)	1.24 (0.19)
(100, 20)	1.2e-02 (2.0e-04)	4.0e-03 (4.7e-04)	2.1e-02 (2.2e-04)	1.8e-02 (2.7e-03)	0.87 (0.13)
(200, 5)	1.6e-02 (5.1e-04)	3.0e-03 (4.3e-04)	2.2e-02 (4.6e-04)	3.1e-02 (6.6e-03)	1.42 (0.23)
(200, 10)	1.3e-02 (4.0e-04)	4.7e-03 (4.8e-04)	2.1e-02 (3.9e-04)	2.7e-02 (3.8e-03)	1.30 (0.17)
(200, 15)	1.2e-02 (2.1e-04)	3.7e-03 (4.7e-04)	2.1e-02 (2.0e-04)	1.6e-02 (2.4e-03)	0.86 (0.14)
(200, 20)	1.2e-02 (1.6e-04)	4.2e-03 (4.9e-04)	2.1e-02 (1.5e-04)	1.5e-02 (2.0e-03)	0.75 (0.11)
Normal	MSE	selected κ	first item	second item	ratio
(100, 5)	1.2e-02 (5.6e-04)	4.9e-03 (4.7e-04)	1.8e-02 (5.5e-04)	2.5e-02 (4.0e-03)	1.74 (0.32)
(100, 10)	9.6e-03 (3.6e-04)	4.8e-03 (4.9e-04)	1.8e-02 (3.5e-04)	1.5e-02 (1.7e-03)	1.01 (0.13)
(100, 15)	9.0e-03 (3.2e-04)	5.5e-03 (4.9e-04)	1.8e-02 (3.2e-04)	1.6e-02 (1.6e-03)	1.09 (0.12)
(100, 20)	9.9e-03 (2.1e-04)	4.6e-03 (4.9e-04)	1.9e-02 (2.1e-04)	1.2e-02 (1.4e-03)	0.70 (0.09)
(200, 5)	1.1e-02 (3.2e-04)	5.2e-03 (4.7e-04)	1.7e-02 (3.0e-04)	1.9e-02 (2.2e-03)	1.36 (0.20)
(200, 10)	9.9e-03 (2.6e-04)	5.5e-03 (4.9e-04)	1.8e-02 (2.5e-04)	1.7e-02 (1.6e-03)	1.03 (0.11)
(200, 15)	9.4e-03 (2.8e-04)	5.3e-03 (4.9e-04)	1.8e-02 (2.7e-04)	1.5e-02 (1.5e-03)	0.93 (0.10)
(200, 20)	9.7e-03 (2.1e-04)	4.9e-03 (5.0e-04)	1.9e-02 (2.1e-04)	1.2e-02 (1.4e-03)	0.71 (0.08)
LogNormal	MSE	selected κ	first item	second item	ratio
(100, 5)	1.5e-02 (5.7e-04)	3.4e-03 (4.2e-04)	2.0e-02 (5.2e-04)	2.5e-02 (3.2e-03)	1.40 (0.20)
(100, 10)	1.3e-02 (4.3e-04)	4.5e-03 (4.6e-04)	2.0e-02 (4.3e-04)	2.4e-02 (3.2e-03)	1.35 (0.18)
(100, 15)	1.2e-02 (3.2e-04)	4.6e-03 (4.7e-04)	2.0e-02 (3.2e-04)	1.7e-02 (2.0e-03)	0.98 (0.13)
(100, 20)	1.2e-02 (3.5e-04)	4.9e-03 (4.9e-04)	2.1e-02 (3.5e-04)	1.9e-02 (2.3e-03)	0.97 (0.14)
(200, 5)	1.4e-02 (4.3e-04)	4.7e-03 (4.7e-04)	2.0e-02 (4.2e-04)	3.0e-02 (4.0e-03)	1.85 (0.27)
(200, 10)	1.2e-02 (3.2e-04)	4.5e-03 (4.9e-04)	2.0e-02 (3.1e-04)	1.9e-02 (2.4e-03)	1.13 (0.16)
(200, 15)	1.2e-02 (2.9e-04)	4.4e-03 (4.9e-04)	2.0e-02 (2.8e-04)	1.8e-02 (2.3e-03)	1.01 (0.14)
(200, 20)	1.2e-02 (2.5e-04)	4.2e-03 (4.9e-04)	2.0e-02 (2.5e-04)	1.5e-02 (1.9e-03)	0.85 (0.12)

Fig. S1. The figures depict the Covid-19 infection counts across all $n = 33$ provinces. Each figure corresponds to a single province with the province name below. In the i-th figure, the red and blue dotted lines represent the recovered trajectories $\widehat{I}_i{}'$ of infection counts derived from the sparsified data under $m' = 20, 30$, whereas the black solid line denotes the infection counts I_i obtained from the full dataset.

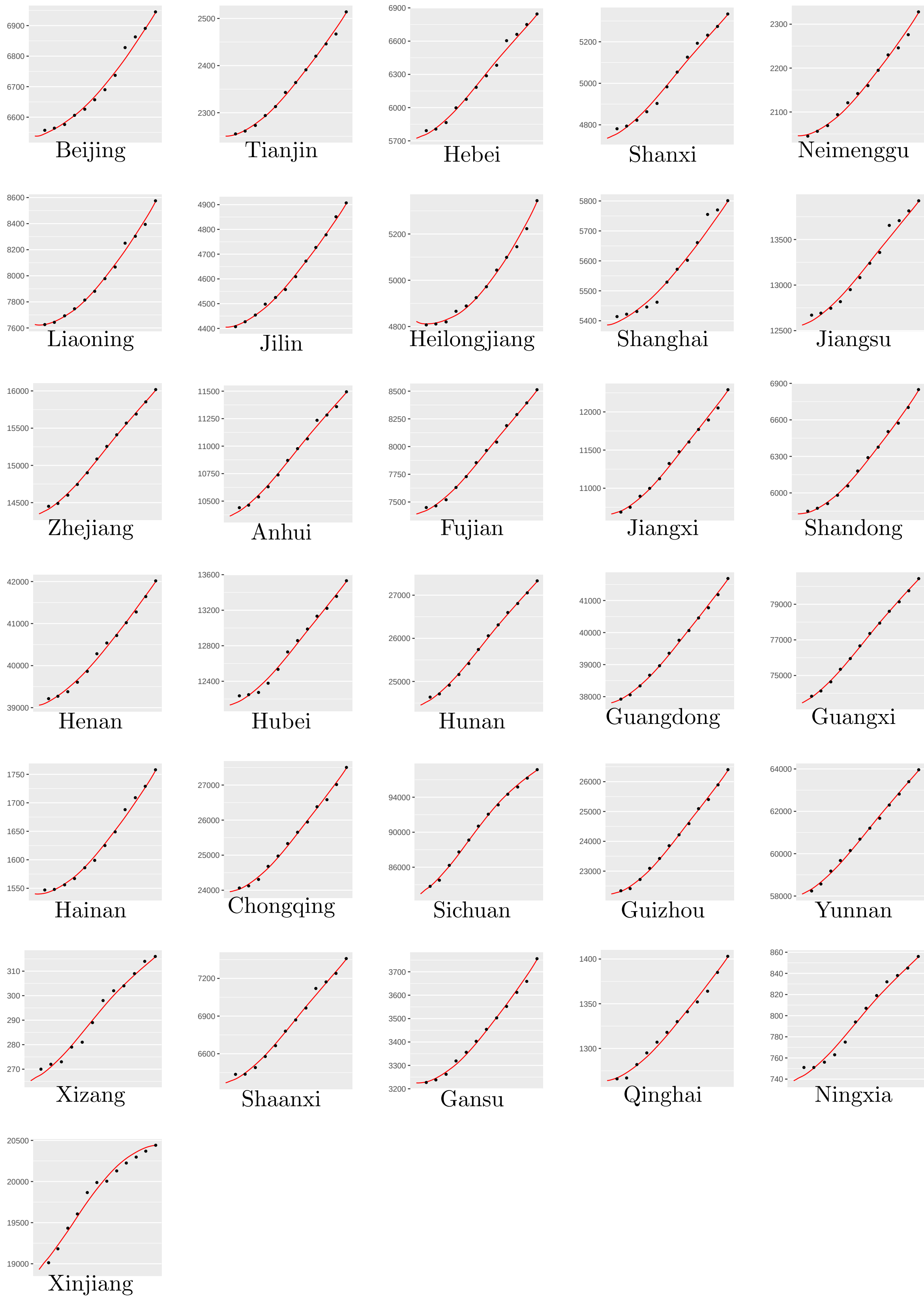

Fig. S2. These figures illustrate the AIDS infection counts across all $n = 31$ provinces. Each figure corresponds to a specific province, with the province name labeled below it. The red solid lines represent the recovered trajectories of infection counts, while the black dots indicate the sparsely observed data points.

S2. Proofs of Theorems

PROOF (PROOF OF THEOREM 1).

Step 1: We present the proof for $q = 1$ and $\alpha = (1, 0)$ for simplicity of notation; the extension to general case is technically straightforward. The estimate $\widehat{\mathbb{E}}\, X^{(\alpha)}(t)$ in (8) shares the following explicit expression

$$\widehat{\mathbb{E}}\, X^{(\alpha)}(t) = \widehat{w}_{10}\widehat{R}_{00} + \widehat{w}_{20}\widehat{R}_{10} + \widehat{w}_{11}\widehat{R}_{01},$$

where the mid-terms are

$$\widehat{R}_{ab} := \tfrac{1}{nm(m-1)} \textstyle\sum_{i,j_1\neq j_2} K_h(T_{ij_1} - t)K_h(T_{ij_2} - t) \times (T_{ij_1} - t)^a (T_{ij_2} - t)^b X_{ij_1} X_{ij_2},$$

$$\widehat{S}_{ab} := \tfrac{1}{nm(m-1)} \textstyle\sum_{i,j_1\neq j_2} K_h(T_{ij_1} - t)K_h(T_{ij_2} - t) \times (T_{ij_1} - t)^a (T_{ij_2} - t)^b,$$

$$\widehat{D} := \widehat{S}_{00}(\widehat{S}_{20}\widehat{S}_{02} - \widehat{S}_{11}\widehat{S}_{11}) + \widehat{S}_{10}(\widehat{S}_{01}\widehat{S}_{11} - \widehat{S}_{10}\widehat{S}_{02}) + \widehat{S}_{01}(\widehat{S}_{10}\widehat{S}_{11} - \widehat{S}_{01}\widehat{S}_{20}),$$

$$\widehat{w}_{10} := (\widehat{S}_{11}\widehat{S}_{01} - \widehat{S}_{10}\widehat{S}_{02})\widehat{D}^{-1}, \quad \widehat{w}_{20} := (\widehat{S}_{00}\widehat{S}_{02} - \widehat{S}_{01}\widehat{S}_{01})\widehat{D}^{-1}, \quad \widehat{w}_{11} := (\widehat{S}_{10}\widehat{S}_{01} - \widehat{S}_{00}\widehat{S}_{11})\widehat{D}^{-1}.$$

Define the collection of time points $\mathbb{T} = \{T_{ij} : 1 \leq i \leq n, 1 \leq j \leq m\}$. Consider the decomposition of $\widehat{\mathbb{E}}\, X^{(\alpha)} - \mathbb{E}\, X^{(\alpha)}$ into the deviation part and the bias part

$$\begin{aligned}
&\widehat{\mathbb{E}}\, X^{(\alpha)} - \mathbb{E}\, X^{(\alpha)} \\
=&\widehat{\mathbb{E}}\, X^{(\alpha)} - \mathbb{E}\left(\widehat{\mathbb{E}}\, X^{(\alpha)}|\mathbb{T}\right) + \mathbb{E}\left(\widehat{\mathbb{E}}\, X^{(\alpha)}|\mathbb{T}\right) - \mathbb{E}\, X^{(\alpha)} \\
=&\widehat{w}_{10}\big(\widehat{R}_{00} - \mathbb{E}\left(\widehat{R}_{00}|\mathbb{T}\right)\big) + \widehat{w}_{20}\big(\widehat{R}_{10} - \mathbb{E}\left(\widehat{R}_{10}|\mathbb{T}\right)\big) + \widehat{w}_{11}\big(\widehat{R}_{01} - \mathbb{E}\left(\widehat{R}_{01}|\mathbb{T}\right)\big) && \text{(S1)} \\
&+ \widehat{w}_{10}\big(\mathbb{E}\left(\widehat{R}_{00}|\mathbb{T}\right) - \widehat{S}_{00}C(t,t) - \widehat{S}_{10}\mathbb{E}\, X^{(\alpha)} - \widehat{S}_{01}\partial_{s_2}C(t,t)\big) && \text{(S2)} \\
&+ \widehat{w}_{20}\big(\mathbb{E}\left(\widehat{R}_{10}|\mathbb{T}\right) - \widehat{S}_{10}C(t,t) - \widehat{S}_{20}\mathbb{E}\, X^{(\alpha)} - \widehat{S}_{11}\partial_{s_2}C(t,t)\big) && \text{(S3)} \\
&+ \widehat{w}_{11}\big(\mathbb{E}\left(\widehat{R}_{01}|\mathbb{T}\right) - \widehat{S}_{01}C(t,t) - \widehat{S}_{11}\mathbb{E}\, X^{(\alpha)} - \widehat{S}_{02}\partial_{s_2}C(t,t)\big). && \text{(S4)}
\end{aligned}$$

The arguments in Zhang and Wang [2016] show that $\widehat{S}_{ab} \lesssim h^{a+b}$, $\widehat{D} \asymp h^4$, $\widehat{w}_{10} \lesssim h^{-1}$, $\widehat{w}_{20} \lesssim h^{-2}$ and $\widehat{w}_{11} \lesssim h^{-2}$. Then Taylor expansion of $C(T_{ij_1}, T_{ij_2})$ suggests that the bias part containing (S2)(S3)(S4) shares the rate $O_p(h) = O_p(h^{1+q-\|\alpha\|_1})$.

Step 2: The three items in the deviation part (S1) share the same convergence rate. We present the analysis first item and the other two items are dealt with by similar approach. The first item admits the decomposition

$$\widehat{w}_{10}\big(\widehat{R}_{00} - \mathbb{E}\left(\widehat{R}_{00}|\mathbb{T}\right)\big) = (\widehat{w}_{10} - \widetilde{w}_{10})\big(\widehat{R}_{00} - \mathbb{E}\left(\widehat{R}_{00}|\mathbb{T}\right)\big) + \widetilde{w}_{10}\big(\widehat{R}_{00} - \mathbb{E}\left(\widehat{R}_{00}|\mathbb{T}\right)\big) \quad \text{(S5)}$$

where $\widetilde{w}_{10} := (\widetilde{S}_{11}\widetilde{S}_{01} - \widetilde{S}_{10}\widetilde{S}_{02})\widetilde{D}^{-1}$, $\widetilde{w}_{20} := (\widetilde{S}_{00}\widetilde{S}_{02} - \widetilde{S}_{01}\widetilde{S}_{01})\widetilde{D}^{-1}$ and $\widetilde{w}_{11} := (\widetilde{S}_{10}\widetilde{S}_{01} - \widetilde{S}_{00}\widetilde{S}_{11})\widetilde{D}^{-1}$ with $\widetilde{S}_{ab} := \mathbb{E}\, K_h(T_{ij_1} - t)K_h(T_{ij_2} - t) \times (T_{ij_1} - t)^a(T_{ij_2} - t)^b$ and $\widetilde{D} := \widetilde{S}_{00}(\widetilde{S}_{20}\widetilde{S}_{02} - \widetilde{S}_{11}\widetilde{S}_{11}) + \widetilde{S}_{10}(\widetilde{S}_{01}\widetilde{S}_{11} - \widetilde{S}_{10}\widetilde{S}_{02}) + \widetilde{S}_{01}(\widetilde{S}_{10}\widetilde{S}_{11} - \widetilde{S}_{01}\widetilde{S}_{20})$. The arguments in Zhang and Wang [2016] also suggest that $\widehat{S}_{ab} - \widetilde{S}_{ab} = O_p(h^{a+b}\sqrt{\frac{1}{n} + \frac{1}{n(mh)^{\|\alpha\|_0}}})$, $\widehat{w}_{ab} - \widetilde{w}_{ab} = O_p(h^{-(a+b)}\sqrt{\frac{1}{n} + \frac{1}{n(mh)^{\|\alpha\|_0}}})$, $\widetilde{w}_{ab} \lesssim h^{-(a+b)}$ and $\widehat{R}_{ab} - \mathbb{E}\left(\widehat{R}_{ab}|\mathbb{T}\right) = O_p(h^{a+b}\sqrt{\frac{1}{n} + \frac{1}{n(mh)^{\|\alpha\|_0}}})$. Therefore, the first item in the right side of (S5) shares the rate

$$(\widehat{w}_{10} - \widetilde{w}_{10})\big(\widehat{R}_{00} - \mathbb{E}\left(\widehat{R}_{00}|\mathbb{T}\right)\big) = O_p\big(\tfrac{1}{nh}(1 + \tfrac{1}{(mh)^2})\big) = O_p\big(\tfrac{1}{nh^{\|\alpha\|_1}}(1 + \tfrac{1}{(mh)^{\|\alpha\|_0}})\big)$$

that is relatively smaller compared with our target.

Step 3: It remains to consider the second item in the right side of (S5). With the aid of integration operation $\int_H^{1-H}(\cdot)dt$, the second moment of that is

$$\mathbb{E}\left(\int c\widetilde{w}_{10}\big(\widehat{R}_{00} - \mathbb{E}\left(\widehat{R}_{00}|\mathbb{T}\right)\big)dt\right)^2 = \mathbb{E}\left(\tfrac{1}{nm(m-1)}\int c\widetilde{w}_{10}\textstyle\sum_{i,j_1\neq j_2} K_h(T_{ij_1} - t)K_h(T_{ij_2} - t)U_{ij_1j_2}dt\right)^2$$

where $U_{ij_1j_2} = X_{ij_1}X_{ij_2} - \mathbb{E}\left(X_{ij_1}X_{ij_2}|\mathbb{T}\right)$. Taking the square off, the summation goes over (i, j_1, j_2) and (i', j_1', j_2') but only these three kinds of items are of non-zero expectation:

(1) $i = i'$, $\quad j_1, j_2, j_1', j_2'$ are distinct;

(2) $i = i'$, $\quad j_1 = j_1'$ and j_2, j_1', j_2' are distinct;

(3) $i = i'$, $\quad j_1 = j_1'$ and $j_2 = j_2'$.

The first kind that includes $O(nm^4)$ items equals

$$\begin{aligned}
&\mathbb{E}\int \widetilde{w}_{10}(t)c(t)K_h(T_{ij_1} - t)K_h(T_{ij_2} - t)U_{ij_1j_2}dt \int \widetilde{w}_{10}(s)c(s)K_h(T_{ij_1'} - s)K_h(T_{ij_2'} - s)U_{ij_1'j_2'}ds \\
=&\int\int \widetilde{w}_{10}(t)c(t)\widetilde{w}_{10}(s)c(s)\mathbb{E}\, K_h(T_{ij_1} - t)K_h(T_{ij_2} - t)K_h(T_{ij_1'} - s)K_h(T_{ij_2'} - s)U_{ij_1,j_2}U_{ij_1',j_2'}dtds \\
\lesssim& h^{-2}\int\int \mathbb{E}\, K_h(T_{ij_1} - t)K_h(T_{ij_2} - t)K_h(T_{ij_1'} - s)K_h(T_{ij_2'} - s)dtds \\
\lesssim& h^{-2}\int\int \mathbb{E}\, K_h(T_{ij_1} - t)\mathbb{E}\, K_h(T_{ij_2} - t)\mathbb{E}\, K_h(T_{ij_1'} - s)\mathbb{E}\, K_h(T_{ij_2'} - s)dtds \\
=&O(h^{-2}).
\end{aligned}$$

The second kind that includes $O(nm^3)$ items equals

$$\mathbb{E}\int \widetilde{w}_{10}(t)c(t)K_h(T_{ij_1} - t)K_h(T_{ij_2} - t)U_{ij_1j_2}dt \int \widetilde{w}_{10}(s)c(s)K_h(T_{ij_1'} - s)K_h(T_{ij_2'} - s)U_{ij_1j_2'}ds$$

$$
\begin{aligned}
&= \int\int \widetilde{w}_{10}(t)c(t)\widetilde{w}_{10}(s)c(s)\mathbb{E}\, K_h(T_{ij_1}-t)K_h(T_{ij_2}-t)K_h(T_{ij_1}-s)K_h(T_{ij_2'}-s)U_{ij_1,j_2}U_{ij_1',j_2'}dtds \\
&\lesssim h^{-2}\int\int \mathbb{E}\, K_h(T_{ij_1}-t)K_h(T_{ij_2}-t)K_h(T_{ij_1}-s)K_h(T_{ij_2'}-s)dtds \\
&\lesssim h^{-2}\int\int I_{|t-s|\le 2h}\mathbb{E}\, K_h(T_{ij_1}-t)K_h(T_{ij_1}-s)\mathbb{E}\, K_h(T_{ij_2}-t)\mathbb{E}\, K_h(T_{ij_2'}-s)dtds \\
&= O(h^{-2}).
\end{aligned}
$$

The third kind that includes $O(nm^2)$ items equals

$$
\begin{aligned}
&\mathbb{E}\int \widetilde{w}_{10}(t)c(t)K_h(T_{ij_1}-t)K_h(T_{ij_2}-t)U_{ij_1j_2}dt\int \widetilde{w}_{10}(s)c(s)K_h(T_{ij_1}-s)K_h(T_{ij_2}-s)U_{ij_1j_2}ds \\
&= \int\int \widetilde{w}_{10}(t)c(t)\widetilde{w}_{10}(s)c(s)\mathbb{E}\, K_h(T_{ij_1}-t)K_h(T_{ij_2}-t)K_h(T_{ij_1}-s)K_h(T_{ij_2}-s)U^2_{ij_1j_2}dtds \\
&\lesssim h^{-2}\int\int \mathbb{E}\, K_h(T_{ij_1}-t)K_h(T_{ij_2}-t)K_h(T_{ij_1}-s)K_h(T_{ij_2}-s)dtds \\
&\lesssim h^{-2}\int\int I_{|t-s|\le 2h}\mathbb{E}\, K_h(T_{ij_1}-t)K_h(T_{ij_1}-s)\mathbb{E}\, K_h(T_{ij_2}-t)K_h(T_{ij_2}-s)dtds \\
&= O(h^{-2}h^{-1}).
\end{aligned}
$$

Combine above three kinds together to obtain

$$
\mathbb{E}\left(\int c\widetilde{w}_{10}\big(\widehat{R}_{00}-\mathbb{E}\,(\widehat{R}_{00}|\mathbb{T})\big)dt\right)^2 = O\big(\tfrac{1}{nh^2}\big(1+\tfrac{1}{m^2h}\big)\big) = O\big(\tfrac{1}{nh^{2\|\alpha\|_1}}\big(1+\tfrac{1}{m^{\|\alpha\|_0}h^{\|\alpha\|_0-1}}\big)\big).
$$

Step 4: The final convergence rate is achieved by inserting the results in Steps (2) and (3) into the deviation part (S1).

Proof (Proof of Theorem 2).

Step 1: We first show that $(\beta_1,\ldots,\beta_p)^T = -M^{-1}M_0$ where the matrix $M=(M_{kl})_{1\le k,l\le p}$ and the vector $M_0=(M_{0l})_{1\le l\le p}$ are composed by the integration

$$
M_{kl} = \int_H^{1-H}\mathbb{E}\, X^{(\alpha_k)}(t)g_k(t)\mathbb{E}\, X^{(\alpha_l)}(t)g_l(t)dt.
$$

Consider the function $A(\theta_1,\ldots,\theta_p)=\int_H^{1-H}\left(\sum_{k=0}^p\theta_k\mathbb{E}\, X^{(\alpha_k)}g_k\right)^2dt$ with $\theta_0=1$. Obviously $A\ge 0$ and reaches 0 when $\theta_k=\beta_k$ due to the model (1) and thus $\partial_{\theta_l}A(\beta_1,\ldots,\beta_p)=0$ for $l=1,\ldots,p$. The expression $(\beta_1,\ldots,\beta_p)^T=-M^{-1}M_0$ is immediately obtained by these linear equations and the non-singularity condition in Assumption 1 (3).

Step 2: The deviation of the parameter estimates are measured by

$$
(\widehat{\beta}_1-\beta_1,\ldots,\widehat{\beta}_p-\beta_p)^T = -(\widehat{M}^{-1}\widehat{M}_0 - M^{-1}M_0). \tag{S6}
$$

Consider the following decomposition for each entry $\widehat{M}_{kl} - M_{kl}$

$$\int_H^{1-H} \widehat{\mathbb{E}}\, X^{(\alpha_k)} g_k(t) \widehat{\mathbb{E}}\, X^{(\alpha_l)} g_l(t) dt - \int_H^{1-H} \mathbb{E}\, X^{(\alpha_k)} g_k(t) \mathbb{E}\, X^{(\alpha_l)} g_l(t) dt$$

$$= \int_H^{1-H} \left(\widehat{\mathbb{E}}\, X^{(\alpha_k)} g_k(t) - \mathbb{E}\, X^{(\alpha_k)} g_k(t)\right) \mathbb{E}\, X^{(\alpha_l)} g_l(t) dt \tag{S7}$$

$$+ \int_H^{1-H} \mathbb{E}\, X^{(\alpha_k)} g_k(t) \left(\widehat{\mathbb{E}}\, X^{(\alpha_l)} g_l(t) - \mathbb{E}\, X^{(\alpha_l)} g_l(t)\right) dt \tag{S8}$$

$$+ \int_H^{1-H} \left(\widehat{\mathbb{E}}\, X^{(\alpha_k)} g_k(t) - \mathbb{E}\, X^{(\alpha_k)} g_k(t)\right) \left(\widehat{\mathbb{E}}\, X^{(\alpha_l)} g_l(t) - \mathbb{E}\, X^{(\alpha_l)} g_l(t)\right) dt. \tag{S9}$$

The convergence rate of items (S7)(S8) are discussed in Theorem 1. The item (S9) is bounded by $\|\widehat{\mathbb{E}}\, X^{(\alpha_k)} - X^{(\alpha_k)}\|_{L^2} \|\widehat{\mathbb{E}}\, X^{(\alpha_l)} - \mathbb{E}\, X^{(\alpha_l)}\|_{L^2}$ according to Cauchy inequality, and is further described by the square of L^2 convergence rate discussed in Zhang and Wang [2016], which is relatively smaller to items (S7)(S8).

The convergence rate in (S6) is immediately obtained from Lemma S1 and non-singularity in Assumption 1 (3).

Lemma S1. *Suppose A is a $p \times p$ non-singular matrix and v is a p-dimensional vector. The perturbations A' and v' admit the entrywise condition* $\max_{1 \le k,l \le p} |A'_{kl} - A_{kl}| = O_p(r_n)$ *and* $\max_{1 \le k \le p} |v'_k - v_k| = O_p(r_n)$*, respectively, for a converging series $r_n \to 0$. Then the vector $(A')^{-1} v'$ shares the same convergence rate*

$$\|(A')^{-1} v' - A^{-1} v\| = O_p(r_n).$$

Proof (Proof of Lemma S1). Recall that $A^{-1} = \frac{A^*}{\det(A)}$ where A^* refers to the classic adjoint matrix of A. Simple computation shows $\det(A') - \det(A) = O_p(r_n)$ and $\|(A')^* - A^*\| = O_p(r_n)$. Then the result $\|(A')^{-1} v' - A^{-1} v\| = O_p(r_n)$ is immediately obtained from non-singularity.

Proof (Proof of Theorem 3).

Step 1: Recall that the principal scores for the ith curve X_i are estimated by $(\widehat{\xi_{i1}}, \ldots, \widehat{\xi_{il}}) = \arg\min_{(\theta_{i1}, \ldots, \theta_{iL})} \widehat{S}(\theta_{i1}, \ldots, \theta_{iL})$ in (12) where

$$\widehat{S}(\theta_{i1}, \ldots, \theta_{iL}) = \tfrac{1}{m} \textstyle\sum_{j=1}^m \left(\widehat{Z}_i(T_{ij}) - X_{ij}\right)^2 + \tfrac{\kappa}{1-2H} \int_H^{1-H} \left(F_{\widehat{\beta}}(\widehat{Z}_i)\right)^2 dt$$

for $\widehat{Z}_i(\theta_{i1}, \ldots, \theta_{iL})(t) = \widehat{\mathbb{E}}\, X(t) + \sum_{1 \le l \le L} \theta_{il} \widehat{\phi}_l(t)$ and the observations are $X_{ij} = \mathbb{E}\, X(T_{ij}) + \sum_{1 \le l \le \infty} \xi_{il} \phi_l(T_{ij}) + \varepsilon_{ij}$.

Consider the non-plugged-in version $(\widetilde{\xi}_{i1},\dots,\widetilde{\xi}_{iL}) = \arg\min_{(\theta_{i1},\dots,\theta_{iL})}\widetilde{S}(\theta_{i1},\dots,\theta_{iL})$ where

$$\widetilde{S}(\theta_{i1},\dots,\theta_{iL}) = \frac{1}{m}\sum_{j=1}^{m}\big(\widetilde{Z}_i(T_{ij}) - X_{ij}\big)^2 + \frac{\kappa}{1-2H}\int_H^{1-H}\big(F_\beta(\widetilde{Z}_i)\big)^2 dt$$

for $\widetilde{Z}_i(\theta_{i1},\dots,\theta_{iL})(t) = \mathbb{E}\,X(t) + \sum_{1\le l\le L}\theta_{il}\phi_l(t)$ and the parameters $\{\beta_k\}_{k=0}^{p}$. Then convergence of score estimates is analyzed by the following decomposition

$$(\widehat{\xi}_{i1}-\xi_{i1},\dots,\widehat{\xi}_{il}-\xi_{iL}) = (\widehat{\xi}_{i1}-\widetilde{\xi}_{i1},\dots,\widehat{\xi}_{il}-\widetilde{\xi}_{iL}) + (\widetilde{\xi}_{i1}-\xi_{i1},\dots,\widetilde{\xi}_{iL}-\xi_{iL})$$

and each item is addressed in the following steps.

Step 2: Notice that $\widetilde{S}(\theta_{i1},\dots,\theta_{iL})$ only relies on the data $\{X_{ij}\}_{1\le j\le m}$ from the i-th curve X_i. Applying Lagrange's Mean Value Theorem to $\langle V,\nabla\widetilde{S}\rangle$ for $V\in\mathbb{R}^L$ to obtain

$$\begin{aligned}&\langle V,\nabla\widetilde{S}(\widetilde{\xi}_{i1},\dots,\widetilde{\xi}_{iL}) - \nabla\widetilde{S}(\xi_{i1},\dots,\xi_{iL})\rangle\\ =&V^T(\mathtt{H}\widetilde{S})(\widetilde{\xi}_{i1}-\xi_{i1},\dots,\widetilde{\xi}_{iL}-\xi_{iL})^T + O\big(\|V\|\times\|(\widetilde{\xi}_{i1}-\xi_{i1},\dots,\widetilde{\xi}_{iL}-\xi_{iL})\|^2\big),\end{aligned}$$

where $\mathtt{H}\widetilde{S}$ denotes the Hessian matrix at $(\xi_{i1},\dots,\xi_{iL})\in\mathbb{R}^L$. The smallest eigenvalue of Hessian matrix $\mathtt{H}\widetilde{S}$, defined by $\mathtt{H}\widetilde{S}_{\mathrm{se}} = \inf_{\gamma\in\mathbb{R}^L,\|\gamma\|=1}\gamma^T\mathtt{H}\widetilde{S}\gamma$, is bounded below in probability when $\kappa\to 0$ with the aid of Assumption 2 (2) since

$$\begin{aligned}\mathtt{H}\widetilde{S}_{\mathrm{se}} =& \inf_{\gamma\in\mathbb{R}^L,\|\gamma\|=1}\sum_{1\le l_1,l_2\le L}\gamma_{l_1}\gamma_{l_2}\partial_{\theta_{il_1}}\partial_{\theta_{il_2}}\widetilde{S}\\ =& \inf_{\gamma\in\mathbb{R}^L,\|\gamma\|=1}\Big(\frac{2}{m}\sum_{j=1}^{m}\sum_{1\le l_1,l_2\le L}\gamma_{l_1}\gamma_{l_2}\phi_{l_1}(T_{ij})\phi_{l_2}(T_{ij})\\ &+\frac{2\kappa}{1-2H}\int_H^{1-H}\langle d^2\big(F_\beta(\widetilde{Z}_i(\xi_{i1},\dots,\xi_{iL}))\big)^2,\sum_{1\le l_1\le L}\gamma_{l_1}\phi_{l_1},\sum_{1\le l_2\le L}\gamma_{l_2}\phi_{l_2}\rangle dt\Big)\\ \gtrsim& \inf_{\gamma\in\mathbb{R}^L,\|\gamma\|=1}\Big(\frac{2}{m}\sum_{j=1}^{m}\big(\sum_{1\le l_1\le L}\gamma_{l_1}\phi_{l_1}(T_{ij})\big)^2\\ &-\frac{2\kappa}{1-2H}\big\|d^2\big(F_\beta(\widetilde{Z}_i(\xi_{i1},\dots,\xi_{iL}))\big)^2\big\|\times\big\|\sum_{1\le l_1\le L}\gamma_{l_1}\phi_{l_1}\big\|_{H^{\nu_0}}^2\Big)\gtrsim_p 1.\end{aligned}\tag{S10}$$

In the last inequality (S10), the item $\frac{1}{m}\sum_{j=1}^{m}\big(\sum_{1\le l_1\le L}\gamma_{l_1}\phi_{l_1}(T_{ij})\big)^2$ converges to the integration $\int\big(\sum_{1\le l\le L}\gamma_l\phi_l(t)\big)^2 f_T(t)dt \ge \inf_t f_T(t) > 0$ that can be bounded below. The second-order Fréchet derivative $\big\|d^2\big(F_\beta(\widetilde{Z}_i(\xi_{i1},\dots,\xi_{iL}))\big)^2\big\|$ is bounded above in probability since

$$\|X_i - \widetilde{Z}_i(\xi_{i1},\dots,\xi_{iL}))\|_{H^{\nu_0}} \le \sum_{l=L+1}^{\infty}|\xi_{il}|l^{\nu_0 b} \lesssim_p L^{-a+\nu_0 b+1}\to 0$$

and $\|d^2\big(F_\beta(X_i)\big)^2\| = O_p(1)$ from Assumption 2 (2). The norm $\big\|\sum_{1\le l_1\le L}\gamma_{l_1}\phi_{l_1}\big\|_{H^{\nu_0}}$ is controlled by

$$\big\|\sum_{1\le l_1\le L}\gamma_{l_1}\phi_{l_1}\big\|_{H^{\nu_0}} \le \sum_{1\le l_1\le L}\gamma_{l_1} l_1^{\nu_0 b} = O(L^{\nu_0 b+0.5}).$$

Combing above bounds and the tuning parameter constraint $\kappa L^{2\nu_0 b+1} = o(1)$ to obtain $\mathtt{H}\widetilde{S}_{\mathrm{se}} \gtrsim_p 1$.

Step 3: Then $(\widetilde{\xi}_{i1}-\xi_{i1},\ldots,\widetilde{\xi}_{iL}-\xi_{iL})$ is bounded above by $\langle V, \nabla\widetilde{S}(\widetilde{\xi}_{i1},\ldots,\widetilde{\xi}_{iL}) - \nabla\widetilde{S}(\xi_{i1},\ldots,\xi_{iL})\rangle$. The l'-th component $(1\le l'\le L)$ of the gradient $\nabla\widetilde{S}$ is

$$\begin{aligned}
&\partial_{\theta_{il'}}\widetilde{S}(\widetilde{\xi}_{i1},\ldots,\widetilde{\xi}_{iL}) - \partial_{\theta_{il'}}\widetilde{S}(\xi_{i1},\ldots,\xi_{iL}) = \partial_{\theta_{il'}}\widetilde{S}(\xi_{i1},\ldots,\xi_{iL})\\
&= -\tfrac{2}{m}\sum_{j=1}^m\big(\sum_{l=L+1}^\infty \xi_{il}\phi_l(T_{ij})\big)\phi_{l'}(T_{ij}) - \tfrac{2}{m}\sum_{j=1}^m \varepsilon_{ij}\phi_{l'}(T_{ij})\\
&\quad - \tfrac{2\kappa}{1-2H}\int_H^{1-H} F_\beta(\widetilde{Z}_i(\xi_{i1},\ldots,\xi_{iL}))\langle dF_\beta(\widetilde{Z}_i(\xi_{i1},\ldots,\xi_{iL})), \phi_{l'}(T_{ij})\rangle dt.
\end{aligned}$$

The norm of the first part is bounded by the decaying scores

$$\sum_{1\le l'\le L}\big|\tfrac{2}{m}\sum_{j=1}^m\sum_{l=L+1}^\infty \xi_{il}\phi_l(T_{ij})\phi_{l'}(T_{ij})\big|^2 \lesssim \sum_{1\le l'\le L}\big(\sum_{l=L+1}^\infty|\xi_{il}|\big)^2 = O_p(L^{-2a+3}).$$

The norm of the second part reflects the influences from the noise

$$\mathbb{E}\sum_{1\le l'\le L}\big(\tfrac{2}{m}\sum_{j=1}^m\varepsilon_{ij}\phi_{l'}(T_{ij})\big)^2 = \tfrac{4\mathrm{Var}(\varepsilon_{11})}{m^2}\mathbb{E}\sum_{1\le l'\le L}\sum_{j=1}^m\phi_{l'}^2(T_{ij}) = O(Lm^{-1}).$$

In the third part, consider $F_\beta(X_i) - F_\beta(\widetilde{Z}_i(\xi_{i1},\ldots,\xi_{iL}))$ that is mainly determined by $X_i^{(\nu)} - \widetilde{Z}_i^{(\nu)}(\xi_{i1},\ldots,\xi_{iL}) = \sum_{l=L+1}^\infty \xi_{il}\phi_l^{(\nu)}$ and recall $F_\beta(X_i) = 0$

$$\begin{aligned}
&|F_\beta(\widetilde{Z}_i(\xi_{i1},\ldots,\xi_{iL}))| \le |\langle dF_\beta(\acute{\xi}), X_i - \widetilde{Z}_i(\xi_{i1},\ldots,\xi_{iL})\rangle| = |\langle dF_\beta(\acute{\xi}), \textstyle\sum_{l=L+1}^\infty\xi_{il}\phi_l\rangle|\\
&\le \|dF_\beta(\acute{\xi})\|\times\|\textstyle\sum_{l=L+1}^\infty\xi_{il}\phi_l\|_{H^{\nu_0}} \le \|dF_\beta(\acute{\xi})\|\times(\sum_{l=L+1}^\infty|\xi_{il}|l^{\nu_0 b})
\end{aligned}$$

where $\acute{\xi}$ is a point between X_i and $\widetilde{Z}_i(\xi_{i1},\ldots,\xi_{iL})$. Similar arguments to *Step 2* show that $\|dF_\beta(\acute{\xi})\|$ and $\|dF_\beta(\widetilde{Z}_i(\xi_{i1},\ldots,\xi_{iL}))\|$ are bounded above with probability tending to one. Then the third part enjoys the bound

$$\begin{aligned}
&\sum_{1\le l'\le L}\big(\tfrac{2\kappa}{1-2H}\int_H^{1-H}F_\beta(\widetilde{Z}_i(\xi_{i1},\ldots,\xi_{iL}))\langle dF_\beta(\widetilde{Z}_i(\xi_{i1},\ldots,\xi_{iL})), \phi_{l'}(T_{ij})\rangle dt\big)^2\\
&\lesssim_p \sum_{1\le l'\le L}\big(\kappa L^{-a+\nu_0 b+1}\times l'^{\nu_0 b}\big)^2 \lesssim \kappa^2 L^{-2a+4\nu_0 b+3} \lesssim L^{-2a+1}.
\end{aligned}$$

Combining above three bounds and the results in *Step 2* gives

$$\|(\widetilde{\xi}_{i1}-\xi_{i1},\ldots,\widetilde{\xi}_{iL}-\xi_{iL})\| = O_p(L^{-a+1.5} + L^{0.5}m^{-0.5}). \tag{S11}$$

Step 4: It remains to study $(\widehat{\xi_{i1}} - \widetilde{\xi}_{i1}, \ldots, \widehat{\xi}_{il} - \widetilde{\xi}_{iL})$. Simple computation shows

$$\begin{aligned}
&\left|\widehat{S}(\theta_{i1}, \ldots, \theta_{iL}) - \widetilde{S}(\theta_{i1}, \ldots, \theta_{iL})\right| \\
=&\tfrac{1}{m}\textstyle\sum_{j=1}^{m}\left(\widehat{Z}_i(T_{ij}) - X_{ij}\right)^2 + \tfrac{\kappa}{1-2H}\int_H^{1-H}\left(F_{\widehat{\beta}}(\widehat{Z}_i)\right)^2 dt \\
&- \tfrac{1}{m}\textstyle\sum_{j=1}^{m}\left(\widetilde{Z}_i(T_{ij}) - X_{ij}\right)^2 + \tfrac{\kappa}{1-2H}\int_H^{1-H}\left(F_{\beta}(\widetilde{Z}_i)\right)^2 dt \\
=&\tfrac{1}{m}\textstyle\sum_{j=1}^{m}\left(\widehat{Z}_i(T_{ij}) + \widetilde{Z}_i(T_{ij}) - 2X_{ij}\right)\left(\widehat{Z}_i(T_{ij}) - \widetilde{Z}_i(T_{ij})\right) \qquad \text{(S12)} \\
&+ \tfrac{\kappa}{1-2H}\int_H^{1-H}\left(F_{\widehat{\beta}}(\widehat{Z}_i) + F_{\beta}(\widetilde{Z}_i)\right)\left(F_{\widehat{\beta}}(\widehat{Z}_i) - F_{\beta}(\widetilde{Z}_i)\right) dt. \qquad \text{(S13)}
\end{aligned}$$

The part in (S12) converges to its expectation on $\{T_{ij}\}_{1\le j\le m}$ with rate $O_p(m^{-0.5})$, and thus is bounded by $\|\widehat{\mathbb{E}}\, X - \mathbb{E}\, X\|_{L^2}$ and $\sum_{1\le l\le L}\theta_{il}\|\widehat{\phi}_l - \phi_l\|_{L^2}$ plus $O_p(m^{-0.5})$. Note that the two potential orientations of eigenfunction estimates ($\widehat{\phi}_l$ and $-\widehat{\phi}_l$) would lead to the opposite score estimate ($\widehat{\xi}_{il}$ and $-\widehat{\xi}_{il}$) and finally result in the identical recovery estimates $\widehat{X}_i$ as in (13). Consequently, we consider the choice $\widehat{\phi}_l$ with $\int_T \widehat{\varphi}_l\varphi_l > 0$ to simplify the notations in this proof. Under this convention, Bosq [2000] achieves the L^2 convergence rate for eigenfunction estimate $\|\widehat{\phi}_l - \phi_l\|_{L^2} = O_p(l^{2a+1}r_{\Sigma,0})$ and eigenvalue estimate $\|\widehat{\lambda}_l - \lambda_l\|_{L^2} = O_p(l^{-2a}r_{\Sigma,0})$ that rely on the diverging index l. The notation $r_{\Sigma,0} = h_\Sigma + \sqrt{\frac{1}{n}\left(1 + \frac{1}{m^2h_\Sigma^2}\right)}$ is the L^2 convergence rate of covariance estimate in (14) and h_Σ is the bandwidth. The L^2 convergence rate for mean estimate $\widehat{\mathbb{E}}\, X$ is relatively smaller than $r_{\Sigma,0}$, i.e. converging faster [Zhang and Wang, 2016]. Therefore, the part in (S12) enjoys the rate $O_p(m^{-0.5} + L^{a+2}r_{\Sigma,0})$ uniformly over $\theta_{il} = O_p(l^{-a})$.

The part in (S13) is dominated by $\max_{1\le k\le p}\kappa|\widehat{\beta}_k - \beta_k|$ and $\kappa\|\widehat{Z}_i - \widetilde{Z}_i\|_{H^{\nu_0}}$ according to Assumption 2 (2). Theorem 2 suggests $\max_{1\le k\le p}|\widehat{\beta}_k - \beta_k| = O_p(r_\beta)$ with

$$r_\beta = \max_{0\le k\le p}\left\{h_k + \sqrt{\tfrac{1}{nh_k^{2\|\alpha_k\|_1}}\left(1 + \tfrac{1}{m^{\|\alpha_k\|_0}h_k^{\|\alpha_k\|_0 - 1}}\right)}\right\}.$$

According to the eigenfunction derivative estimate in (15), the item $\|\widehat{Z}_i - \widetilde{Z}_i\|_{H^{\nu_0}}$ is controlled by

$$\begin{aligned}
\|\widehat{Z}_i - \widetilde{Z}_i\|_{H^{\nu_0}} \lesssim &\textstyle\sum_{0\le\nu\le\nu_0}\|\widehat{\mathbb{E}}\, X^{(\nu)} - \mathbb{E}\, X^{(\nu)}\|_{L^2} + \sum_{1\le l\le L}\theta_{il}\sum_{0\le\nu\le\nu_0}\|\widehat{\partial_\nu\phi_l} - \partial_\nu\phi_l\|_{L^2} \\
=&O_p(L^{a+1}r_{\Sigma,\nu_0} + L^{3a+2}r_{\Sigma,0})
\end{aligned}$$

uniformly over $\theta_{il} = O_p(l^{-a})$, where r_{Σ,ν_0} refers to the L^2 rate of estimate $\widehat{\partial_t^{\nu_0}\Sigma}$ in (14).

Above analysis gives $\left|\widehat{S}(\theta_{i1},\ldots,\theta_{iL})-\widetilde{S}(\theta_{i1},\ldots,\theta_{iL})\right| = O_p(m^{-0.5}+L^{a+2}r_{\Sigma,0})$ uniformly over $\theta_{il}=O_p(l^{-a})$. Combing the convexity of $\widetilde{S}$ in *Step 2* and the binary-partition/chain technique in the proof of [Shao and Yao, 2023, Theorem 4.1] to obtain

$$\|(\widehat{\xi_{i1}}-\widetilde{\xi}_{i1},\ldots,\widehat{\xi}_{il}-\widetilde{\xi}_{iL})\| = O_p\big(m^{-0.5}+L^{a+2}r_{\Sigma,0}+\kappa(r_\beta+L^{a+1}r_{\Sigma,\nu_0}+L^{3a+2}r_{\Sigma,0})\big). \quad \text{(S14)}$$

Step 5: Finally, the recovery $\widehat{X}_i=\widehat{\mathbb{E}}\,X+\sum_{1\le l\le L}\widehat{\xi}_{il}\widehat{\phi}_l$ shares the following rate based on the results (S11) and (S14)

$$\begin{aligned}\|\widehat{X}_i-X_i\|_{L^2}\le&\|\widehat{\mathbb{E}}\,X-\mathbb{E}\,X\|_{L^2}+\|\textstyle\sum_{1\le l\le L}(\widehat{\xi}_{il}-\widetilde{\xi}_{iL})\widehat{\phi}_l\|_{L^2}+\|\textstyle\sum_{1\le l\le L}(\widetilde{\xi}_{iL}-\xi_{il})\widehat{\phi}_l\|_{L^2}\\&+\|\textstyle\sum_{1\le l\le L}\xi_{il}(\widehat{\phi}_l-\phi_l)\|_{L^2}+\|\textstyle\sum_{L+1\le l<\infty}\xi_{il}\phi_l\|_{L^2}\\=&O_p(L^{-a+2}+Lm^{-0.5}+L^{a+2.5}r_{\Sigma,0}),\end{aligned}$$

where the tuning parameter κ satisfies $\kappa(r_\beta+L^{a+1}r_{\Sigma,\nu_0}+L^{3a+2}r_{\Sigma,0})=O_p(L^{-a+1.5}+L^{0.5}m^{-0.5}+L^{a+2}r_{\Sigma,0})$.

References

D. Bosq. *Linear Proceses in Function Spaces.* Lecture Notes in Statistics. Springer, 2000.

L. Shao and F. Yao. Robust functional data analysis for discretely observed data. *arxiv*, page 2305.16236, 2023.

X. Zhang and J. L. Wang. From sparse to dense functional data and beyond. *The Annals of Statistics*, 44:2281–2321, 2016.