## **Functional Additive Models**

...- XA

 T
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I
 I</t



Т y y у y 1 y y y ( ı Т ) ı y у Т 1 1 X. y y Ņ Т y y ¥ Т I X. functional additive model ( ) Т Т I ı y I Drosophila **X**. y () (,) ( ()() () () ( ) ) ( ) I. . . ı

s 5 complately (g)-0.3(idg)-44[(and)-4158(measur(mnentr)]TJ 0 -1.9252 TD 0.0003 Tc [(are)-625.3(contmbinaend)-361.5(with)-363 mAdg ar (4)5 )55 2 0 . 3 5 dresh(x) - 315.7(x) - 0.4ce esums = 0.25 +X 5 ) 1 T J 0 0 4 6 5 6 1 7 7 1 2 Δ 5 7 1 1 7 7 1

- $E(Y |_{k}) = b_{k | k}$   $E(\zeta_{m} |_{k}) = b_{k m | k}$  ()
- y.
- $f_{k(k)} = \frac{f_{k(k)}}{f_{km}(\cdot)} + \frac{f_{k(k)}}{k} = \frac{f_{k(k)}}{k} + \frac{$
- y T i i i () () <sup>i</sup> y k -

$$E(Y|X) = Y + \sum_{k=1}^{\infty} f_k(x_k) \tag{()}$$

 $E(Y(t)|X) = Y(t) + \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} f_{km}(x_k) m(t), \quad ()$ 

y \_\_\_\_\_

$$Ef_k(k) = 0, \quad k = 0, \dots,$$
  
 $Ef_{km}(k) = 0, \quad k = 0, \dots, m = 0, \dots, m = 0, \dots$  (1)

 $E_{Jkm}(k) \equiv , \qquad k \equiv , \quad , \dots, m \equiv , \quad , \dots$ 

k - ' y '

x y i T i T

$$E(Y - |y||_{k}) = E\left\{E(Y - |y|X)|_{k}\right\}$$
$$= E\left\{\sum_{j=1}^{\infty} f_{j}(|j||_{k}\right\} = f_{k}(|k|), \quad (-)$$

1 , 1 y.

1

ı.

T

$$E(\zeta_m|_k) = E\left\{E(\zeta_m|X)|_k\right\}$$
$$= E\left\{\sum_{j=1}^{\infty} f_{jm}(j)|_k\right\} = f_{km}(k). \quad ()$$

T  $f_k(k) = E(Y - |Y||_k)$   $f_{km}(k) = E(\zeta_m|_k)$   $k, m = 1, \dots,$ 

$$ik$$
 $ik$ 
 $y$ 
 $y$ 
 $y$ 
 $y$ 
 $y$ 
 $y$ 
 $y$ 
 $y$ 
 $k$ 
 $i$ 
 $k$ 
 $j$ 
 $k$ 

tional expectation (E)K = K

$$\hat{f}_k$$
  $\hat{f}_{km}$   $f_k$   $f_k$   $f_k$ 

$$\sum_{i=1}^{n} K\left(\frac{\hat{i}_{k}-x}{h_{k}}\right) \{Y_{i}-\beta -\beta (x-\hat{i}_{k})\}$$
()

$$\beta \qquad \beta \qquad \hat{f}_k(x) = \hat{\beta} (x) - \bar{Y} \qquad h_k$$

(

( )

$$\{\hat{i}_{k}, \hat{\zeta}_{im}\}_{i=\dots,n}^{y} \quad ( )$$

$$\sum_{i=1}^{n} K\left(\frac{\hat{i}_{k}-x}{h_{mk}}\right)\{\hat{\zeta}_{im}-\beta -\beta (x-\hat{i}_{k})\} \quad ( )$$

$$\beta = \beta \quad \hat{f}_{k}(x) = \hat{\beta}_{k}(x)$$

 $\beta \qquad \beta \qquad f_{mk}(x) = \beta (x)$  $h_{mk} \qquad T$ 

$$\widehat{E}(Y|X) = \overline{Y} + \sum_{k=1}^{K} \widehat{f}_k(x_k). \tag{(1)}$$

T y R (\_\_\_\_)

$$\widehat{E}\{Y(t)|X\} = \widehat{Y}(t) + \sum_{m=1}^{M} \sum_{k=1}^{K} \widehat{f}_{mk}(k) \widehat{f}_{m}(t), \qquad t \in \mathcal{T},$$
((1))

 $\begin{array}{l}
\mathbf{y} \\
\mathbf{R} = -\frac{\sum_{i=}^{n} \int \left[Y_{i}(t) - E\{Y_{i}(t)|X_{i}\}\right] dt}{\sum_{i=}^{n} \int \left\{Y_{i}(t) - Y_{i}(t)\right\} dt}, \\
\mathbf{\widehat{R}} = -\frac{\sum_{i=}^{n} \sum_{l=}^{m_{i}} \left[V_{il} - \widehat{E}\{Y_{i}(t_{il})|X_{i}\}\right] (t_{il} - t_{i,l-})}{\sum_{i=}^{n} \sum_{l=}^{m_{i}} \left\{V_{il} - Y_{i}(t_{il})\right\} (t_{il} - t_{i,l-})} \\
\mathbf{\widehat{E}}(Y_{i}(t)|X_{i}) \qquad i \qquad I \qquad I \qquad I \\
\end{array}$ ( )

$$\begin{cases} \hat{i}_{ik}, Y_i \} \\ \hat{j}_{ik}, \hat{\zeta}_{im} \end{cases} \quad i = , ..., n \\ y = 1 \\ K (n) \to \infty \\ ( ) \\ K (n) \to \infty \\ ( ) \\ ($$

$$\hat{f}_k(x) - f_k(x) \xrightarrow{p} , \qquad ()$$

$$\hat{f}_{km}(x) - f_{km}(x) \xrightarrow{p} , \qquad ()$$

( )

$$|\hat{f}_k(x) - f_k(x)| \stackrel{!}{\longrightarrow} \stackrel{\tilde{\vartheta}_{mk}(x) = |\hat{f}_{mk}(x) - f_{mk}(x)|}{( ) }$$

) 
$$\widehat{E}(Y|X) - E(Y|X) \xrightarrow{p} , \qquad ()$$

$$E(Y|X) = Y + \sum_{k=1}^{n} f_k(x) \quad ( )$$

$$\widehat{E}\{Y(t)|X\} - E\{Y(t)|X\} \xrightarrow{p} , \qquad ( )$$

$$\widehat{E}\{Y(t)|X\} = \widehat{Y}(t) + \sum_{k=1}^{K} \sum_{m=1}^{M} \widehat{f}_{mk}(\widehat{x}) \widehat{f}_{m}(t)$$

$$\widehat{f}_{mk}(\widehat{x}) \qquad ( )$$

$$|\widehat{E}(Y|X) - E(Y|X)| \qquad \vartheta_n^* = |\widehat{E}(Y(t)|X) - E(Y(t)|X)|$$

$$\begin{bmatrix} x_{i} & y_{i} & y_$$

$$f_{km} y \ \hat{f}_{km}(x) = \sum_{i=1}^{n} (\hat{\xi}_{im} - \bar{\xi}_{\cdot m})(\hat{i}_{ik} - \bar$$

$$Y_{i}^{*} \qquad V_{il} \qquad V_{il}^{*} \qquad V_{il}^{*} \qquad V_{il}^{*}$$

$$U_{ij} \qquad V_{il} \qquad T \qquad Y \qquad (E)$$

$$E_{i} = \frac{(Y_{i}^{*} - \hat{Y}_{i}^{*})}{Y_{i}^{*}}$$

$$E_{i,f} = \frac{\int (Y_{i}^{*}(t) - \hat{Y}_{i}^{*}(t)) \ dt}{\int Y_{i}^{*}(t) \ dt},$$

$$(-)$$

Τι Ε ...Τ...

y fk fkm '' y ''' T ''y (T ) T y T y

)

T () () () () () ()

-	Т	۲ı	
	•	•	
r		•	
	•	•	
~			
		•	
1			
	•	•	

**E T** ( ) (

Т	I.	1		Т	. 1	I.	~	~	
			I		( .	)			
-									

~	Тт		Тт	
				•
I				•
				•
*				•
				•
I				•
		•		•

X.

y<sub>E</sub> y i













ı.







y = ( ) K = $y = -) \qquad % ($   $, - \cdots$ ( ) M =, ()E % ( % ( ). , -·-·) у ()Е , , I, ( )\_ % ( % ( ) --) % (  $-\cdot - \cdot$ %( 1 У **,** · · · ·)

Т X. Х. T. y . I , I, Х. I. y y, ı T y ı Y





- !- --

 $\mathbf{y} = \mathbf{x}$   $(\zeta_m, m = \ , \ , \ )$ 

Т

Predictor: Embryo Phase

Phase

Response: Pupa

(



т . . т. т. т. т.







у 2.5 2 1.5 1.51.50.5 0.5 0 0 -0.5 -0.5 -0.5 -1 -1 -1 -1.5 -1.5 -1.5 -2 -2 -2 10 15 Time(hours) 10 15 Time(hours) 10 15 Time(hours) 20 0 0 20 0 5 20 5 5 -0.2 -0.2 -0.2 -0.4 -0.4 -0.4 -0.6 -0.6 -0.6 -0.8 -0.8 -0.8 -1 -1 -1.2 -1.2 -1.2 -1.4 -1.4 -1.4 -1.6 -1.6 -1.6 -1.8 -1.8 -1.8 -2 -2 -2.2 -2.2 -2.2 160 180 Time(hours) 160 180 Time(hours) 160 180 Time(hours) 140 200 120 140 200 120 140 200 120



Т  $k, k = 1, \dots, n$  $(A_{G_X}f)(s) =$ X ı. f(s) $f \in L (S)$   $(A_{G_X} f)(t) = \int f(s) \times$ . 1 1  $G_X(s,t) ds$ ) T y- (  $\int j(s) k(s) ds = \delta_{jk}$  $\delta_{jk} = \int^{y} j = k$  $j \neq k$  T = т — т.  $\stackrel{'}{\geq} \geq \cdots$ . I I 1  $G_X(s, s) = \sum_{k=k}^{j} X_k \times X_k$ y · Y т  $\{Y, k\}$ 

 $G_X(s, s) = \sum_{k \in K} G_X(s, s) = \sum_{k \in K} K$   $G_X(s, s) = \sum_{k \in K} K$ 

$$X(s) = \chi(s) + \sum_{j=1}^{\infty} j_{j}(s)$$
  

$$Y(t) = \gamma(t) + \sum_{k=1}^{\infty} \zeta_{k-k}(t).$$
((

)

)

$$\{k, k = 1, \dots\} \quad \{m, m = 1, \dots\}$$

$$\sum_{k=}^{\infty} \beta_{k-k}(t) \qquad \begin{array}{c} \mathbf{x} & \beta_{k} & \beta_{km-1} & \beta(s) = \\ \beta(s,t) = \sum_{k=}^{\infty} \sum_{m=}^{\infty} \beta_{km-k}(s) & m(t) \\ ( ) & \mathbf{y} & \mathbf{y} \\ ( ) & \mathbf{y} \\ ($$

$$\begin{split} U_{ij} &= X_i(s_{ij}) + \epsilon_{ij} \\ &= X(s_{ij}) + \sum_{k=1}^{\infty} i_k k(s_{ij}) + \epsilon_{ij}, \\ &s_{ij} \in S, \quad \leq i \leq n, \quad \leq j \leq n_i, \end{split}$$

$$V_{il} = Y_{i}(t_{il}) + \varepsilon_{il}$$

$$= Y(t_{il}) + \sum_{m=1}^{\infty} \zeta_{im-k}(t_{il}) + \varepsilon_{il},$$

$$s_{il} \in T, \leq i \leq n, \leq l \leq m_{i}. \quad ()$$

$$U_{i} = (U_{i}, ..., U_{in_{i}})^{T}, X_{i} = (X(s_{i}), ..., X(s_{in_{i}}))^{T}, X_{i} = (X(s_{i}), ..., X(s_{in_{i}})^{T}, X_{i})^{T}, X_{i} = (X(s_{i}), ..., X(s_{i}))^{T}, X_{i} = (X(s_{i}), ..., X(s_{i}), X(s_{i})^{T}, X(s_{i})^{T}, X(s_{i}), X(s_{i})^{T}, X(s_{i})^{T$$

 $\sum_{U_{i}} (J_{i}) \sum_{U_{i}} (\Sigma_{U_{i}})_{j,l} = G_{X}(s_{ij}, s_{il}) + (Y_{ij})_{j,l} = G_{X}(s_{ij}, s_{il}) + (Y_{$ 

$$\hat{\zeta}_{ik}^{I} = \sum_{j=1}^{n_{i}} (U_{ij} - \hat{\chi}(t_{ij})) \hat{\chi}(s_{ij})(s_{ij} - s_{i,j-}),$$

$$\hat{\zeta}_{im}^{I} = \sum_{j=1}^{n_{i}} (V_{ij} - \hat{\chi}(t_{ij})) \hat{\chi}(t_{ij})(t_{ij} - t_{i,j-}),$$

$$( )$$

 $\zeta_{im} = \int \{ X_i(s) - X_i(s) \}_k(s) \, ds \qquad \zeta_{im} = \int \{ Y_i(t) - Y_i(t) \}_{m(t)} \, dt \, T = E$ 

 $\{ \begin{array}{c} & & \\ & &$ 

$$\hat{X} = \int \{\widehat{V}_X(s) - \widetilde{G}_X(s)\} ds / |S| \qquad ()$$

 $\hat{X} > \hat{X} = E$   $\{ k, k\}_{k \ge 1}$   $\{ k, k = 1 \}$ 

[Received August 2007. Revised June 2008.]

K E E K ( ) E x Y Drosophila melanogaster, Science ( ) Topics in Stochastic Processes

y ( ) Probability and Measure ( ) T ( ) Annals of Statistics

( ) E Journal of Multivariate Analysis

y ( ) T Scandinavian Journal of Statistics

T X ( ) T Canadian Journal of Statistics

## Supplement for "Functional Additive Models"

## 1. NOTATIONS AND AUXILIARY RESULTS

Covariance operators are denoted by  $\mathcal{G}_{\mathbf{X}}$ ,  $\widehat{\mathcal{G}}_{\mathbf{X}}$ , generated by kernels  $G_{\mathbf{X}}$ ,  $\widehat{G}_{\mathbf{X}}$ ; i.e.,  $\mathcal{G}_{\mathbf{X}}(f) = \int_{\mathcal{T}} G_{\mathbf{X}}(s,t) f(s) ds$ ,  $\widehat{\mathcal{G}}_{\mathbf{X}}(f) = \int_{\mathcal{T}} \widehat{G}_{\mathbf{X}}(s,t) f(s) ds$  for any  $f \in L^2(\mathcal{T})$ . Define

$$D_{\mathbf{X}} = \int_{\mathcal{T}^2} \{ \widehat{G}_{\mathbf{X}}(s, t) - G_{\mathbf{X}}(s, t) \}^2 ds dt, \qquad \delta_{\mathbf{k}}^{\mathbf{x}} = \min_{1 \le \mathbf{j} \le \mathbf{k}} (\lambda_{\mathbf{j}} - \lambda_{\mathbf{j}+1}), \qquad (32)$$
$$K_0 = \inf\{ j \ge \mathbf{1} : \lambda_{\mathbf{j}} - \lambda_{\mathbf{j}+1} \le 2D_{\mathbf{X}} \} - \mathbf{1}, \qquad \pi_{\mathbf{k}}^{\mathbf{x}} = \mathbf{1}/\lambda_{\mathbf{k}} + \mathbf{1}/\delta_{\mathbf{k}}^{\mathbf{x}}.$$

Let K = K(n) denote the numbers of leading eigenfunctions included to approximate X as sample size n varies; i.e.,  $\hat{X}_i(s) = \hat{\mu}_X(s) + \sum_{k=1}^{K} \hat{\xi}_{ik} \hat{\phi}_k(s)$ . Analogously, define the quantities  $\mathcal{G}_Y$ ,  $\hat{\mathcal{G}}_Y$ ,  $D_Y$ ,  $\delta_m^y \pi_m^y$ ,  $M_0$  and M for the process Y, for the case of functional responses. The following lemma gives the weak uniform convergence rates for the estimators of the FPCs, setting the stage for the subsequent developments. The proof is in Section 2.

Lemma 1 nder 
$$A$$
  $A$   $C$   $C$  \_nd  $C$   

$$\sup_{\mathbf{t}\in\mathcal{S}} |\hat{\mu}_{\mathbf{X}}(s) - \mu_{\mathbf{X}}(s)| = O_{\mathbf{p}}(\frac{1}{\sqrt{nb_{\mathbf{X}}}}), \quad \sup_{\mathbf{s}_{1},\mathbf{s}_{2}\in\mathcal{S}} |\hat{G}_{\mathbf{X}}(s_{1},s_{2}) - G_{\mathbf{X}}(s_{1},s_{2})| = O_{\mathbf{p}}(\frac{1}{\sqrt{nh_{\mathbf{X}}^{2}}}), (33)$$
and  $s$  consequence  $\hat{\sigma}_{\mathbf{X}}^{2} - \sigma_{\mathbf{X}}^{2} = O_{\mathbf{p}}(n^{-1/2}h_{\mathbf{X}}^{-2} + n^{-1/2}h_{\mathbf{X}}^{*-1})$  Considering eigenv\_ ues  $\lambda_{\mathbf{k}}$  of, u tip icity one  $\hat{\phi}_{\mathbf{k}}$  c\_n e chosen such th\_t

 $P(\sup_{1 \le \mathbf{k} \le \mathbf{K}_{0}} |\hat{\lambda}_{\mathbf{k}} - \lambda_{\mathbf{k}}| \le D_{\mathbf{X}}) = \mathbf{1}, \quad \sup_{\mathbf{s} \in S} |\hat{\phi}_{\mathbf{k}}(s) - \phi_{\mathbf{k}}(s)| = O_{\mathbf{p}}(\frac{\pi_{\mathbf{k}}^{\mathbf{x}}}{\sqrt{nh_{\mathbf{X}}^{2}}}), \quad k = 1, \dots, K_{0}, (34)$ here  $D_{\mathbf{X}} = \pi_{\mathbf{k}}^{\mathbf{x}} \ \text{and} \ K_{0} \ \text{are de ned in}$ ,  $An_{\mathbf{x}} \ ogous \ y \ under \ B = B, \quad C = C \quad \text{and} \ C$ 

 $\sup_{\mathbf{t}\in\mathcal{T}} |\hat{\mu}_{\mathbf{Y}}(t) - \mu_{\mathbf{Y}}(t)| = O_{\mathbf{p}}(\frac{1}{\sqrt{n}b_{\mathbf{Y}}}), \quad \sup_{\mathbf{t}_{1},\mathbf{t}_{2}\in\mathcal{T}} |\hat{G}_{\mathbf{Y}}(t_{1},t_{2}) - G_{\mathbf{Y}}(t_{1},t_{2})| = O_{\mathbf{p}}(\frac{1}{\sqrt{n}h_{\mathbf{Y}}^{2}}), \quad (35)$   $\neg nd \neg s \neg consequence \quad \hat{\sigma}_{\mathbf{Y}}^{2} - \sigma_{\mathbf{Y}}^{2} = O_{\mathbf{p}}(n^{-1/2}h_{\mathbf{Y}}^{-2} + n^{-1/2}h_{\mathbf{Y}}^{*-1}) \quad Considering \ eigenv \ ues$  $\rho_{\mathbf{m}} \ of \ u \ tip \ icity \ one \ \hat{\psi}_{\mathbf{m}} \ c_{-}n \ e \ chosen \ such \ th_{-}t$ 

$$P(\sup_{1 \le \mathbf{m} \le \mathbf{M}_0} |\hat{\rho}_{\mathbf{m}} - \rho_{\mathbf{m}}| \le D_{\mathbf{Y}}) = \mathbf{1}, \sup_{\mathbf{t} \in \mathcal{T}} |\hat{\psi}_{\mathbf{m}}(\mathbf{t}) - \psi_{\mathbf{m}}(\mathbf{t})| = O_{\mathbf{p}}(\frac{\pi_{\mathbf{k}}^{\mathbf{y}}}{\sqrt{n}h_{\mathbf{Y}}^2}), m = \mathbf{1}, \dots, M_0, (36)$$

here  $D_{\mathbf{Y}} = \pi_{\mathbf{k}}^{\mathbf{y}}$  and  $M_0$  are dened and ogous y to for process Y

Recall that  $||f||_{\infty} = \sup_{\mathbf{x} \in \mathcal{A}} |f(t)|$  for an arbitrary function f with support  $\mathcal{A}$ , and  $||g|| = \sqrt{\int_{\mathcal{A}} g^2(t) dt}$  for any  $g \in L^2(\mathcal{A})$  and define

$$\begin{aligned} \theta_{\mathbf{i}\mathbf{k}}^{(1)} &= c_1 \|X_{\mathbf{i}}\| + c_2 \|X_{\mathbf{i}}X_{\mathbf{i}}'\|_{\infty} \quad \overset{*}{\mathbf{x}} + c_3, \quad Z_{\mathbf{k}}^{(1)} = \sup_{\mathbf{s}\in\mathcal{S}} |\hat{\phi}_{\mathbf{k}}(s) - \phi_{\mathbf{k}}(s)|, \\ \theta_{\mathbf{i}\mathbf{k}}^{(2)} &= \mathbf{1} + \|\phi_{\mathbf{k}}\phi_{\mathbf{k}}'\|_{\infty} \quad \overset{*}{\mathbf{x}}, \qquad Z_{\mathbf{k}}^{(2)} = \sup_{\mathbf{s}\in\mathcal{S}} |\hat{\mu}_{\mathbf{x}}(s) - \mu_{\mathbf{x}}(s)|, \\ \theta_{\mathbf{i}\mathbf{k}}^{(3)} &= c_4 \|X_{\mathbf{i}}\|_{\infty} + c_5 \|X_{\mathbf{i}}'\|_{\infty} + c_6, \qquad Z_{\mathbf{k}}^{(3)} = \|\phi_{\mathbf{k}}'\|_{\infty} \quad \overset{*}{\mathbf{x}}, \\ \theta_{\mathbf{i}\mathbf{k}}^{(4)} &= |\sum_{\mathbf{j}=2}^{\mathbf{n}_i} \epsilon_{\mathbf{i}\mathbf{j}}\phi_{\mathbf{k}}(s_{\mathbf{i}\mathbf{j}})(s_{\mathbf{i}\mathbf{j}} - s_{\mathbf{i},\mathbf{j}-1})|, \quad Z_{\mathbf{k}}^{(4)} \equiv \mathbf{1}, \\ \theta_{\mathbf{i}\mathbf{k}}^{(5)} &= \sum_{\mathbf{j}=2}^{\mathbf{n}_i} |\epsilon_{\mathbf{i}\mathbf{j}}|(s_{\mathbf{i}\mathbf{j}} - s_{\mathbf{i},\mathbf{j}-1}), \\ \mathbf{k} & /, \end{aligned}$$

 $\psi t \phi$ 

**Lemma 2** For  $\theta_{ik}^{()} Z_k^{()} \vartheta_{im}^{()} \_nd Q_m^{()} \_s de ned in , \_nd ,$ 

$$|\hat{\xi}_{\mathbf{ik}}^{\mathbf{l}} - \xi_{\mathbf{ik}}| \le \sum_{=1}^{5} \theta_{\mathbf{ik}}^{(\,)} Z_{\mathbf{k}}^{(\,)}, \qquad |\hat{\zeta}_{\mathbf{im}}^{\mathbf{l}} - \zeta_{\mathbf{im}}| \le \sum_{=1}^{5} \vartheta_{\mathbf{im}}^{(\,)} Q_{\mathbf{m}}^{(\,)}.$$
(39)

The proof is in Section 2. In the sequel we suppress the superscript I in the FPC estimates  $\hat{\xi}_{ik}^{I}$  and  $\hat{\zeta}_{im}^{I}$ .

Recall that the sequences of bandwidths  $h_k$  and  $h_{mk}$  are employed to obtain the estimates  $\hat{f}_k$  and  $\hat{f}_{mk}$  for the regression functions  $f_k$  and  $f_{mk}$ , and that the density of  $\xi_k$  is denoted by  $p_k$ . Define

$$\theta_{\mathbf{k}}(x) = p_{\mathbf{k}}(x) \{ \frac{\pi_{\mathbf{k}}^{\mathbf{x}}}{\sqrt{n}h_{\mathbf{X}}^{2}} + \frac{1}{\sqrt{n}b_{\mathbf{X}}} + \sqrt{\frac{*}{\mathbf{x}}} \}, \\ \vartheta_{\mathbf{mk}}(x) = p_{\mathbf{k}}(x) \{ \frac{\pi_{\mathbf{m}}^{\mathbf{y}}}{\sqrt{n}h_{\mathbf{Y}}^{2}} + \frac{1}{\sqrt{n}b_{\mathbf{Y}}} + \sqrt{\frac{*}{\mathbf{Y}}} \}.$$
(40)

The weak convergence rates  $\tilde{\theta}_k$  and  $\tilde{\vartheta}_{mk}$  of the regression function estimators  $\hat{f}_k(x)$  and  $\hat{f}_{mk}(x)$  (see Theorem 1) are as follows,

$$\tilde{\theta}_{\mathbf{k}}(x) = \frac{\theta_{\mathbf{k}}(x)}{h_{\mathbf{k}}} + \frac{1}{2} |f_{\mathbf{k}}''(x)| h_{\mathbf{k}}^{2} + \sqrt{\frac{\operatorname{var}(Y|x) ||K_{1}||^{2}}{p_{\mathbf{k}}(x) n h_{\mathbf{k}}}}, \qquad (41)$$

$$\tilde{\vartheta}_{\mathbf{mk}}(x) = \frac{\theta_{\mathbf{k}}(x)}{h_{\mathbf{mk}}} + \vartheta_{\mathbf{mk}}(x) + \frac{1}{2} |f_{\mathbf{mk}}''(x)| h_{\mathbf{mk}}^{2} + \sqrt{\frac{\operatorname{var}(\zeta_{\mathbf{m}}|x) ||K_{1}||^{2}}{p_{\mathbf{k}}(x) n h_{\mathbf{mk}}}}.$$

Considering the predictions  $\widehat{E}(Y|X)$  for the scalar response case and  $\widehat{E}\{Y(t)|X\}$ for the functional response case, the numbers of eigenfunctions K and M used for approximating the infinite dimensional processes X and Y generally tend to infinity as the sample size n increases. We require  $K \leq K_0$  and  $M \leq M_0$  in (A6). Since it follows from (35) that  $K_0 \rightarrow \infty$ , as long as all eigenvalues  $\lambda_j$  are of multiplicity 1, and analogously for  $M_0$ , this is not a strong restriction. Denote the set of positive integers by  $\mathcal{N}$  and  $\mathcal{N}_k = \{1, \ldots, k\}$ . Convergence rates  $\theta_n^*$  and  $\vartheta_n^*$  for the predictions (20) and (21) are as follows,

$$\theta_{n}^{*} = \sum_{k=1}^{K} \left\{ \frac{\theta_{k}(\xi_{k})}{h_{k}} + \frac{1}{2} |f_{k}''(\xi_{k})|h_{k}^{2} + \sqrt{\frac{\operatorname{var}(Y|\xi_{k})||K_{1}||^{2}}{p_{k}(\xi_{k})nh_{k}}} \right\} + \left| \sum_{k \geq K+1} f_{k}(\xi_{k}) \right|,$$

$$\theta_{n}^{*} = \sum_{k=1}^{K} \sum_{m=1}^{M} \left\{ \left( \frac{\theta_{k}(\xi_{k})}{h_{mk}} + \vartheta_{mk}(\xi_{k}) \right) |\psi_{m}(t)| + \frac{1}{2} |f_{mk}''(\xi_{k})|\psi_{m}(t)|h_{mk}^{2} + \sqrt{\frac{\operatorname{var}(\zeta_{m}|\xi_{k})||K_{1}||^{2}}{p_{k}(\xi_{k})nh_{k}}} |\psi_{m}(t)| + \frac{\pi_{m}^{Y}|f_{mk}(\xi_{k})|}{p_{k}(\xi_{k})nh_{k}} + \frac{\pi_{m}^{Y}|f_{mk}(\xi_{k})|}{p_{k}(\xi_{k})nh_{k}}} + \frac{\pi_{m}^{Y}|f_{mk}(\xi_{k})|}{p_{k}(\xi_{k})nh_{k}} + \frac{\pi_{m}^{Y}|f_{mk}(\xi_{k})|}{p_{k}(\xi_{k})nh_{k}} + \frac{\pi_{m}^{Y}|f_{mk}(\xi_{k})|}{p_{k}(\xi_{k})nh_{k}} + \frac{\pi_{m}^{Y}|f_{mk}(\xi_{k})|}{p_{k}(\xi_{k})nh_{k}}} + \frac{\pi_{m}^{Y}|f_{mk}(\xi_{k})|}{p_{k}(\xi_{k})nh_{k}} + \frac{\pi_{m}^{Y}|f_{mk}(\xi_{k})nh_{k}} + \frac{\pi_{m}$$

Noting  $\hat{\xi}_{ik} = \hat{\eta}_{ik} + \hat{\tau}_{ik}$ , one finds

$$|\hat{\xi}_{i\mathbf{k}} - \xi_{i\mathbf{k}}| \le \{|\hat{\eta}_{i\mathbf{k}} - \tilde{\eta}_{i\mathbf{k}}| + |\tilde{\eta}_{i\mathbf{k}} - \xi_{i\mathbf{k}}| + |\hat{\tau}_{i\mathbf{k}}|\}.$$
(43)

Without loss of generality, assume  $\|\phi_k\|_{\infty} \ge 1$ ,  $\|\phi'_k\|_{\infty} \ge 1$ ,  $\|X_i\|_{\infty} \ge 1$  and  $\|X'_i\|_{\infty} \ge 1$ . For  $\theta_{ik}^{(\)}$  and  $Z_k^{(\)}$  (37),  $\ell = 1, \ldots, 5$ , the first term on the r.h.s. of (43) is bounded by

$$\begin{split} &\{\sum_{\mathbf{j}=2}^{\mathbf{n}_{i}} [|X_{\mathbf{i}}(s_{\mathbf{ij}}) - \hat{\mu}(s_{\mathbf{ij}})| \cdot |\hat{\phi}_{\mathbf{k}}(s_{\mathbf{ij}}) - \phi_{\mathbf{k}}(s_{\mathbf{ij}})| + |\hat{\mu}(s_{\mathbf{ij}}) - \mu(s_{\mathbf{ij}})| \cdot |\phi_{\mathbf{k}}(s_{\mathbf{ij}})|](s_{\mathbf{ij}} - s_{\mathbf{i,j}-1})\} \\ &\leq \{\sum_{\mathbf{j}=1}^{\mathbf{n}_{i}} [|X_{\mathbf{i}}(s_{\mathbf{ij}})| + |\mu(s_{\mathbf{ij}})| + 1]^{2} (s_{\mathbf{ij}} - s_{\mathbf{i,j}-1})\}^{1/2} \{\sum_{\mathbf{j}=2}^{\mathbf{n}_{i}} [\hat{\phi}_{\mathbf{k}}(s_{\mathbf{ij}}) - \phi_{\mathbf{k}}(s_{\mathbf{ij}})]^{2} (s_{\mathbf{ij}} - s_{\mathbf{i,j}-1})\}^{1/2} \\ &+ \{\sum_{\mathbf{j}=1}^{\mathbf{n}_{i}} [\hat{\mu}(s_{\mathbf{ij}}) - \mu(s_{\mathbf{ij}})]^{2} (s_{\mathbf{ij}} - s_{\mathbf{i,j}-1})\}^{1/2} \{\sum_{\mathbf{j}=2}^{\mathbf{n}_{i}} \phi_{\mathbf{k}}^{2} (s_{\mathbf{ij}}) (s_{\mathbf{ij}} - s_{\mathbf{i,j}-1})\}^{1/2} \\ &\leq \theta_{\mathbf{ik}}^{(1)} Z_{\mathbf{k}}^{(1)} + \theta_{\mathbf{ik}}^{(2)} Z_{\mathbf{k}}^{(2)}. \end{split}$$

The second term on the r.h.s. of (43) has the upper bound

$$|\tilde{\eta}_{\mathbf{i}\mathbf{j}} - \xi_{\mathbf{i}\mathbf{k}}| \le \| (X_{\mathbf{i}} + \mu)' \phi_{\mathbf{k}} + (X_{\mathbf{i}} + \mu) \phi_{\mathbf{k}}' \|_{\infty} \quad \overset{*}{\mathbf{X}} \le \theta_{\mathbf{i}\mathbf{k}}^{(3)} Z_{\mathbf{k}}^{(3)}.$$

From the above, the third term on the r.h.s. of (43) is bounded by  $(\theta_{ik}^{(4)}Z_k^{(4)} + \theta_{ik}^{(5)}Z_k^{(5)})$ .

Proof of Theore, For simplicity, denote " $\sum_{i=1}^{n}$ " by " $\sum_{i}$ ",  $w_i = K_1\{(x-\xi_{ik})/h_k\}/(nh_k)$ ,  $\hat{w}_i = K_1\{(x-\hat{\xi}_{ik})/h_k\}/(nh_k)$ , and write  $\theta_k = \theta_k(x)$ . From (12), the local linear estimator  $\hat{f}_k(x)$  of the regression function  $f_k(x)$  can be explicitly written as

$$\hat{f}_{\mathbf{k}}(t) = \frac{\sum_{\mathbf{i}} \hat{w}_{\mathbf{i}} Y_{\mathbf{i}}}{\sum_{\mathbf{i}} \hat{w}_{\mathbf{i}}} - \frac{\sum_{\mathbf{i}} \hat{w}_{\mathbf{i}} (\hat{\xi}_{\mathbf{ik}} - x)}{\sum_{\mathbf{i}} \hat{w}_{\mathbf{i}}} \hat{f}'_{\mathbf{k}}(x), \qquad (44)$$

where

$$\hat{f}'_{k}(x) = \frac{\sum_{i} \hat{w}_{i}(\hat{\xi}_{ik} - x)Y_{i} - \{\sum_{i} \hat{w}_{i}(\hat{\xi}_{ik} - x)\sum_{i} \hat{w}_{i}Y_{i}\}/\sum_{i} \hat{w}_{i}}{\sum_{i} \hat{w}_{i}(\hat{\xi}_{ik} - x)^{2} - \{\sum_{i} \hat{w}_{i}(\hat{\xi}_{ik} - x)\}^{2}/\sum_{i} \hat{w}_{i}}.$$
(45)

Let  $\tilde{f}_k(x)$  be a hypothetical estimator, obtained by substituting the true values  $w_i$  and  $\xi_{ik}$  for  $\hat{w}_i$ ,  $\hat{\xi}_{ik}$  in (44) and (45). To evaluate  $|\hat{f}_k(x) - \tilde{f}_k(x)|$ , one has to quantify the

orders of the di erences

$$D_1 = \sum_{\mathbf{i}} (\hat{w}_{\mathbf{i}} - w_{\mathbf{i}}), \quad D_2 = \sum_{\mathbf{i}} (\hat{w}_{\mathbf{i}} - w_{\mathbf{i}})Y_{\mathbf{i}},$$
$$D_3 = \sum_{\mathbf{i}} (\hat{w}_{\mathbf{i}}\hat{\xi}_{\mathbf{ik}} - w_{\mathbf{i}}\xi_{\mathbf{ik}}), \quad D_4 = \sum_{\mathbf{i}} (\hat{w}_{\mathbf{i}}\hat{\xi}_{\mathbf{ik}}^2 - w_{\mathbf{i}}\xi_{\mathbf{ik}}^2)$$

Considering  $D_1$ , without loss of generality, assume the compact support of  $K_1$  is [-1, 1]. Since  $K_1$  is Lipschitz continuous on its support,

$$D_1 \leq \frac{c}{nh_{\mathbf{k}}^2} \sum_{\mathbf{i}} |\hat{\xi}_{\mathbf{i}\mathbf{k}} - \xi_{\mathbf{i}\mathbf{k}}| \{ I(|x - \xi_{\mathbf{i}\mathbf{k}}| \leq h_{\mathbf{k}}) + I(|x - \hat{\xi}_{\mathbf{i}\mathbf{k}}| \leq h_{\mathbf{k}}) \},$$
(46)

for some c > 0, where  $I(\cdot)$  is an indicator function. Lemma 2 implies for the first term on the r.h.s. of (46)

$$\frac{1}{nh_{\mathbf{k}}^{2}}\sum_{\mathbf{i}}|\hat{\xi}_{\mathbf{i}\mathbf{k}}-\xi_{\mathbf{i}\mathbf{k}}|I(|x-\xi_{\mathbf{i}\mathbf{k}}|\leq h_{\mathbf{k}})\leq \sum_{i=1}^{5}Z_{\mathbf{k}}^{(\cdot)}\frac{1}{nh_{\mathbf{k}}^{2}}\sum_{\mathbf{i}}\theta_{\mathbf{i}\mathbf{k}}^{(\cdot)}I(|x-\xi_{\mathbf{i}\mathbf{k}}|\leq h_{\mathbf{k}})$$

Applying the central limit theorem for a random number of summands (Billingsley, 1995, page 380), observing  $\sum_{i} I(|x - \xi_{ik}| \le h_k)/(nh_k) \xrightarrow{p} 2p_k(x)$ , one finds

$$\frac{1}{nh_{\mathbf{k}}} \sum_{\mathbf{i}} \theta_{\mathbf{i}\mathbf{k}}^{()} I(|x - \xi_{\mathbf{i}\mathbf{k}}| \le h_{\mathbf{k}}) \xrightarrow{\mathbf{p}} 2p_{\mathbf{k}}(x) E(\theta_{\mathbf{i}\mathbf{k}}^{()}),$$
(47)

provided that  $E(\theta_{ik}^{()}) < \infty$  for  $\ell = 1, ..., 5$ . Note that  $E\theta_{ik}^{(1)} < \infty$ ,  $E\theta_{ik}^{(3)} < \infty$  by (A4),  $E\theta_{ik}^{(4)} \leq 2\sigma_X \sqrt{\frac{*}{X}}$  and  $E\theta_{ik}^{(5)} \leq |S|\sigma_X$  by the Cauchy-Schwarz inequality. Then

$$Z_{\mathbf{k}}^{(1)} \frac{1}{nh_{\mathbf{k}}^{2}} \sum_{\mathbf{i}} \theta_{\mathbf{i}\mathbf{k}}^{(1)} I(|x - \xi_{\mathbf{i}\mathbf{k}}| \le h_{\mathbf{k}}) = O_{\mathbf{p}}\{\frac{\pi_{\mathbf{k}}^{\mathbf{x}}}{\sqrt{nh_{\mathbf{X}}^{2}h_{\mathbf{k}}}} p_{\mathbf{k}}(x)\},$$

$$Z_{\mathbf{k}}^{(2)} \frac{1}{nh_{\mathbf{k}}^{2}} \sum_{\mathbf{i}} \theta_{\mathbf{i}\mathbf{k}}^{(2)} I(|x - \xi_{\mathbf{i}\mathbf{k}}| \le h_{\mathbf{k}}) = O_{\mathbf{p}}\{\frac{1}{\sqrt{nb_{\mathbf{X}}h_{\mathbf{k}}}} p_{\mathbf{k}}(x)\},$$

$$Z_{\mathbf{k}}^{(3)} \frac{1}{nh_{\mathbf{k}}^{2}} \sum_{\mathbf{i}} \theta_{\mathbf{i}\mathbf{k}}^{(3)} I(|x - \xi_{\mathbf{i}\mathbf{k}}| \le h_{\mathbf{k}}) = O_{\mathbf{p}}\{\frac{\|\phi_{\mathbf{k}}\|_{\infty}}{h_{\mathbf{k}}} p_{\mathbf{k}}(x)\},$$

$$Z_{\mathbf{k}}^{(4)} \frac{1}{nh_{\mathbf{k}}^{2}} \sum_{\mathbf{i}} \theta_{\mathbf{i}\mathbf{k}}^{(4)} I(|x - \xi_{\mathbf{i}\mathbf{k}}| \le h_{\mathbf{k}}) = O_{\mathbf{p}}\{\frac{\sqrt{\frac{\mathbf{x}}{\mathbf{x}}}}{h_{\mathbf{k}}} p_{\mathbf{k}}(x)\},$$

$$Z_{\mathbf{k}}^{(5)} \frac{1}{nh_{\mathbf{k}}^{2}} \sum_{\mathbf{i}} \theta_{\mathbf{i}\mathbf{k}}^{(5)} I(|x - \xi_{\mathbf{i}\mathbf{k}}| \le h_{\mathbf{k}}) = O_{\mathbf{p}}\{\frac{\pi_{\mathbf{k}}^{\mathbf{x}}}{\sqrt{nh_{\mathbf{X}}^{2}h_{\mathbf{k}}}} p_{\mathbf{k}}(x)\}.$$
(48)

We now obtain  $(nh_k^2)^{-1} \sum_i |\hat{\xi}_{ik} - \xi_{ik}| I(|x - \xi_{ik}| \le h_k) = O_p(\theta_k h_k^{-1})$ . The asymptotic rate of the second term can be derived analogously, observing

$$\frac{1}{nh_{\mathbf{k}}}\sum_{\mathbf{i}}I(|x-\hat{\xi}_{\mathbf{i}\mathbf{k}}| \le h_{\mathbf{k}}) \le \frac{1}{nh_{\mathbf{k}}}\sum_{\mathbf{i}}\{I(|x-\xi_{\mathbf{i}\mathbf{k}}| \le 2h_{\mathbf{k}}) + I(\sum_{i=1}^{5}\theta_{\mathbf{i}\mathbf{k}}^{(i)}Z_{\mathbf{k}}^{(i)} > h_{\mathbf{k}})\} \xrightarrow{\mathbf{p}} 4p_{\mathbf{k}}(x)$$
  
leading to  $(nh_{\mathbf{k}}^{2})^{-1}\sum_{\mathbf{i}}|\hat{\xi}_{\mathbf{i}\mathbf{k}} - \xi_{\mathbf{i}\mathbf{k}}|I(|x-\hat{\xi}_{\mathbf{i}\mathbf{k}}| \le h_{\mathbf{k}}) = O_{\mathbf{p}}(\theta_{\mathbf{k}}h_{\mathbf{k}}^{-1}).$  Then  $D_{1} = O_{\mathbf{p}}(\theta_{\mathbf{k}}h_{\mathbf{k}}^{-1})$  follows.

Analogously, one shows  $D_2 = O_p(\theta_k h_k^{-1})$ , applying the Cauchy-Schwarz inequality for  $\theta_{ik}^{()}$ ,  $\ell = 1, 3$ , and observing the independence between  $Y_i$  and  $\theta_{ik}^{()}$  for  $\ell = 2, 4, 5$ , given the moment condition (A4). For  $D_3$ , observe

$$D_{3} = \sum_{i} \{ (\hat{w}_{i} - w_{i})\xi_{ik} + (\hat{w}_{i} - w_{i})(\hat{\xi}_{ik} - \xi_{ik}) + w_{i}(\hat{\xi}_{ik} - \xi_{ik}) \} \equiv D_{31} + D_{32} + D_{33}$$

Then  $D_{31} = O_p(\theta_k h_k^{-1})$ , analogously to  $D_1$ . It is easy to see that  $D_{32} = o_p(D_{31})$ . Since  $D_{33} \le c \sum_{i=1}^5 Z_k^{(i)} (nh_k)^{-1} \sum_i \theta_{ik}^{(i)} I(|x - \xi_{ik}| \le h_k)$  for some c > 0, one also has  $D_{33} = o_p(D_{31})$ . This results in  $D_3 = O_p(\theta_k h_k^{-1})$ . Observing  $|\hat{\xi}_{ik}^2 - \xi_{ik}^2| \le |\hat{\xi}_{ik} - \xi_{ik}| \cdot |\xi_{ik}| + (\hat{\xi}_{ik} - \xi_{ik})^2$ , one can show  $D_4 = O_p(\theta_k h_k^{-1})$ , using similar arguments as for  $D_3$ , and  $E\xi_{ik}^4 < \infty$  from (A4). Combining the results for  $D_i$ ,  $\ell = 1, \ldots, 4$ , and applying Slutsky's Theorem leads to  $|\hat{f}_k(x) - \tilde{f}_k(x)| = O_p(\theta_k h_k^{-1})$ . Using (A5), and applying standard asymptotic results for the hypothetical local linear smoother  $\tilde{f}_k(x)$  completes the proof of (18).

To derive (19), additionally one only needs to consider  $\sum_{i} (\hat{w}_{i}\hat{\zeta}_{im} - w_{i}\zeta_{im}) = \sum_{i} (\{\hat{w}_{i} - w_{i})\zeta_{im} + (\hat{w}_{i} - w_{i})(\hat{\zeta}_{im} - \zeta_{im}) + w_{i}(\hat{\zeta}_{im} - \zeta_{im})\}$ , where the third term yields an extra term of order  $O_{p}(\vartheta_{mk})$  by observing

$$|\sum_{\mathbf{i}} w_{\mathbf{i}}(\hat{\zeta}_{\mathbf{im}} - \zeta_{\mathbf{im}})| \leq \sum_{i=1}^{5} Q_{\mathbf{m}}^{(i)} \sum_{\mathbf{i}} w_{\mathbf{i}} \vartheta_{\mathbf{im}}^{(i)} \leq \frac{1}{nh_{\mathbf{mk}}} \sum_{i=1}^{5} Q_{\mathbf{m}}^{(i)} \sum_{\mathbf{i}} \vartheta_{\mathbf{im}}^{(i)} I(|x - \xi_{\mathbf{ik}}| \leq h_{\mathbf{mk}})$$

Similar arguments as above complete this derivation.

*Proof of Theore,* Using (A7), the derivation of  $\theta_n^*$  in (42) is straightforward,

following the above arguments. To obtain (21), note that

$$\hat{E}\{Y(t)|X\} - E\{Y(t)|X\} \\
\leq \sum_{k=1}^{K} \sum_{m=1}^{M} |\hat{f}_{mk}(\xi_{k})\hat{\psi}_{m}(t) - f_{mk}(\xi_{k})\psi_{m}(t)| + |\sum_{k\geq K+1} \sum_{m\geq M+1} f_{mk}(\xi_{k})\psi_{m}(t)| \\
\leq \sum_{k=1}^{K} \sum_{m=1}^{M} [|\hat{f}_{mk}(\xi_{k}) - f_{mk}(\xi_{k})|\{|\psi_{m}(t)| + |\hat{\psi}_{m}(t) - \psi_{m}(t)|\} + |f_{mk}(\xi_{k})| \cdot |\hat{\psi}_{m}(t) - \psi_{m}(t)|] \\
+ |\sum_{(k,m)\in\mathcal{N}^{2}\setminus\mathcal{N}_{K}\times\mathcal{N}_{M}} f_{mk}(\xi_{k})\psi_{m}(t)|.$$

This implies the convergence rate  $\vartheta_n^*$  in (42).