Functional Additive Models

Hans-Georg MÜLLER and Fang YAO

1. INTRODUCTION

functional additive model (M)

M

Drosophila

Y ζ_m β K M

3. FUNCTIONAL ADDITIVE MODELING

ξ_{ik}

$$E(Y-\mu_Y|\xi_k)=b_k\xi_k \qquad E(\zeta_m|\xi_k)=b_{km}\xi_k \qquad (\ )$$

$b_k\xi_k$ $b_{km}\xi_k$ ξ_k $f_k(\xi_k)$ $f_k(\cdot)$ $k=\ ,\ ,\ldots$ $f_{km}(\cdot)$ $k,m=\ ,\ ,\ldots$ ξ_k

$$E(Y|X)=\mu_Y+\sum_{k=}^{\infty} f_k(\xi_k) \qquad (\ )$$

$$E(Y(t)|X)=\mu_Y(t)+\sum_{k=}^{\infty}\sum_{m=}^{\infty} f_{km}(\xi_k)\psi_m(t), \qquad (\ )$$

$$Ef_k(\xi_k)=\ , \quad k=\ ,\ ,\ldots,$$
$$Ef_{km}(\xi_k)=\ , \quad k=\ ,\ ,\ldots,m=\ ,\ ,\ldots. \qquad (\ )$$

Y ξ_k M

$$E(Y-\mu_Y|\xi_k)=E\{E(Y-\mu_Y|X)|\xi_k\}$$
$$=E\left\{\sum_{j=}^{\infty} f_j(\xi_j)\Big|\xi_k\right\}=f_k(\xi_k), \qquad (\ )$$

$$E(\zeta_m|\xi_k)=E\{E(\zeta_m|X)|\xi_k\}$$
$$=E\left\{\sum_{j=}^{\infty} f_{jm}(\xi_j)\Big|\xi_k\right\}=f_{km}(\xi_k). \qquad (\ )$$

$f_k(\xi_k)=E(Y-\mu_Y|\xi_k)$ $f_{km}(\xi_k)=E(\zeta_m|\xi_k)$ $k,m=\ ,\ ,\ldots,$

f_k f_{km} M

4. FITTING OF FUNCTIONAL ADDITIVE MODELS

$\hat\varphi_k$ $\hat\lambda_k$ $\hat\xi_{ik}$

principal analysis by conditional expectation (PACE) K

$\hat f_k$ $\hat f_{km}$ f_k f_{km} f_k $\{\hat\xi_{ik},Y_i\}_{i=\ ,\ldots,n}$ $\hat\xi_{ik}$

$$\sum_{i=}^{n} K\left(\frac{\hat\xi_{ik}-x}{h_k}\right)\{Y_i-\beta\ -\beta\ (x-\hat\xi_{ik})\} \qquad (\ )$$

β β $\hat f_k(x)=\hat\beta\ (x)-\bar Y$ h_k K f_{mk}

$\{\hat\xi_{ik}, \hat\zeta_{im}\}_{i=\ ,\dots,n}$ ( )

$$\sum_{i=}^{n} K\left(\frac{\hat\xi_{ik}-x}{h_{mk}}\right)\{\hat\zeta_{im}-\beta \ -\beta \ (x-\hat\xi_{ik})\} \qquad (\ )$$

β β $\hat f_{mk}(x)=\hat\beta\ (x)$. h_{mk} M ( )

$$\widehat{E}(Y|X)=\bar Y+\sum_{k=}^{K}\hat f_k(\xi_k). \qquad (\ )$$

R ( )

$$R \ = \ -\frac{\sum_{i=}^{n}\{Y_i-E(Y_i|X_i)\}}{\sum_{i=}(Y_i-\mu_Y)}$$
$$\widehat{R} \ = \ -\frac{\sum_{i=}^{n}\{Y_i-\widehat{E}(Y_i|X_i)\}}{\sum_{i=}(Y_i-\bar Y)}, \qquad (\ )$$

$E(Y_i|X_i)$ $\widehat{E}(Y_i|X_i)$ i ( ) ( ).

M ( )

$$\widehat{E}\{Y(t)|X\}=\hat\mu_Y(t)+\sum_{m=}^{M}\sum_{k=}^{K}\hat f_{mk}(\xi_k)\hat\psi_m(t), \qquad t\in\mathcal{T}, \qquad (\ )$$

$$R \ = \ -\frac{\sum_{i=}^{n}\int [Y_i(t)-E\{Y_i(t)|X_i\}]\ dt}{\sum_{i=}^{n}\int \{Y_i(t)-\mu_Y(t)\}\ dt},$$
$$\widehat{R} \ = \ -\frac{\sum_{i=}^{n}\sum_{l=}^{m_i}[V_{il}-\widehat{E}\{Y_i(t_{il})|X_i\}]\ (t_{il}-t_{i,l-}\ )}{\sum_{i=}^{n}\sum_{l=}^{m_i}\{V_{il}-\mu_Y(t_{il})\}\ (t_{il}-t_{i,l-}\ )} \qquad (\ )$$

$E\{Y_i(t)|X_i\}$ $\widehat{E}(Y_i(t)|X_i)$ i ( ) ( ) t_{il}

5. THEORETICAL RESULTS

E ξ_{ik} $\hat\xi_{ik}$ ζ_{im} $\hat\zeta_{im}$ $k=\ ,\dots,K$ $m=\ ,\dots,M$ M f_k f_{km}

( ) ( ) (H H) $|\hat\xi_{ik}-\xi_{ik}|$ $|\hat\zeta_{im}-\zeta_{im}|$ i i $i=\ ,\dots,n$ ( ). f_k ( ) f_{mk} ( )

$\{\hat\xi_{ik}, Y_i\}$ $\{\hat\xi_{ik}, \hat\zeta_{im}\}$ $i=\ ,\dots,n$. $n\to\infty$ K M $K=K(n)\to\infty$ $M=M(n)\to\infty$ ( ). ( ) ( ) ( ) ( )

Theorem 1. U ( ) ( ) ( ) ( ) ( ) ( ) ( ) $k\geq$ $\lambda_j, j\leq k$

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

( ). U ( ) ( ) ( ) k m $\lambda_j, j\leq k$ $\rho_l, l\leq m$

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

( ).

$\tilde\theta_k(x)=|\hat f_k(x)-f_k(x)|$ $\tilde\vartheta_{mk}(x)=|\hat f_{mk}(x)-f_{mk}(x)|$ ( ). M.

Theorem 2. U ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )

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

$\widehat{E}(Y|X)=\bar Y+\sum_{k=}^{K}\hat f_k(\xi_k)$ ( ). U ( ) ( ) ( ) ( ) ( ) $t\in\mathcal{T}$

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

$\widehat{E}\{Y(t)|X\}=\hat\mu_Y(t)+\sum_{k=}^{K}\sum_{m=}^{M}\hat f_{mk}(\xi_k)\hat\psi_m(t)$ $\hat f_{mk}(\xi_k)$ ( ).

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

6. SIMULATION STUDIES

M ( ) ( ) $n=$ X_i $\mu_X(s)=s+\ (s)$ $\leq s\leq$ $\varphi\ (s)=-\ (\pi s/\ )/\sqrt{\ }$ $\varphi\ (s)=\ (\pi s/\ )/\sqrt{\ }$ $\leq s\leq$ $\lambda\ =$ $\lambda\ =$ $\lambda_k=$ $k\geq$ ( ) $_{ij}\sim\ (\ ,\ )$. ( ) $\xi_{ik}\sim\mathcal{N}(\ ,\lambda_k)$ ( ) $\xi_{ik}=\sqrt{\lambda_k}(Z_{ik}-\ )/$ $Z_{ik}\sim$ (,) $k=\ ,$

[illegible] $[\;,\;]$ [illegible] $\{\;,\dots,\;\}$ [illegible] $\{\;,\dots,\;\}$ [illegible] X_i^* [illegible] U_{ij}^* [illegible] U_{ij} [illegible] (Y_i^*). [illegible] $Y_i = \sum_{k=} f_k(\xi_{ik}) + \varepsilon_i$ [illegible] f_k, [illegible] ξ_{ik} [illegible] X_i [illegible] Y_i [illegible] $\varepsilon_i \sim (\;,.\;)$ [illegible] $\mu_Y =$ [illegible] $\sigma_Y = .$ [illegible] M [illegible] $f_k(x) = x - \lambda_k$ [illegible] $k = \;,$ [illegible] $f\;(x) = x$ [illegible] $f\;(x) = x/$ [illegible] f_k [illegible] $\hat f_k(x) = \sum_{i=}^n (Y_i - \bar Y)(\hat\xi_{ik} - \bar{\hat\xi}_{\cdot k})\hat\lambda_k^{-} x$ [illegible] $\bar Y_i = \sum_{i=}^n Y_i/n$, $\bar{\hat\xi}_{\cdot k} = \sum_{i=}^n \hat\xi_{ik}/n$, $\hat\lambda_k$ [illegible] k [illegible] X [illegible] $\hat\xi_{ik}$, [illegible] E [illegible] M [illegible] f_k [illegible] M [illegible]

[illegible] f_{km} [illegible] $\hat f_{km}(x) = \sum_{i=}^n (\hat\zeta_{im} - \bar\zeta_{\cdot m})(\hat\xi_{ik} - \bar{\hat\xi}_{\cdot k})\hat\lambda_k^{-} x$. [illegible] $Y_i(t) = \mu_Y(t) + \zeta_i\,\psi\,(t)$ [illegible] $\mu_Y(t) = t + \;(t)$, $\psi\,(t) = -\;(\pi t/\;)/\sqrt{\;}$ [illegible] ζ_i [illegible] $\le t \le$ [illegible] ζ_i [illegible] $f_{km}, k = \;,\;, m = \;(\;)$ [illegible] $f_k\,(x) = x - \lambda_k$ [illegible] $k = \;,$ [illegible] $\zeta_i = \sum_{k=} (\xi_{ik} - \lambda_k)$ [illegible] $f\;(x) = x$, $f\;(x) = x/$ [illegible] $\zeta_i = \xi_i + \xi_i\,/$ [illegible] ε_{il} [illegible] $\mathcal{N}(\;,.\;)$. [illegible] X_i^* [illegible] Y_i^* [illegible] U_{ij}^* [illegible] V_{il}^* [illegible] U_{ij} [illegible] V_{il}. [illegible] (E)

$$\mathrm{E}_i = \frac{(Y_i^* - \hat Y_i^*)}{Y_i^*}$$

$$\mathrm{E}_{i,f} = \frac{\int (Y_i^*(t) - \hat Y_i^*(t))\,dt}{\int Y_i^*(t)\,dt}, \quad (\;)$$

[illegible] E [illegible] M [illegible] f_k [illegible] f_{km} [illegible] M [illegible] M [illegible] M [illegible] M [illegible]

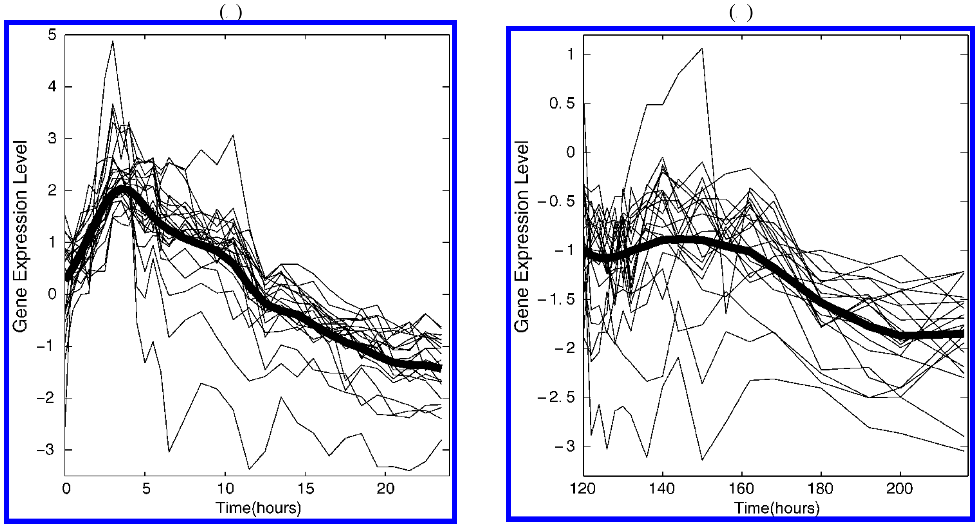
	M								
	M	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	M	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	M	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	M	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

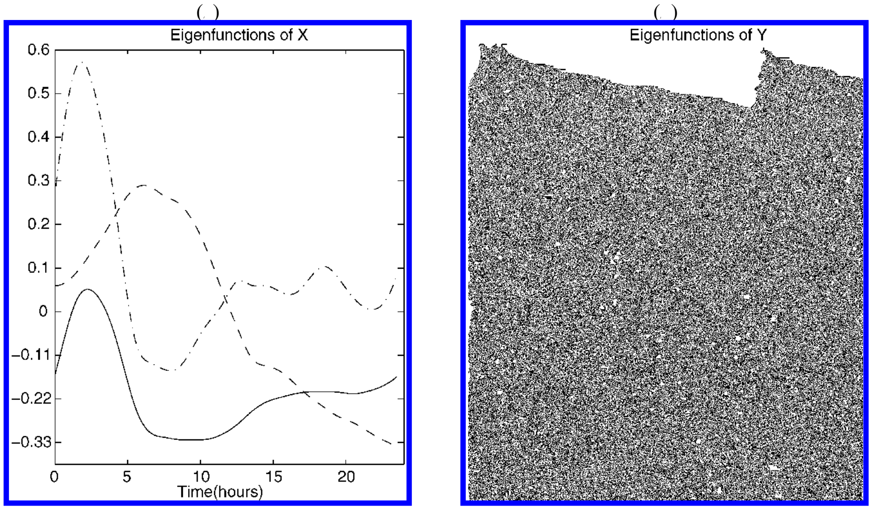
E [illegible] M [illegible] $n =$ [illegible] M [illegible] E (). [illegible]

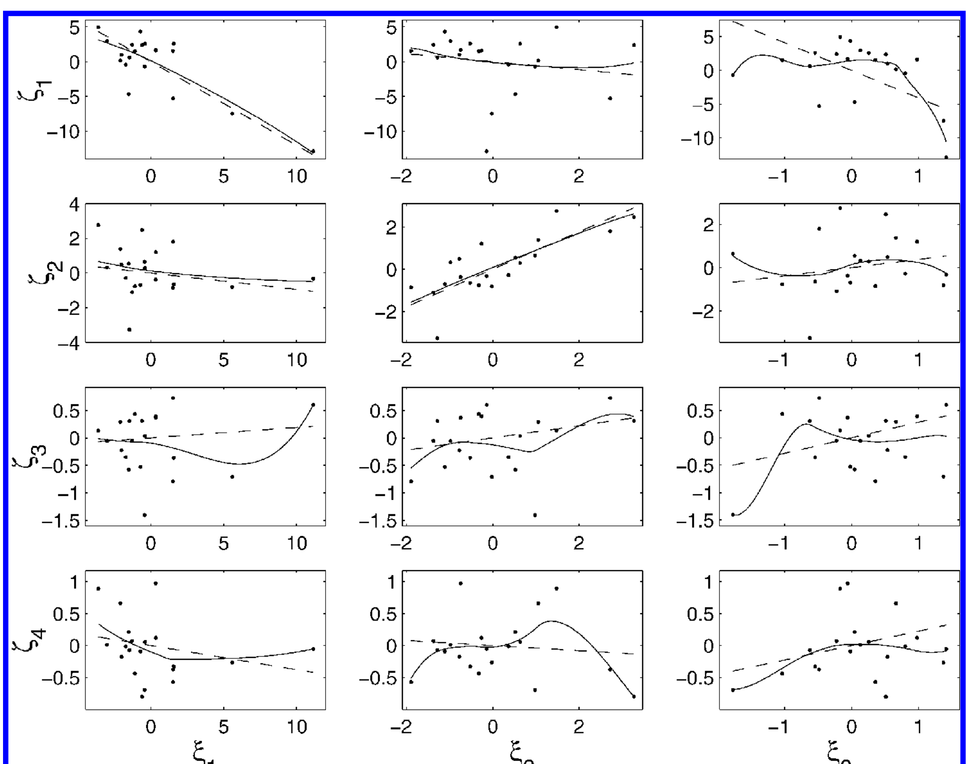
7. APPLICATION TO GENE EXPRESSION TIME COURSE DATA

Drosophila

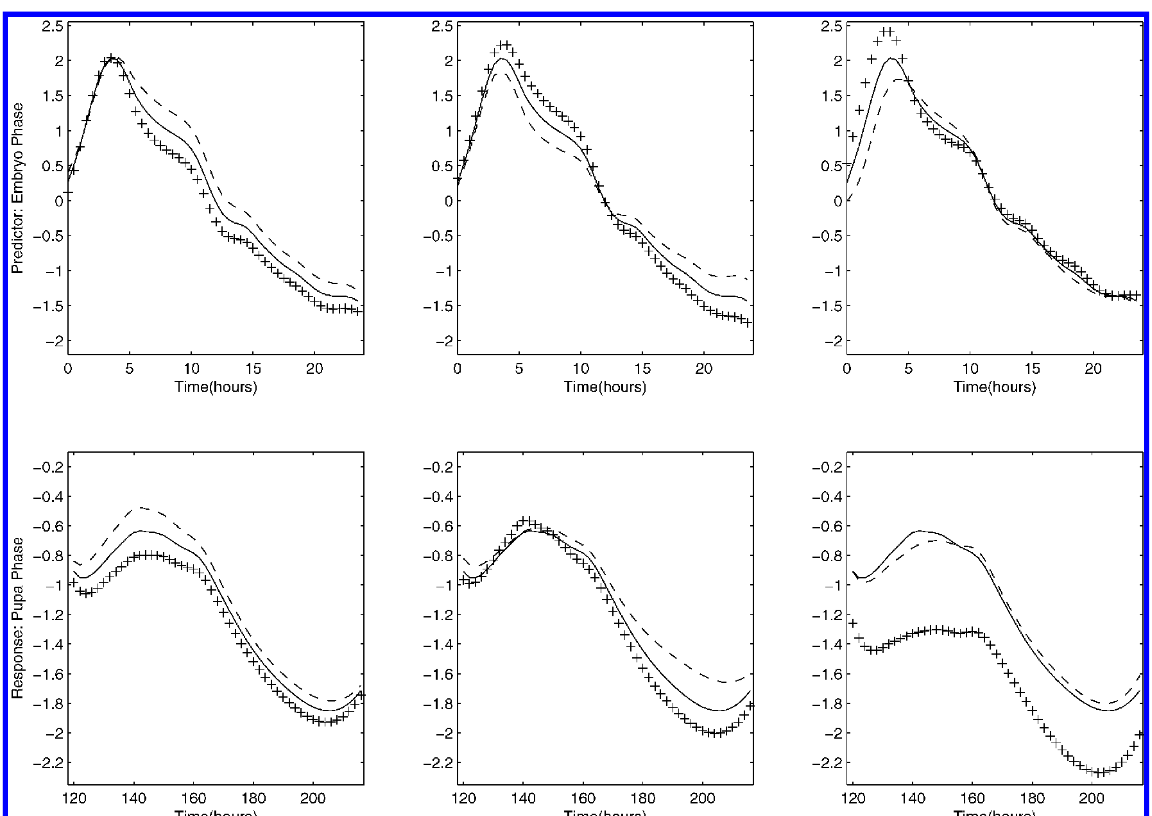
$n =$




Eigenfunctions of X
0.6
0.5
0.4
0.3
0.2
0.1
0
−0.11
−0.22
−0.33
0
5
10
15
20
Time(hours)
Eigenfunctions of Y


ζ1
ζ2
ζ3
ζ4
ξ1
ξ2
ξ3

8. CONCLUDING REMARKS



$\hat{\mu}_X(s)+\alpha\hat{\varphi}_k(s)$ $\alpha =$ ( ) $\alpha =$ (−−) $\alpha = -$ (++) $k =$ ( ) $k =$ ( ) $k =$ ( ) $\hat{\mu}_Y(t)+\sum_{m=}^{M}\{\hat{f}_{km}(\alpha)+\sum_{\neq k}\hat{f}_{m}(\,)\}\hat{\psi}_m(t)$

APPENDIX: ESTIMATION PROCEDURES AND ASSUMPTIONS

A.1 Eigenrepresentations and Estimating Functional Principal Component Scores

$\varphi_k, k = \;,\;, \ldots,$ X $(A_{G_X} f)(s) = \lambda f(s)$ $f \in L\;(S)$ $(A_{G_X} f)(t) = \int f(s) \times G_X(s,t)\,ds$ H $\int \varphi_j(s)\varphi_k(s)\,ds = \delta_{jk}$ $\delta_{jk} = \;\; j = k \;\; = \;\; j \neq k$. φ_j $\lambda \geq \lambda \geq \cdots$. Y ψ_m ρ_m. $\{\mu_X, \varphi_j\}$ $\{\mu_Y, \psi_k\}$ $G_X(s\;,s\;) = \sum_k \lambda_k \times \varphi_k(s\;)\varphi_k(s\;)$ $G_Y(t\;,t\;) = \sum_m \rho_m \psi_m(t\;)\psi_m(t\;)$ X Y K X Y

$$X(s) = \mu_X(s) + \sum_{j=}^{\infty} \xi_j \varphi_j(s)$$

$$Y(t) = \mu_Y(t) + \sum_{k=}^{\infty} \zeta_k \psi_k(t). \qquad (\;.\;)$$

$\{\varphi_k, k = \;,\;, \ldots\}$ $\{\psi_m, m = \;,\;, \ldots\}$ β_k β_{km} $\beta(s) = \sum_{k=}^{\infty} \beta_k \varphi_k(t)$ $\beta(s,t) = \sum_{k=}^{\infty} \sum_{m=}^{\infty} \beta_{km}\varphi_k(s)\psi_m(t)$. (.) ( ) ( )

$$\begin{aligned} U_{ij} &= X_i(s_{ij}) + \;_{ij} \\ &= \mu_X(s_{ij}) + \sum_{k=}^{\infty} \xi_{ik}\varphi_k(s_{ij}) + \;_{ij}, \\ & s_{ij} \in S, \;\; \leq i \leq n, \;\; \leq j \leq n_i, \end{aligned} \qquad (\;.\;)$$

$$\begin{aligned} V_{il} &= Y_i(t_{il}) + \varepsilon_{il} \\ &= \mu_Y(t_{il}) + \sum_{m=}^{\infty} \zeta_{im}\psi_k(t_{il}) + \varepsilon_{il}, \\ & s_{il} \in T, \;\; \leq i \leq n, \;\; \leq l \leq m_i. \end{aligned} \qquad (\;.\;)$$

M $\mathbf{U}_i = (U_i\;, \ldots, U_{in_i})^T$ $_{X_i} = (\mu_X(s_i\;), \ldots, \mu_X(s_{in_i}))^T$ $_{ik} = (\varphi_k(s_i\;), \ldots, \varphi_k(s_{in_i}))^T$ ξ_{ik} λ_k $^T_{ik}\boldsymbol{\Sigma}^-_{U_i}(\mathbf{U}_i - \;_{X_i})$ (j,l) $n_i \times n_i$ $\boldsymbol{\Sigma}_{U_i}$ $(\boldsymbol{\Sigma}_{U_i})_{j,l} = G_X(s_{ij}, s_{il}) + \sigma_X \delta_{jl}$ $\delta_{jl} =$ $j = l$ $\delta_{jl} =$ $j \neq l$. ξ_{ik} $_{X_i}$ λ_k $_{ik}$ $\boldsymbol{\Sigma}_{X_i}$ ( )

$$\hat{\xi}_{ik} = \hat{\lambda}_k \hat{\;}^T_{ik} \widehat{\boldsymbol{\Sigma}}^-_{U_i}(\mathbf{U}_i - \hat{\;}_{X_i}), \qquad (\;.\;)$$

(j,l) $\widehat{\boldsymbol{\Sigma}}_{U_i}$ $(\widehat{\boldsymbol{\Sigma}}_{U_i})_{j,l} = \widehat{G}_X(s_{ij}, s_{il}) + \hat{\sigma}_Y(s_{ij})\delta_{jl}$. M ( ) (E) $\hat{\xi}_{ik}, \hat{\zeta}_{im}$ (.)

$$\hat{\xi}^I_{ik} = \sum_{j=}^{n_i} (U_{ij} - \hat{\mu}_X(t_{ij}))\hat{\varphi}_k(s_{ij})(s_{ij} - s_{i,j-}\;),$$

$$\hat{\zeta}^I_{im} = \sum_{j=}^{n_i} (V_{ij} - \hat{\mu}_Y(t_{ij}))\hat{\psi}_k(t_{ij})(t_{ij} - t_{i,j-}\;), \qquad (\;.\;)$$

$\xi_{ik} = \int\{X_i(s) - \mu_X(s)\}\varphi_k(s)\,ds$ $\zeta_{im} = \int\{Y_i(t) - \mu_Y(t)\}\psi_m(t)\,dt$. E

A.2 Estimation Procedures

( ) E ( )E μ_X

$$\sum_{i=}^{n}\sum_{j=}^{n_i} K\left(\frac{s_{ij} - s}{b_X}\right)\{U_{ij} - \beta^X - \beta^X(s - s_{ij})\} \qquad (\;.\;)$$

β^X β^X $\hat{\mu}_X(s) = \hat{\beta}^X(s)$. K b_X E $G_X(s\;,s\;)$ X $^X_i(s_{ij}\;, s_{ij}\;) = (U_{ij}\; - \hat{\mu}_X(s_{ij}\;))(U_{ij}\; - \hat{\mu}_X(s_{ij}\;))$

$$\sum_{i=}^{n} \sum_{\leq j \neq j \leq n_i} K\left(\frac{s_{ij}\; - s}{h_X}, \frac{s_{ij}\; - s}{h_X}\right) \times \left\{\;^X_i(s_{ij}\;, s_{ij}\;) - f\left(\beta^X, (s\;,s\;), (s_{ij}\;, s_{ij}\;)\right)\right\}, \qquad (\;.\;)$$

$f(\beta^X, (s\;,s\;), (s_{ij}\;, s_{ij}\;)) = \beta^X + \beta^X(s\; - s_{jl}\;) + \beta^X(s\; - s_{ij}\;)$ $^X = (\beta^X, \beta^X, \beta^X)^T$ $\widehat{G}_X(s\;,s\;) = \hat{\beta}^X(s\;,s\;)$. H K h_X $j = j$ σ_X G_X $\widetilde{G}_X(s)$ $\{G_X(s,s) + \sigma_X\}$ $\widehat{V}_X(s)$

h^*_X $a = \{s\ s \in S\}$ $b = \{s\ s \in S\}$
$|S| = b - a$ $S = [a + |S|/\ , b - |S|/\ ]$ σ_X

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

$\hat{\sigma}_X >$ $\hat{\sigma}_X =$ E
$\{\lambda_k, \varphi_k\}_{k\geq}$ $\{\hat{\lambda}_k, \{$

ξ_k $t \in T$

( ) $\sum_{k=}^{K} [|f_k''(\xi_k)|h_k + n^{- /} \{ (Y|\xi_k)\}^{/} p_k^{- /}(\xi_k) \times h_k^{- /}] \to$

( ) $\sum_{k=}^{K} \sum_{m=}^{M} [|f_{mk}''(\xi_k)\psi_m(t)|h_{mk} + n^{- /} \{ (\zeta_m|\xi_k)\}^{/} \times p_k^{- /}(\xi_k)|\psi_m(t)|h_{mk}^{- /}] \to$

[Received August 2007. Revised June 2008.]

REFERENCES

M E E M E H K M M K ( ) E *Drosophila melanogaster Science*

M ( ) *Topics in Stochastic Processes*

( ) *Probability and Measure* ( )

H ( ) *The Annals of Statistics*

H ( ) E M *Journal of Multivariate Analysis*

H M ( ) H M *Scandinavian Journal of Statistics*

M ( ) *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

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

Let $K = K(n)$ denote the numbers of leading eigenfunctions included to approximate X as sample size n varies; i.e., $\widehat{X}_{\mathbf{i}}(s) = \hat{\mu}_{\mathbf{X}}(s) + \sum_{\mathbf{k}=1}^{\mathbf{K}} \hat{\xi}_{\mathbf{ik}}\hat{\phi}_{\mathbf{k}}(s)$. Analogously, define the quantities $\mathcal{G}_{\mathbf{Y}}$, $\widehat{\mathcal{G}}_{\mathbf{Y}}$, $D_{\mathbf{Y}}$, $\delta_{\mathbf{m}}^{\mathbf{y}}$ $\pi_{\mathbf{m}}^{\mathbf{y}}$, 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*

$$\sup_{\mathbf{t}\in\mathcal{S}} |\hat{\mu}_{\mathbf{X}}(s) - \mu_{\mathbf{X}}(s)| = O_{\mathbf{p}}(\frac{1}{\sqrt{n}b_{\mathbf{X}}}), \quad \sup_{\mathbf{s}_1,\mathbf{s}_2\in\mathcal{S}} |\widehat{G}_{\mathbf{X}}(s_1,s_2) - G_{\mathbf{X}}(s_1,s_2)| = O_{\mathbf{p}}(\frac{1}{\sqrt{n}h_{\mathbf{X}}^2}), \tag{33}$$

nd 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*

$$P(\sup_{1\le \mathbf{k}\le \mathbf{K}_0} |\hat{\lambda}_{\mathbf{k}} - \lambda_{\mathbf{k}}| \le D_{\mathbf{X}}) = 1, \quad \sup_{\mathbf{s}\in\mathcal{S}} |\hat{\phi}_{\mathbf{k}}(s) - \phi_{\mathbf{k}}(s)| = O_{\mathbf{p}}(\frac{\pi_{\mathbf{k}}^{\mathbf{x}}}{\sqrt{n}h_{\mathbf{X}}^2}), \quad k = 1,\dots,K_0, \tag{34}$$

here $D_{\mathbf{X}}$ $\pi_{\mathbf{k}}^{\mathbf{x}}$ *nd* K_0 *re de ned in*

An ogous y under B B C C nd C

$$\sup_{\mathbf{t}\in\mathcal{T}} |\hat{\mu}_{\mathbf{Y}}(t) - \mu_{\mathbf{Y}}(t)| = O_{\mathbf{p}}(\frac{1}{\sqrt{n}b_{\mathbf{Y}}}), \quad \sup_{\mathbf{t}_1,\mathbf{t}_2\in\mathcal{T}} |\widehat{G}_{\mathbf{Y}}(t_1,t_2) - G_{\mathbf{Y}}(t_1,t_2)| = O_{\mathbf{p}}(\frac{1}{\sqrt{n}h_{\mathbf{Y}}^2}), \tag{35}$$

nd s consequence $\hat{\sigma}_{\mathbf{Y}}^2 - \sigma_{\mathbf{Y}}^2 = O_{\mathbf{p}}(n^{-1/2}h_{\mathbf{Y}}^{-2} + n^{-1/2}h_{\mathbf{Y}}^{*\,-1})$ *Considering eigenv ues* $\rho_{\mathbf{m}}$ *of u tip icity one* $\hat{\psi}_{\mathbf{m}}$ *c n e chosen such th t*

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

here $D_{\mathbf{Y}}$ $\pi_{\mathbf{k}}^{\mathbf{y}}$ *nd* M_0 *re de ned n 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**

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

Lemma 2 *For* $\theta_{ik}^{(\ell)}$, $Z_k^{(\ell)}$, $\vartheta_{im}^{(\ell)}$ *and* $Q_m^{(\ell)}$ *as defined in*   *and*

$$|\hat{\xi}_{ik}^{I} - \xi_{ik}| \le \sum_{\ell=1}^{5} \theta_{ik}^{(\ell)} Z_k^{(\ell)}, \qquad |\hat{\zeta}_{im}^{I} - \zeta_{im}| \le \sum_{\ell=1}^{5} \vartheta_{im}^{(\ell)} Q_m^{(\ell)}. \tag{39}$$

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 ξ_k is denoted by p_k. Define

$$\begin{aligned}
\theta_k(x) &= p_k(x)\{\frac{\pi_k^X}{\sqrt{n}h_X^2} + \frac{1}{\sqrt{n}b_X} + \sqrt{\Delta_X^*}\}, \\
\vartheta_{mk}(x) &= p_k(x)\{\frac{\pi_m^Y}{\sqrt{n}h_Y^2} + \frac{1}{\sqrt{n}b_Y} + \sqrt{\Delta_Y^*}\}.
\end{aligned} \tag{40}$$

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,

$$\begin{aligned}
\tilde{\theta}_k(x) &= \frac{\theta_k(x)}{h_k} + \frac{1}{2}|f_k''(x)|h_k^2 + \sqrt{\frac{\text{var}(Y|x)\|K_1\|^2}{p_k(x)nh_k}}, \qquad (41) \\
\tilde{\vartheta}_{mk}(x) &= \frac{\theta_k(x)}{h_{mk}} + \vartheta_{mk}(x) + \frac{1}{2}|f_{mk}''(x)|h_{mk}^2 + \sqrt{\frac{\text{var}(\zeta_m|x)\|K_1\|^2}{p_k(x)nh_{mk}}}.
\end{aligned}$$

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 $K \le K_0$ and $M \le M_0$ in (A6). Since it follows from (35) that $K_0 \to \infty$, as long as all eigenvalues λ_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 θ_n^* and ϑ_n^* for the predictions (20) and

(21) are as follows,

$$\theta_n^* = \sum_{k=1}^{K}\{\frac{\theta_k(\xi_k)}{h_k} + \frac{1}{2}|f_k''(\xi_k)|h_k^2 + \sqrt{\frac{var(Y|\xi_k)\|K_1\|^2}{p_k(\xi_k)nh_k}}\} + |\sum_{k\geq K+1} f_k(\xi_k)|, \quad (42)$$

$$\vartheta_n^* = \sum_{k=1}^{K}\sum_{m=1}^{M}\{(\frac{\theta_k(\xi_k)}{h_{mk}} + \vartheta_{mk}(\xi_k))|\psi_m(t)| + \frac{1}{2}|f_{mk}''(\xi_k)|\psi_m(t)|h_{mk}^2 + \sqrt{\frac{var(\zeta_m|\xi_k)\|K_1\|^2}{p_k(\xi_k)nh_k}}|\psi_m(t)|$$

$$+ \frac{\pi_m^y|f_{mk}(\xi_k)|}{}$$

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

$$|\hat{\xi}_{ik} - \xi_{ik}| \leq \{|\hat{\eta}_{ik} - \tilde{\eta}_{ik}| + |\tilde{\eta}_{ik} - \xi_{ik}| + |\hat{\tau}_{ik}|\}. \tag{43}$$

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

$$\begin{aligned}
&\{\sum_{j=2}^{n_i}[|X_i(s_{ij}) - \hat{\mu}(s_{ij})| \cdot |\hat{\phi}_k(s_{ij}) - \phi_k(s_{ij})| + |\hat{\mu}(s_{ij}) - \mu(s_{ij})| \cdot |\phi_k(s_{ij})|](s_{ij} - s_{i,j-1})\} \\
&\leq \{\sum_{j=1}^{n_i}[|X_i(s_{ij})| + |\mu(s_{ij})| + 1]^2(s_{ij} - s_{i,j-1})\}^{1/2}\{\sum_{j=2}^{n_i}[\hat{\phi}_k(s_{ij}) - \phi_k(s_{ij})]^2(s_{ij} - s_{i,j-1})\}^{1/2} \\
&\quad + \{\sum_{j=1}^{n_i}[\hat{\mu}(s_{ij}) - \mu(s_{ij})]^2(s_{ij} - s_{i,j-1})\}^{1/2}\{\sum_{j=2}^{n_i} \phi_k^2(s_{ij})(s_{ij} - s_{i,j-1})\}^{1/2} \\
&\leq \theta_{ik}^{(1)} Z_k^{(1)} + \theta_{ik}^{(2)} Z_k^{(2)}.
\end{aligned}$$

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

$$|\tilde{\eta}_{ij} - \xi_{ik}| \leq \|(X_i + \mu)'\phi_k + (X_i + \mu)\phi_k'\|_\infty \Delta_X^* \leq \theta_{ik}^{(3)} Z_k^{(3)}.$$

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

$$\hat{f}_k(t) = \frac{\sum_i \hat{w}_i Y_i}{\sum_i \hat{w}_i} - \frac{\sum_i \hat{w}_i(\hat{\xi}_{ik} - x)}{\sum_i \hat{w}_i}\hat{f}_k'(x), \tag{44}$$

where

$$\hat{f}_k'(x) = \frac{\sum_i \hat{w}_i(\hat{\xi}_{ik} - x)Y_i - \{\sum_i \hat{w}_i(\hat{\xi}_{ik} - x)\sum_i \hat{w}_i Y_i\}/\sum_i \hat{w}_i}{\sum_i \hat{w}_i(\hat{\xi}_{ik} - x)^2 - \{\sum_i \hat{w}_i(\hat{\xi}_{ik} - x)\}^2/\sum_i \hat{w}_i}. \tag{45}$$

Let $\tilde{f}_k(x)$ be a hypothetical estimator, obtained by substituting the true values w_i and ξ_{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 differences

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

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,

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

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)

$$\frac{1}{nh^2_{\mathbf{k}}} \sum_{\mathbf{i}} |\hat{\xi}_{\mathbf{ik}} - \xi_{\mathbf{ik}}| I(|x - \xi_{\mathbf{ik}}| \le h_{\mathbf{k}}) \le \sum_{\ell=1}^{5} Z^{(\ell)}_{\mathbf{k}} \frac{1}{nh^2_{\mathbf{k}}} \sum_{\mathbf{i}} \theta^{(\ell)}_{\mathbf{ik}} I(|x - \xi_{\mathbf{ik}}| \le h_{\mathbf{k}}).$$

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

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

provided that $E(\theta^{(\ell)}_{\mathbf{ik}}) < \infty$ for $\ell = 1, \ldots, 5$. Note that $E\theta^{(1)}_{\mathbf{ik}} < \infty$, $E\theta^{(3)}_{\mathbf{ik}} < \infty$ by (A4), $E\theta^{(4)}_{\mathbf{ik}} \le 2\sigma_{\mathbf{X}} \sqrt{\Delta^*_{\mathbf{X}}}$ and $E\theta^{(5)}_{\mathbf{ik}} \le |\mathcal{S}| \sigma_{\mathbf{X}}$ by the Cauchy-Schwarz inequality. Then

$$Z^{(1)}_{\mathbf{k}} \frac{1}{nh^2_{\mathbf{k}}} \sum_{\mathbf{i}} \theta^{(1)}_{\mathbf{ik}} I(|x - \xi_{\mathbf{ik}}| \le h_{\mathbf{k}}) = O_{\mathbf{p}}\{ \frac{\pi^{\mathbf{x}}_{\mathbf{k}}}{\sqrt{n} h^2_{\mathbf{X}} h_{\mathbf{k}}} p_{\mathbf{k}}(x) \},$$
$$Z^{(2)}_{\mathbf{k}} \frac{1}{nh^2_{\mathbf{k}}} \sum_{\mathbf{i}} \theta^{(2)}_{\mathbf{ik}} I(|x - \xi_{\mathbf{ik}}| \le h_{\mathbf{k}}) = O_{\mathbf{p}}\{ \frac{1}{\sqrt{n} b_{\mathbf{X}} h_{\mathbf{k}}} p_{\mathbf{k}}(x) \},$$
$$Z^{(3)}_{\mathbf{k}} \frac{1}{nh^2_{\mathbf{k}}} \sum_{\mathbf{i}} \theta^{(3)}_{\mathbf{ik}} I(|x - \xi_{\mathbf{ik}}| \le h_{\mathbf{k}}) = O_{\mathbf{p}}\{ \frac{\|\phi_{\mathbf{k}}\|_\infty \Delta^*_{\mathbf{X}}}{h_{\mathbf{k}}} p_{\mathbf{k}}(x) \},$$
$$Z^{(4)}_{\mathbf{k}} \frac{1}{nh^2_{\mathbf{k}}} \sum_{\mathbf{i}} \theta^{(4)}_{\mathbf{ik}} I(|x - \xi_{\mathbf{ik}}| \le h_{\mathbf{k}}) = O_{\mathbf{p}}\{ \frac{\sqrt{\Delta^*_{\mathbf{X}}}}{h_{\mathbf{k}}} p_{\mathbf{k}}(x) \},$$
$$Z^{(5)}_{\mathbf{k}} \frac{1}{nh^2_{\mathbf{k}}} \sum_{\mathbf{i}} \theta^{(5)}_{\mathbf{ik}} I(|x - \xi_{\mathbf{ik}}| \le h_{\mathbf{k}}) = O_{\mathbf{p}}\{ \frac{\pi^{\mathbf{x}}_{\mathbf{k}}}{\sqrt{n} h^2_{\mathbf{X}} h_{\mathbf{k}}} p_{\mathbf{k}}(x) \}. \tag{48}$$

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

$$\frac{1}{nh_k}\sum_i I(|x-\hat{\xi}_{ik}|\leq h_k)\leq\frac{1}{nh_k}\sum_i\{I(|x-\xi_{ik}|\leq 2h_k)+I(\sum_{\ell=1}^{5}\theta_{ik}^{(\ell)}Z_k^{(\ell)}>h_k)\}\xrightarrow{p}4p_k(x),$$

leading to $(nh_k^2)^{-1}\sum_i|\hat{\xi}_{ik}-\xi_{ik}|I(|x-\hat{\xi}_{ik}|\leq h_k)=O_p(\theta_k h_k^{-1})$. Then $D_1=O_p(\theta_k h_k^{-1})$ follows.

Analogously, one shows $D_2=O_p(\theta_k h_k^{-1})$, applying the Cauchy-Schwarz inequality for $\theta_{ik}^{(\ell)}$, $\ell=1,3$, and observing the independence between Y_i and $\theta_{ik}^{(\ell)}$ 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 $D_{33}\leq c\sum_{\ell=1}^{5}Z_k^{(\ell)}(nh_k)^{-1}\sum_i\theta_{ik}^{(\ell)}I(|x-\xi_{ik}|\leq h_k)$ 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 $|\hat{\xi}_{ik}^2-\xi_{ik}^2|\leq|\hat{\xi}_{ik}-\xi_{ik}|\cdot|\xi_{ik}|+(\hat{\xi}_{ik}-\xi_{ik})^2$, 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_ℓ, $\ell=1,\dots,4$, 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

$$|\sum_i w_i(\hat{\zeta}_{im}-\zeta_{im})|\leq\sum_{\ell=1}^{5}Q_m^{(\ell)}\sum_i w_i\vartheta_{im}^{(\ell)}\leq\frac{1}{nh_{mk}}\sum_{\ell=1}^{5}Q_m^{(\ell)}\sum_i\vartheta_{im}^{(\ell)}I(|x-\xi_{ik}|\leq h_{mk})$$

Similar arguments as above complete this derivation. $\square$

Proof of Theore Using (A7), the derivation of θ_n^* in (42) is straightforward, following the above arguments. To obtain (21), note that

$$
\begin{aligned}
& \widehat{E}\{Y(t)|X\} - E\{Y(t)|X\} \\
\leq & \sum_{k=1}^{K}\sum_{m=1}^{M} |\hat{f}_{mk}(\xi_k)\hat{\psi}_m(t) - f_{mk}(\xi_k)\psi_m(t)| + |\sum_{k\geq K+1}\sum_{m\geq M+1} f_{mk}(\xi_k)\psi_m(t)| \\
\leq & \sum_{k=1}^{K}\sum_{m=1}^{M} [|\hat{f}_{mk}(\xi_k) - f_{mk}(\xi_k)|\{|\psi_m(t)| + |\hat{\psi}_m(t) - \psi_m(t)|\} + |f_{mk}(\xi_k)| \cdot |\hat{\psi}_m(t) - \psi_m(t)|] \\
& + |\sum_{(k,m)\in\mathcal{N}^2\setminus\mathcal{N}_K\times\mathcal{N}_M} f_{mk}(\xi_k)\psi_m(t)|.
\end{aligned}
$$

This implies the convergence rate ϑ_n^* in (42).