Technical Report

Tikhonov Regularization for Nonparametric Instrumental Variable Estimators

P. Gagliardini and O. Scaillet

August 2011

This technical report contains the proofs of the technical Lemmas A.1-A.8 in the paper entitled "Tikhonov Regularization for Nonparametric Instrumental Variable Estimators" and written by P. Gagliardini and O. Scaillet. We gather these proofs in Sections 1-8. Section 9 contains further proofs of technical lemmas B and C used in Sections 1-8. Finally, in Section 10 we discuss the characterization of operator $\mathcal{E}$ when $l, d_{X_2} \geq 1$ and the Sobolev embedding condition $2l > d_{X_2}$ is satisfied. Equations labelled as (n) refer to the paper, and Equations labelled as (TR.n) refer to the technical report.

1 Proof of Lemma A.1

(i) Consider the function

$$u = \sum_{j \in \mathbb{N}^{d_{X_2}}} \frac{1}{\xi_j} \left\langle \phi, \chi_j \right\rangle_{L^2(\mathcal{X}_2)} \chi_j. \tag{TR.1}$$

Then, $u \in H_0^{2l}(\mathcal{X}_2)$ and $\mathcal{D}u = \phi$, that is, u solves the PDE. Let us now show uniqueness of the solution. Suppose that u_1 and u_2 are two solutions in $H_0^{2l}(\mathcal{X}_2)$, and let us show that $u_1 = u_2$. For this purpose, we use that for any $\varphi_1, \varphi_2 \in H_0^{2l}(\mathcal{X}_2)$ we have:

$$\begin{aligned} \langle \varphi_1, \mathcal{D}\varphi_2 \rangle_{L^2(\mathcal{X}_2)} &= \sum_{|\alpha| \leq l} \left\langle \varphi_1, \left(-\nabla^2\right)^{\alpha} \varphi_2 \right\rangle_{L^2(\mathcal{X}_2)} \\ &= \sum_{|\alpha| \leq l} \langle \nabla^{\alpha}\varphi_1, \nabla^{\alpha}\varphi_2 \rangle_{L^2(\mathcal{X}_2)} = \langle \varphi_1, \varphi_2 \rangle_{H^l(\mathcal{X}_2)}, \end{aligned} \tag{TR.2}$$

from partial integration and the boundary conditions. Setting $\varphi_1 = \varphi_2 = u_1 - u_2$ in (TR.2) yields $\|u_1 - u_2\|_{H^l(\mathcal{X}_2)} = 0$, which implies $u_1 = u_2$.

(ii) From (TR.2) we have $\left\langle \chi_j, \chi_k \right\rangle_{H^l(\mathcal{X}_2)} = \left\langle \chi_j, \mathcal{D}\chi_k \right\rangle_{L^2(\mathcal{X}_2)} = \xi_k \delta_{j,k}$. Then, from (TR.1) we get for $\phi \in L^2(\mathcal{X}_2)$ and $u = \mathcal{D}^{-1}\phi$:

$$\left\|\mathcal{D}^{-1}\phi\right\|_{\mathcal{S}}^2 = \|u\|_{\mathcal{S}}^2 = \sum_{j \in \mathbb{N}^{d_{X_2}}} \left[\xi_j \left\langle u, \chi_j \right\rangle_{L^2(\mathcal{X}_2)}\right]^2 = \sum_{j \in \mathbb{N}^{d_{X_2}}} \left\langle \phi, \chi_j \right\rangle_{L^2(\mathcal{X}_2)}^2 = \|\phi\|_{L^2(\mathcal{X}_2)}^2.$$

Thus, operator $\mathcal{D}^{-1}$ is bounded and hence continuous.

(iii) For given $u \in H_0^{2l}(\mathcal{X}_2)$, let us consider the linear functional $\mathcal{T}_u(\varphi) = \langle \mathcal{D}u, \varphi \rangle_{L^2(\mathcal{X}_2)}$, for $\varphi \in H^l(\mathcal{X}_2)$. By using $\|\mathcal{D}u\|_{L^2(\mathcal{X}_2)}^2 = \sum_{j \in \mathbb{N}^{d_{X_2}}} \left[\xi_j \left\langle u, \chi_j \right\rangle_{L^2(\mathcal{X}_2)}\right]^2 = \|u\|_{\mathcal{S}}^2$ and the Cauchy-Schwartz inequality, we have

$$|\mathcal{T}_u(\varphi)| \leq \|\mathcal{D}u\|_{L^2(\mathcal{X}_2)} \|\varphi\|_{L^2(\mathcal{X}_2)} \leq \|u\|_{\mathcal{S}} \|\varphi\|_{H^l(\mathcal{X}_2)}.$$

Thus, $\mathcal{T}_u$ is continuous, with norm bounded by $\|u\|_{\mathcal{S}}$. From the Rietz representation theorem, for any $u \in H_0^{2l}(\mathcal{X}_2)$ there exists $\mathcal{E}(u) \in H^l(\mathcal{X}_2)$ such that $\mathcal{T}_u(\varphi) = \langle \mathcal{E}(u), \varphi \rangle_{H^l(\mathcal{X}_2)}$ for any $\varphi \in H^l(\mathcal{X}_2)$, and $\|\mathcal{E}(u)\|_{H^l(\mathcal{X}_2)} \leq \|u\|_{\mathcal{S}}$. From the definition of $\mathcal{T}_u$, the mapping $u \to \mathcal{E}(u)$ is linear. Hence, this mapping defines a bounded linear operator $\mathcal{E}$ from $H_0^{2l}(\mathcal{X}_2)$ to $H^l(\mathcal{X}_2)$, with norm bounded by 1.

(iv) Let $\psi \in L^2_{x_1}(\mathcal{Z}_1)$ and $\phi \in H^l(\mathcal{X}_2)$. Define $f = \tilde{A}_{x_1}\psi \in L^2(\mathcal{X}_2)$ and $u = \mathcal{D}^{-1}f \in H_0^{2l}(\mathcal{X}_2)$ from (i). Thus, we have from (iii):

$$\langle \psi, A_{x_1}\phi \rangle_{L^2_{x_1}(\mathcal{Z}_1)} = \left\langle \tilde{A}_{x_1}\psi, \phi \right\rangle_{L^2(\mathcal{X}_2)} = \langle \mathcal{D}u, \phi \rangle_{L^2(\mathcal{X}_2)} = \langle \mathcal{E}u, \phi \rangle_{H^l(\mathcal{X}_2)} = \langle \mathcal{E}\mathcal{D}^{-1}\tilde{A}_{x_1}\psi, \phi \rangle_{H^l(\mathcal{X}_2)}.$$

Then, we deduce $A^*_{x_1} = \mathcal{E}\mathcal{D}^{-1}\tilde{A}_{x_1}$.

2 Proof of Lemma A.2

(i) Let $\varphi \in H^l(\mathcal{X}_2)$ and $\psi \in L^2_{x_1}(\mathcal{Z}_1)$. Then $\int \hat{\Omega}_{x_1}(z_1)\left(\hat{A}_{x_1}\varphi\right)(z_1)\,\psi(z_1)\,\hat{f}_{Z_1|X_1}(z_1|x_1)\,dz_1 = \int \varphi(x_2)\left(\int \hat{\Omega}_{x_1}(z_1)\,\hat{f}_{X_2,Z_1|X_1}(x_2,z_1|x_1)\psi(z_1)\,dz_1\right)dx_2 = \langle \varphi, \widetilde{\hat{A}}_{x_1}\psi \rangle_{L^2(\mathcal{X}_2)}$, where $\widetilde{\hat{A}}_{x_1}$ is defined in (10). From Lemma A.1 (iii), $\int \hat{\Omega}_{x_1}(z_1)\left(\hat{A}_{x_1}\varphi\right)(z_1)\,\psi(z_1)\,\hat{f}_{Z_1|X_1}(z_1|x_1)\,dz_1 = \left\langle \varphi, \mathcal{E}\mathcal{D}^{-1}\widetilde{\hat{A}}_{x_1}\psi \right\rangle_{H^l(\mathcal{X}_2)}$, P-a.s.. Then, the linear operator $\hat{A}^*_{x_1} =: \mathcal{E}\mathcal{D}^{-1}\widetilde{\hat{A}}_{x_1}$ is well-defined, P-a.s., and has the desired properties.

(ii) Let us prove that operator $\hat{A}^*_{x_1}\hat{A}_{x_1} = \mathcal{E}\mathcal{D}^{-1}\widetilde{\hat{A}}_{x_1}\hat{A}_{x_1}$ is a compact operator from $H^l(\mathcal{X}_2)$ in itself. From Lemma A.1 (ii)-(iii), operator $\mathcal{E}\mathcal{D}^{-1} : L^2(\mathcal{X}_2) \to H^l(\mathcal{X}_2)$ is bounded. Thus, it is sufficient to prove that $\widetilde{\hat{A}}_{x_1}\hat{A}_{x_1} : H^l(\mathcal{X}_2) \to L^2(\mathcal{X}_2)$ is compact. To this aim, write

$$\begin{aligned}\left(\widetilde{\hat{A}}_{x_1}\hat{A}_{x_1}\varphi\right)(x_2) &= \int\left(\int \hat{\Omega}_{x_1}(z_1)\,\hat{f}_{X_2|Z}(x_2|z)\,\hat{f}_{X_2|Z}(\xi_2|z)\,\hat{f}_{Z_1|X_1}(z_1|x_1)\,dz_1\right)\varphi(\xi_2)\,d\xi_2 \\ &=: \int \alpha(x_2,\xi_2)\,\varphi(\xi_2)\,d\xi_2, \qquad \text{(TR.3)}\end{aligned}$$

for $\varphi \in H^l(\mathcal{X}_2)$. Since $\int \hat{\Omega}_{x_1}(z_1)\,\hat{f}_{X_2|Z}(x_2|z)^2\,\hat{f}_{Z_1|X_1}(z_1|x_1)\,dx_2dz_1 < \infty$, P-a.s, for almost any $x_1 \in \mathcal{X}_1$, from Assumptions B.2 and B.7 (i), and by using the Cauchy-Schwarz inequality, it follows that $\int \alpha(x_2,\xi_2)^2\,dx_2d\xi_2 < \infty$, P-a.s.. Thus, operator $\widetilde{\hat{A}}_{x_1}\hat{A}_{x_1} : L^2(\mathcal{X}_2) \to L^2(\mathcal{X}_2)$ is compact w.r.t. the norm $\|.\|_{L^2(\mathcal{X}_2)}$, P-a.s.. It follows that it is also a compact operator on $H^l(\mathcal{X}_2)$ w.r.t. the norm $\|.\|_{H^l(\mathcal{X}_2)}$, since bounded sets w.r.t. $\|.\|_{H^l(\mathcal{X}_2)}$ are also bounded w.r.t. $\|.\|_{L^2(\mathcal{X}_2)}$. This concludes the proof of Lemma A.2.

3 Proof of Lemma A.3

Write

$$\begin{aligned}
\mathcal{R}_{x_1,T} &= \left[\left(\lambda_{x_1,T}+\hat{A}_{x_1}^*\hat{A}_{x_1}\right)^{-1}\hat{A}_{x_1}^*\hat{A}_{x_1}-\left(\lambda_{x_1,T}+A_{x_1}^*A_{x_1}\right)^{-1}A_{x_1}^*A_{x_1}\right]\varphi_{0,x_1}\\
&+\left[\left(\lambda_{x_1,T}+\hat{A}_{x_1}^*\hat{A}_{x_1}\right)^{-1}-\left(\lambda_{x_1,T}+A_{x_1}^*A_{x_1}\right)^{-1}\right]A_{x_1}^*\left(\hat{\psi}_{x_1}+\zeta_{x_1}\right)\\
&+\left(\lambda_{x_1,T}+\hat{A}_{x_1}^*\hat{A}_{x_1}\right)^{-1}\left(\hat{A}_{x_1}^*-A_{x_1}^*\right)\left(\hat{\psi}_{x_1}+\zeta_{x_1}\right)\\
&+\left(\lambda_{x_1,T}+\hat{A}_{x_1}^*\hat{A}_{x_1}\right)^{-1}\hat{A}_{x_1}^*\hat{q}_{x_1}\\
&=: \mathcal{R}_{1,x_1,T}+\mathcal{R}_{2,x_1,T}+\mathcal{R}_{3,x_1,T}+\mathcal{R}_{4,x_1,T}. \qquad \text{(TR.4)}
\end{aligned}$$

We bound in probability the Sobolev norms of the terms $\mathcal{R}_{i,x_1,T}$, $i=1,...,4$, separately.

i) Bound of $\mathcal{R}_{1,x_1,T}$. We have:

$$\begin{aligned}
&\left(\lambda_{x_1,T}+\hat{A}_{x_1}^*\hat{A}_{x_1}\right)^{-1}\hat{A}_{x_1}^*\hat{A}_{x_1}-\left(\lambda_{x_1,T}+A_{x_1}^*A_{x_1}\right)^{-1}A_{x_1}^*A_{x_1}\\
= &\left(\lambda_{x_1,T}+\hat{A}_{x_1}^*\hat{A}_{x_1}\right)^{-1}\left(\hat{A}_{x_1}^*\hat{A}_{x_1}-A_{x_1}^*A_{x_1}\right)\\
&+\left[\left(\lambda_{x_1,T}+\hat{A}_{x_1}^*\hat{A}_{x_1}\right)^{-1}-\left(\lambda_{x_1,T}+A_{x_1}^*A_{x_1}\right)^{-1}\right]A_{x_1}^*A_{x_1}\\
= &\left(\lambda_{x_1,T}+\hat{A}_{x_1}^*\hat{A}_{x_1}\right)^{-1}\left(\hat{A}_{x_1}^*\hat{A}_{x_1}-A_{x_1}^*A_{x_1}\right)\\
&-\left(\lambda_{x_1,T}+\hat{A}_{x_1}^*\hat{A}_{x_1}\right)^{-1}\left(\hat{A}_{x_1}^*\hat{A}_{x_1}-A_{x_1}^*A_{x_1}\right)\left(\lambda_{x_1,T}+A_{x_1}^*A_{x_1}\right)^{-1}A_{x_1}^*A_{x_1}\\
= &-\left(\lambda_{x_1,T}+\hat{A}_{x_1}^*\hat{A}_{x_1}\right)^{-1}\left(\hat{A}_{x_1}^*\hat{A}_{x_1}-A_{x_1}^*A_{x_1}\right)\left[\left(\lambda_{x_1,T}+A_{x_1}^*A_{x_1}\right)^{-1}A_{x_1}^*A_{x_1}-1\right].
\end{aligned}$$

Thus, we get $\mathcal{R}_{1,x_1,T}=-\hat{S}_{x_1}\left(\lambda_{x_1,T}\right)\hat{U}_{x_1}\mathcal{B}_{x_1,T}^r$, where $\hat{S}_{x_1}\left(\lambda_{x_1,T}\right):=\left(\lambda_{x_1,T}+\hat{A}_{x_1}^*\hat{A}_{x_1}\right)^{-1}$ and $\hat{U}_{x_1}:=\hat{A}_{x_1}^*\hat{A}_{x_1}-A_{x_1}^*A_{x_1}$. Moreover, using

$$\hat{S}_{x_1}\left(\lambda_{x_1,T}\right)-S_{x_1}\left(\lambda_{x_1,T}\right)=-\left(1+S_{x_1}\left(\lambda_{x_1,T}\right)\hat{U}_{x_1}\right)^{-1}S_{x_1}\left(\lambda_{x_1,T}\right)\hat{U}_{x_1}S_{x_1}\left(\lambda_{x_1,T}\right),$$

where $S_{x_1}\left(\lambda_{x_1,T}\right):=\left(\lambda_{x_1,T}+A_{x_1}^*A_{x_1}\right)^{-1}$, we get:

$$\mathcal{R}_{1,x_1,T}=\left[\left(1+S_{x_1}\left(\lambda_{x_1,T}\right)\hat{U}_{x_1}\right)^{-1}S_{x_1}\left(\lambda_{x_1,T}\right)\hat{U}_{x_1}-1\right]S_{x_1}\left(\lambda_{x_1,T}\right)\hat{U}_{x_1}\mathcal{B}_{x_1,T}^r. \qquad \text{(TR.5)}$$

Thus:

$$\begin{aligned}
\left\|\mathcal{R}_{1,x_1,T}\right\|_{H^l(\mathcal{X}_2)} &\leq \left(\left\|\left(1+S_{x_1}\left(\lambda_{x_1,T}\right)\hat{U}_{x_1}\right)^{-1}\right\|_{H^l(\mathcal{X}_2)}\left\|S_{x_1}\left(\lambda_{x_1,T}\right)\hat{U}_{x_1}\right\|_{H^l(\mathcal{X}_2)}+1\right)\\
&\quad\cdot\left\|S_{x_1}\left(\lambda_{x_1,T}\right)\hat{U}_{x_1}\right\|_{H^l(\mathcal{X}_2)}\left\|\mathcal{B}_{x_1,T}^r\right\|_{H^l(\mathcal{X}_2)},
\end{aligned}$$

where $\|S\|_{H^l(\mathcal{X}_2)} := \|S\|_{\mathcal{L}\left(H^l(\mathcal{X}_2)\right)}$ denotes the operator norm of operator S on $H^l(\mathcal{X}_2)$. Then, $\|\mathcal{R}_{1,x_1,T}\|_{H^l(\mathcal{X}_2)} = o_p\left(\|\mathcal{B}^r_{x_1,T}\|_{H^l(\mathcal{X}_2)}\right)$ uniformly in $x_1 \in \mathcal{X}_1$ follows if we show:

(i) $\sup_{x_1\in\mathcal{X}_1} \left\|\left(1 + S_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1}\right)^{-1}\right\|_{H^l(\mathcal{X}_2)} = O_p(1)$,

(ii) $\sup_{x_1\in\mathcal{X}_1} \left\|S_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1}\right\|_{H^l(\mathcal{X}_2)} = o_p(1)$.

To prove (i)-(ii), we first note that (i) is implied by (ii). Indeed, by (ii) we have that $\sup_{x_1\in\mathcal{X}_1} \left\|S_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1}\right\|_{H^l(\mathcal{X}_2)} \leq 1/2$ w.p.a. 1. When $\sup_{x_1\in\mathcal{X}_1} \left\|S_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1}\right\|_{H^l(\mathcal{X}_2)} \leq 1/2$, by the Neumann series we have $\left(1 + S_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1}\right)^{-1} = \sum_{j=0}^{\infty}\left(S_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1}\right)^j$. Hence, $\sup_{x_1\in\mathcal{X}_1} \left\|\left(1 + S_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1}\right)^{-1}\right\|_{H^l(\mathcal{X}_2)} \leq \sum_{j=0}^{\infty}\left[\sup_{x_1\in\mathcal{X}_1}\left\|S_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1}\right\|_{H^l(\mathcal{X}_2)}\right]^j \leq 2$ w.p.a. 1. Let us now prove (ii). From Lemma A.1 we have $\left\|S_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1}\right\|_{H^l(\mathcal{X}_2)} \leq \lambda^{-1}_{x_1,T}\left\|\hat{U}_{x_1}\right\|_{H^l(\mathcal{X}_2)} \leq C\lambda^{-1}_{x_1,T}\left\|\widetilde{\widehat{A}}_{x_1}\widehat{A}_{x_1} - \tilde{A}_{x_1}A_{x_1}\right\|_{L^2(\mathcal{X}_2)}$. By using $\frac{\log T}{Th^{d_{X_1}}_{x_1,T}h_T^{d_{X_2}\vee d_{Z_1}}} + h_T^{2m} + h^{2m}_{x_1,T} = o\left(\lambda^2_{x_1,T}\right)$ uniformly in $x_1 \in \mathcal{X}_1$, (ii) follows from the next Lemma.

Lemma B. 1 *Under Assumptions B.1, B.2, B.3 (i), B.6, B.7 (i)-(ii), and if* $\frac{\log T}{Th^{d_{X_1}}_{x_1,T}h_T^{d_{Z_1}+d_{X_2}}} = O(1)$ *uniformly in* $x_1 \in \mathcal{X}_1$*, then uniformly in* $x_1 \in \mathcal{X}_1$*:*

$$\left\|\widetilde{\widehat{A}}_{x_1}\widehat{A}_{x_1} - \tilde{A}_{x_1}A_{x_1}\right\|^2_{L^2(\mathcal{X}_2)} = O_p\left(\frac{\log T}{Th^{d_{X_1}}_{x_1,T}h_T^{d_{X_2}\vee d_{Z_1}}} + h_T^{2m} + h^{2m}_{x_1,T}\right).$$

ii) Bound of $\mathcal{R}_{2,x_1,T}$. Similarly to previous lines, we have

$$\begin{aligned}&\left(\lambda_{x_1,T} + \hat{A}^*_{x_1}\hat{A}_{x_1}\right)^{-1} - \left(\lambda_{x_1,T} + A^*_{x_1}A_{x_1}\right)^{-1}\\ = \;& -\left(\lambda_{x_1,T} + \hat{A}^*_{x_1}\hat{A}_{x_1}\right)^{-1}\left(\hat{A}^*_{x_1}\hat{A}_{x_1} - A^*_{x_1}A_{x_1}\right)\left(\lambda_{x_1,T} + A^*_{x_1}A_{x_1}\right)^{-1}.\end{aligned}$$

Thus, we get

$$\begin{aligned}\mathcal{R}_{2,x_1,T} &= -\hat{S}_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1}S_{x_1}(\lambda_{x_1,T})A^*_{x_1}\left(\hat{\psi}_{x_1} + \zeta_{x_1}\right)\\ &= \left[\left(1 + S_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1}\right)^{-1}S_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1} - 1\right]S_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1}\left(\mathcal{V}_{x_1,T} + \mathcal{B}^e_{x_1,T}\right).\end{aligned} \tag{TR.6}$$

Then, $\|\mathcal{R}_{2,x_1,T}\|_{H^l(\mathcal{X}_2)} = o_p\left(\|\mathcal{V}_{x_1,T}\|_{H^l(\mathcal{X}_2)} + \|\mathcal{B}^e_{x_1,T}\|_{H^l(\mathcal{X}_2)}\right)$ uniformly in $x_1 \in \mathcal{X}_1$ follows from the arguments in point (i) above and Lemma B.1.

iii) Bound of $\mathcal{R}_{3,x_1,T}$. From Lemmas A.1 (ii) and A.2 (i) we have:

$$\begin{aligned}\|\mathcal{R}_{3,x_1,T}\|_{H^l(\mathcal{X}_2)} &\leq \left\| \left(\lambda_{x_1,T} + \hat{A}_{x_1}^* \hat{A}_{x_1} \right)^{-1} \right\|_{H^l(\mathcal{X}_2)} \|\mathcal{E}\mathcal{D}^{-1}\|_{HL} \left\| \left(\widetilde{\hat{A}}_{x_1} - \tilde{A}_{x_1} \right) \left(\hat{\psi}_{x_1} + \zeta_{x_1} \right) \right\|_{L^2(\mathcal{X}_2)} \\ &\leq C\lambda_{x_1,T}^{-1} \left\| \left(\widetilde{\hat{A}}_{x_1} - \tilde{A}_{x_1} \right) \left(\hat{\psi}_{x_1} + \zeta_{x_1} \right) \right\|_{L^2(\mathcal{X}_2)}, \qquad \text{(TR.7)}\end{aligned}$$

where $\|S\|_{HL} := \|S\|_{\mathcal{L}\left(L^2(\mathcal{X}_2), H^l(\mathcal{X}_2)\right)}$ for an operator S from $L^2(\mathcal{X}_2)$ to $H^l(\mathcal{X}_2)$. Then, $\|\mathcal{R}_{3,x_1,T}\|_{H^l(\mathcal{X}_2)} = O_p\left(\frac{1}{\lambda_{x_1,T}} \left(\frac{\log T}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}+d_{X_2}}} + h_{x_1,T}^{2m} + h_T^{2m} \right) \right)$, uniformly in $x_1 \in \mathcal{X}_1$, follows from the next lemma.

Lemma B. 2 *Under Assumptions B.1-B.3, B.6, B.7 (i)-(ii), and if* $\frac{\log T}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}}} = O(1)$ *uniformly in* $x_1 \in \mathcal{X}_1$, *then*

$$\left\| \left(\widetilde{\hat{A}}_{x_1} - \tilde{A}_{x_1} \right) \left(\hat{\psi}_{x_1} + \zeta_{x_1} \right) \right\|_{L^2(\mathcal{X}_2)} = O_p\left(\frac{\log T}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}+d_{X_2}}} + h_{x_1,T}^{2m} + h_T^{2m} \right),$$

uniformly in $x_1 \in \mathcal{X}_1$.

iv) Bound of $\mathcal{R}_{4,x_1,T}$. Similarly to previous lines, we have:

$$\|\mathcal{R}_{4,x_1,T}\|_{H^l(\mathcal{X}_2)} \leq C\lambda_{x_1,T}^{-1} \left\| \widetilde{\hat{A}}_{x_1} \hat{q}_{x_1} \right\|_{L^2(\mathcal{X}_2)}. \qquad \text{(TR.8)}$$

Then, $\|\mathcal{R}_{4,x_1,T}\|_{H^l(\mathcal{X}_2)} = O_p\left(\frac{1}{\lambda_{x_1,T}} \left(\frac{\log T}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}}} + h_{x_1,T}^{2m} + h_T^{2m} \right) \right)$, uniformly in $x_1 \in \mathcal{X}_1$, follows from the next lemma.

Lemma B. 3 *Under Assumptions B.1-B.3, B.6, B.7 (i)-(ii), and if* $\frac{\log T}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}}} = O(1)$ *uniformly in* $x_1 \in \mathcal{X}_1$, *then*

$$\left\| \widetilde{\hat{A}}_{x_1} \hat{q}_{x_1} \right\|_{L^2(\mathcal{X}_2)} = O_p\left(\frac{\log T}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}}} + h_{x_1,T}^{2m} + h_T^{2m} \right),$$

uniformly in $x_1 \in \mathcal{X}_1$.

4 Proof of Lemma A.4

Let us first define:

$$\hat{\Psi}_{x_1}(z_1) := \int g_{x_1,0}(w)\Delta\hat{f}(w,z)dw, \tag{TR.9}$$

for $z_1 \in \mathcal{Z}_1$, $x_1 \in \mathcal{X}_1$, where $g_{x_1,0}(w) := y - \varphi_{x_1,0}(x_2)$. The next Lemma C.1 gives a uniform bound for $\hat{\Psi}_{x_1}(z_1)$ and its expectation (see e.g. Newey (1994), Bosq (1998), and Hansen (2008) for similar results). Lemma C.1 will be used in the proofs of Lemma A.4 and other results.

LEMMA C.1: *Let Assumptions B.1-B.3 hold.* (1) *Uniformly in* $x_1 \in \mathcal{X}_1$*:*

$$\sup_{z_1 \in \mathcal{Z}_1} E\left[\hat{\Psi}_{x_1}(z_1)\right] = O\left(h_T^m + h_{x_1,T}^m\right).$$

(2) *Uniformly in* $x_1 \in \mathcal{X}_1$:

$$\sup_{z_1 \in \mathcal{Z}_1} \left|\hat{\Psi}_{x_1}(z_1) - E\left[\hat{\Psi}_{x_1}(z_1)\right]\right| = O_p\left(\sqrt{\frac{\log T}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}}}}\right).$$

(3) *If function* g *on* $\mathcal{Z} \times \mathcal{X}_2$ *is such that* $\|Dg\|_\infty < \infty$*, then uniformly in* $x \in \mathcal{X}$:

$$\int g(z,x_2)\left(\hat{\Psi}_{x_1}(z_1) - E\left[\hat{\Psi}_{x_1}(z_1)\right]\right)dz_1 = O_p\left(\sqrt{\frac{\log T}{T h_{x_1,T}^{d_{X_1}}}}\right).$$

4.1 Proof of Part (i)

We have $\left\|\hat{\psi}_{x_1}\right\|_{L^2_{x_1}(\mathcal{Z}_1)} \leq \sup_{z_1 \in \mathcal{Z}_1}\left|\hat{\psi}_{x_1}(z_1)\right|$ and $\hat{\psi}_{x_1}(z_1) = \left(\hat{\Psi}_{x_1}(z_1) - E\left[\hat{\Psi}_{x_1}(z_1)\right]\right)/f_Z(z)$, uniformly in $x_1 \in \mathcal{X}_1$. From Assumptions B.1 (ii)-(iii) and B.3 (ii), and by using Lemma C.1 (2), the conclusion follows.

4.2 Proof of Part (ii)

From Lemma A.1 (ii)-(iv), we have $\left\|A^*_{x_1}\hat{\psi}_{x_1}\right\|_{H^l(\mathcal{X}_2)} \leq C\left\|\tilde{A}_{x_1}\hat{\psi}_{x_1}\right\|_{L^2(\mathcal{X}_2)} \leq \sup_{x_2 \in \mathcal{X}_2}\left|\tilde{A}_{x_1}\hat{\psi}_{x_1}(x_2)\right|$ uniformly in $x_1 \in \mathcal{X}_1$, and

$$\begin{aligned}
\tilde{A}_{x_1}\hat{\psi}_{x_1}(x_2) &= \int \Omega_{x_1,0}(z_1) f(x_2, z_1|x_1)\hat{\psi}_{x_1}(z_1)dz_1 \\
&= \frac{1}{f_{X_1}(x_1)}\int\int \Omega_{x_1,0}(z_1) f_{X_2|Z}(x_2|z) g_{x_1,0}(w)\left(\hat{f}_{W,Z}(w,z) - E\left[\hat{f}_{W,Z}(w,z)\right]\right)dwdz_1 \\
&= \frac{1}{f_{X_1}(x_1)}\int \Omega_{x_1,0}(z_1) f_{X_2|Z}(x_2|z)\left(\hat{\Psi}_{x_1}(z_1) - E\left[\hat{\Psi}_{x_1}(z_1)\right]\right)dz_1.
\end{aligned}$$

From Assumptions B.1-B.3 and B.6, and Lemma C.1 (3), we have $\sup_{x_2\in\mathcal{X}_2}\left|\tilde{A}_{x_1}\hat{\psi}_{x_1}(x_2)\right| = O_p\left(\sqrt{\frac{\log T}{Th_{x_1,T}^{d_{X_1}}}}\right)$, uniformly in $x_1 \in \mathcal{X}_1$. The conclusion follows.

4.3 Proof of Part (iii)

We have $\left\|\zeta_{x_1}\right\|_{L^2_{x_1}(\mathcal{Z}_1)} \leq \sup_{z\in\mathcal{Z}}\left|\zeta_{x_1}(z_1)\right|$ and $\zeta_{x_1}(z_1) = E\left[\hat{\Psi}_{x_1}(z_1)\right]/f_Z(z)$. From Assumptions B.1 (ii)-(iii) and B.3 (ii), and by applying Lemma C.1 (1), the conclusion follows.

5 Proof of Lemma A.5

From (TR.4)-(TR.8), we get:

$$\begin{aligned}
\|\mathcal{R}_{x_1,T}\|_{L^2(\mathcal{X}_2)} &\leq \|\mathcal{R}_{1,x_1,T}+\mathcal{R}_{2,x_1,T}\|_{L^2(\mathcal{X}_2)} + \|\mathcal{R}_{3,x_1,T}\|_{L^2(\mathcal{X}_2)} + \|\mathcal{R}_{4,x_1,T}\|_{L^2(\mathcal{X}_2)} \\
&\leq \left(\left\|\left(1+S_{x_1}\left(\lambda_{x_1,T}\right)\hat{U}_{x_1}\right)^{-1}\right\|_{L^2(\mathcal{X}_2)}\left\|S_{x_1}\left(\lambda_{x_1,T}\right)\hat{U}_{x_1}\right\|_{L^2(\mathcal{X}_2)}+1\right) \\
&\quad\cdot\left\|S_{x_1}\left(\lambda_{x_1,T}\right)\hat{U}_{x_1}\right\|_{L^2(\mathcal{X}_2)}\left(\|\mathcal{V}_{x_1,T}\|_{L^2(\mathcal{X}_2)}+\|\mathcal{B}^e_{x_1,T}+\mathcal{B}^r_{x_1,T}\|_{L^2(\mathcal{X}_2)}\right) \\
&\quad+C\lambda_{x_1,T}^{-1}\left\|\left(\widetilde{\hat{A}}_{x_1}-\tilde{A}_{x_1}\right)\left(\hat{\psi}_{x_1}+\zeta_{x_1}\right)\right\|_{L^2(\mathcal{X}_2)} \\
&\quad+C\lambda_{x_1,T}^{-1}\left\|\widetilde{\hat{A}}_{x_1}\hat{q}_{x_1}\right\|_{L^2(\mathcal{X}_2)}.
\end{aligned}$$

Then, by the triangular and Cauchy-Schwarz inequalities, we get:

$$\begin{aligned}
E\left[\|\mathcal{R}_{x_1,T}\|^2_{L^2(\mathcal{X}_2)}\right]^{1/2} &\leq \left(E\left[\left\|\left(1+S_{x_1}\hat{U}_{x_1}\right)^{-1}\right\|^{16}_{L^2(\mathcal{X}_2)}\right]^{1/16}E\left[\left\|S_{x_1}\hat{U}_{x_1}\right\|^{16}_{L^2(\mathcal{X}_2)}\right]^{1/16}+1\right) \\
&\quad\cdot E\left[\left\|S_{x_1}\hat{U}_{x_1}\right\|^{8}_{L^2(\mathcal{X}_2)}\right]^{1/8}E\left[\|\mathcal{V}_{x_1,T}\|^4_{L^2(\mathcal{X}_2)}\right]^{1/4} \\
&\quad+\left(E\left[\left\|\left(1+S_{x_1}\hat{U}_{x_1}\right)^{-1}\right\|^{8}_{L^2(\mathcal{X}_2)}\right]^{1/8}E\left[\left\|S_{x_1}\hat{U}_{x_1}\right\|^{8}_{L^2(\mathcal{X}_2)}\right]^{1/8}+1\right) \\
&\quad\cdot E\left[\left\|S_{x_1}\hat{U}_{x_1}\right\|^{4}_{L^2(\mathcal{X}_2)}\right]^{1/4}\|\mathcal{B}^e_{x_1,T}+\mathcal{B}^r_{x_1,T}\|_{L^2(\mathcal{X}_2)} \\
&\quad+C\lambda_{x_1,T}^{-1}E\left[\left\|\left(\widetilde{\hat{A}}_{x_1}-\tilde{A}_{x_1}\right)\left(\hat{\psi}_{x_1}+\zeta_{x_1}\right)\right\|^2_{L^2(\mathcal{X}_2)}\right]^{1/2} \\
&\quad+C\lambda_{x_1,T}^{-1}E\left[\left\|\widetilde{\hat{A}}_{x_1}\hat{q}_{x_1}\right\|^2_{L^2(\mathcal{X}_2)}\right]^{1/2},
\end{aligned}$$

where $S_{x_1} := S_{x_1}(\lambda_{x_1,T})$. Thus, the conclusion follows, if we show:

(i) $E\left[\left\|\left(1+S_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1}\right)^{-1}\right\|_{L^2(\mathcal{X}_2)}^{16}\right] = O(1)$,

(ii) $E\left[\left\|S_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1}\right\|_{L^2(\mathcal{X}_2)}^{16}\right] = o(1)$,

(iii) $E\left[\|\mathcal{V}_{x_1,T}\|_{L^2(\mathcal{X}_2)}^4\right]^{1/4} = O\left(E\left[\|\mathcal{V}_{x_1,T}\|_{L^2(\mathcal{X}_2)}^2\right]^{1/2}\right)$,

(iv) $\lambda_{x_1,T}^{-1}E\left[\left\|\left(\widetilde{\hat{A}}_{x_1}-\tilde{A}_{x_1}\right)\left(\hat{\psi}_{x_1}+\zeta_{x_1}\right)\right\|_{L^2(\mathcal{X}_2)}^2\right]^{1/2} = o\left(\left\|\mathcal{B}_{x_1,T}^e+\mathcal{B}_{x_1,T}^r\right\|_{L^2(\mathcal{X}_2)}\right)$,

(v) $\lambda_{x_1,T}^{-1}E\left[\left\|\widetilde{\hat{A}}_{x_1}\hat{q}_{x_1}\right\|_{L^2(\mathcal{X}_2)}^2\right]^{1/2} = o\left(\left\|\mathcal{B}_{x_1,T}^e+\mathcal{B}_{x_1,T}^r\right\|_{L^2(\mathcal{X}_2)}\right)$.

Let us first prove (i) and (ii). We have $\left\|S_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1}\right\|_{L^2(\mathcal{X}_2)} \leq \left\|S_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1}\right\|_{HL} \leq \|S_{x_1}(\lambda_{x_1,T})\|_{H^l(\mathcal{X}_2)}\left\|\hat{U}_{x_1}\right\|_{HL}$. Using $\|S_{x_1}(\lambda_{x_1,T})\|_{H^l(\mathcal{X}_2)} \leq 1/\lambda_{x_1,T}$ and $\left\|\hat{U}_{x_1}\right\|_{HL} \leq \|\mathcal{E}\mathcal{D}^{-1}\|_{HL}\left\|\widetilde{\hat{A}}_{x_1}\hat{A}_{x_1}-\widetilde{A}_{x_1}A_{x_1}\right\|_{L^2(\mathcal{X}_2)}$, P-a.s., we deduce

$$\left\|S_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1}\right\|_{L^2(\mathcal{X}_2)} \leq \frac{\|\mathcal{E}\mathcal{D}^{-1}\|_{HL}}{\lambda_{x_1,T}}\left\|\widetilde{\hat{A}}_{x_1}\hat{A}_{x_1}-\widetilde{A}_{x_1}A_{x_1}\right\|_{L^2(\mathcal{X}_2)}, \ P\text{-a.s.}. \qquad \text{(TR.10)}$$

Moreover, from a similar argument as in Hall and Horowitz (2005) on p. 2925, we have $\left\|\left(1+S_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1}\right)^{-1}\right\|_{L^2(\mathcal{X}_2)} \leq C\left(1+\frac{1}{\lambda_{x_1,T}}\mathbb{I}_{\|S_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1}\|_{L^2(\mathcal{X}_2)}\geq\frac{1}{2}}\right)$, P-a.s., for a constant C. As in the argument of Hall and Horowitz (2005) in their Inequality (6.27), from Markov inequality it follows

$$E\left[\left\|\left(1+S_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1}\right)^{-1}\right\|_{L^2(\mathcal{X}_2)}^{16}\right] \leq \bar{C}\left(1+\frac{1}{\lambda_{x_1,T}^{16}}E\left[\left\|S_{x_1}(\lambda_{x_1,T})\hat{U}_{x_1}\right\|_{L^2(\mathcal{X}_2)}^{2l}\right]\right), \qquad \text{(TR.11)}$$

for any $l\in\mathbb{N}$, for a constant $\bar{C}$ depending on l but not on T. From (TR.10)-(TR.11), and using $\frac{1}{Th_{x_1,T}^{d_{X_1}}h_T^{d_{X_2}\vee d_{Z_1}}}+h_T^{2m}+h_{x_1,T}^{2m} = O\left(\lambda_{x_1,T}^{2+\varepsilon}\right)$, $\varepsilon>0$, points (i)-(ii) follow from the next Lemma.

Lemma B. 4 *Under Assumptions 5, B.1-B.3 (i), B.6, B.7 (iii) and condition* $\frac{(\log T)^2}{Th_{x_1,T}^{d_{X_1}}h_T^{d_{Z_1}+d_{X_2}}} = O(1)$, *it follows* $E\left[\left\|\widetilde{\hat{A}}_{x_1}\hat{A}_{x_1}-\widetilde{A}_{x_1}A_{x_1}\right\|_{L^2(\mathcal{X}_2)}^{2\zeta}\right] = O\left(a_T^{\zeta}\right)$, *for any* $\zeta\in\mathbb{N}$, *where* $a_T := \frac{1}{Th_{x_1,T}^{d_{X_1}}h_T^{d_{X_2}\vee d_{Z_1}}}+h_T^{2m}+h_{x_1,T}^{2m}$.

Point (iii) follows from the next Lemma B.5.

Lemma B. 5 *Under Assumptions 5, B.1-B.4 and B.8:*

$$E\left[\left\|\left(\lambda_{x_1,T}+A_{x_1}^*A_{x_1}\right)^{-1}A_{x_1}^*\hat{\psi}_{x_1}\right\|_{L^2(\mathcal{X}_2)}^4\right]^{1/4}=O\left(E\left[\left\|\left(\lambda_{x_1,T}+A_{x_1}^*A_{x_1}\right)^{-1}A_{x_1}^*\hat{\psi}_{x_1}\right\|_{L^2(\mathcal{X}_2)}^2\right]^{1/2}\right).$$

Finally, by using that $\left\|\mathcal{B}_{x_1,T}^e+\mathcal{B}_{x_1,T}^r\right\|_{L^2(\mathcal{X}_2)}=b\left(\lambda_{x_1,T},h_{x_1,T}\right)(1+o(1))$ from Lemma A.7, and the condition $\frac{1}{Th_{x_1,T}^{d_{X_1}}h_T^{d_{Z_1}+d_{X_2}}}+h_{x_1,T}^{2m}+h_T^{2m}=o\left(\lambda_{x_1,T}b\left(\lambda_{x_1,T},h_{x_1,T}\right)\right)$, points (iv)-(v) follow from the next Lemmas.

Lemma B. 6 *Under Assumptions 5, B.1-B.3, B.5, B.6 and B.7 (iii) and if* $\frac{1}{Th_{x_1,T}^{d_{X_1}}h_T^{d_{Z_1}}}=O(1)$*, then*

$$E\left[\left\|\left(\widetilde{\hat{A}}_{x_1}-\tilde{A}_{x_1}\right)\left(\hat{\psi}_{x_1}+\zeta_{x_1}\right)\right\|_{L^2(\mathcal{X}_2)}^2\right]^{1/2}=O\left(\frac{1}{Th_{x_1}^{d_{X_1}}h_T^{d_{Z_1}+d_{X_2}}}+h_{x_1,T}^{2m}+h_T^{2m}\right),$$

Lemma B. 7 *Under Assumptions 5, B.1-B.3, B.5, B.6 and B.7 (iii) and if* $\frac{1}{Th_{x_1,T}^{d_{X_1}}h_T^{d_{Z_1}}}=O(1)$*, then*

$$E\left[\left\|\widetilde{\hat{A}}_{x_1}\hat{q}_{x_1}\right\|_{L^2(\mathcal{X}_2)}^2\right]^{1/2}=O\left(\frac{1}{Th_{x_1,T}^{d_{X_1}}h_T^{d_{Z_1}}}+h_{x_1,T}^{2m}+h_T^{2m}\right).$$

6 Proof of Lemma A.6

Let us first expand the function $\left(\lambda_{x_1,T}+A_{x_1}^*A_{x_1}\right)^{-1}A_{x_1}^*\hat{\psi}_{x_1}$ w.r.t. the basis of eigenfunctions $\left\{\phi_{x_1,j}\right\}$ of operator $A_{x_1}^*A_{x_1}$, with eigenvalues $\nu_{x_1,j}$. We have

$$\begin{aligned}\left(\lambda_{x_1,T}+A_{x_1}^*A_{x_1}\right)^{-1}A_{x_1}^*\hat{\psi}_{x_1} &= \sum_{j=1}^{\infty}\langle\phi_{x_1,j},\left(\lambda_{x_1,T}+A_{x_1}^*A_{x_1}\right)^{-1}A_{x_1}^*\hat{\psi}_{x_1}\rangle_{H^l(\mathcal{X}_2)}\phi_{x_1,j}\\ &= \sum_{j=1}^{\infty}\langle\left(\lambda_{x_1,T}+A_{x_1}^*A_{x_1}\right)^{-1}\phi_{x_1,j},A_{x_1}^*\hat{\psi}_{x_1}\rangle_{H^l(\mathcal{X}_2)}\phi_{x_1,j}\\ &= \sum_{j=1}^{\infty}\frac{1}{\lambda_{x_1,T}+\nu_{x_1,j}}\langle\phi_{x_1,j},A_{x_1}^*\hat{\psi}_{x_1}\rangle_{H^l(\mathcal{X}_2)}\phi_{x_1,j}.\end{aligned}$$

Define the variables for $j\in\mathbb{N}$:

$$\begin{aligned}Z_{x_1,j,T}:= &\ \frac{1}{\sqrt{\nu_{x_1,j}}}\langle\phi_{x_1,j},\sqrt{Th_{x_1,T}^{d_{X_1}}}A_{x_1}^*\hat{\psi}_{x_1}\rangle_{H^l(\mathcal{X}_2)}=\frac{1}{\sqrt{\nu_{x_1,j}}}\langle A_{x_1}\phi_{x_1,j},\sqrt{Th_{x_1,T}^{d_{X_1}}}\hat{\psi}_{x_1}\rangle_{L^2_{x_1}(\mathcal{Z}_1)}\\ = &\ \sqrt{Th_{x_1,T}^{d_{X_1}}}\int\int\psi_{x_1,j}\left(z_1\right)\Omega_{x_1,0}\left(z_1\right)g_{\varphi_0}\left(w\right)\left[\hat{f}_{W,Z}(w,z)-E\hat{f}_{W,Z}(w,z)\right]dwdz_1, \quad \text{(TR.12)}\end{aligned}$$

where $g_{x_1,0}(w) := y - \varphi_{x_1,0}(x_2)$ and $\psi_{x_1,j} := \frac{1}{\sqrt{\nu_{x_1,j}}} A_{x_1}\phi_{x_1,j}$. Then, we can write $\left(\lambda_{x_1,T} + A_{x_1}^* A_{x_1}\right)^{-1} A_{x_1}^* \hat{\psi}_{x_1} = \frac{1}{\sqrt{Th_{x_1,T}^{d_{X_1}}}} \sum_{j=1}^{\infty} \frac{\sqrt{\nu_{x_1,j}}}{\lambda_{x_1,T} + \nu_{x_1,j}} Z_{x_1,j,T}\phi_{x_1,j}$. We deduce

$$\begin{aligned} & E\left[\left\| \left(\lambda_{x_1,T} + A_{x_1}^* A_{x_1}\right)^{-1} A_{x_1}^* \hat{\psi}_{x_1} \right\|_{L^2(\mathcal{X}_2)}^2\right] \\ = & \frac{1}{Th_{x_1,T}^{d_{X_1}}} \sum_{j,l=1}^{\infty} \frac{\sqrt{\nu_{x_1,j}}}{\lambda_{x_1,T} + \nu_{x_1,j}} \frac{\sqrt{\nu_{x_1,l}}}{\lambda_{x_1,T} + \nu_{x_1,l}} \left\langle \phi_{x_1,j}, \phi_{x_1,l} \right\rangle_{L^2(\mathcal{X}_2)} E\left[Z_{x_1,j,T} Z_{x_1,l,T}\right]. \quad \text{(TR.13)} \end{aligned}$$

To derive the asymptotic behaviour of the RHS, we need Lemmas B.10 and B.11. The proof of Lemma B.10 builds on the next Lemmas B.8 and B.9.

Lemma B. 8 *Let g be a function with support $\mathcal{S} \subset \mathbb{R}^d$, $\mathcal{S}$ convex, and such that $g \in L^2(F)$, where the distribution F has density f. Let $k(.)$ be a function in $\mathbb{R}^d$ such that $\int k(y)dy = 1$, $w_1 := \int |k(y)|\, dy < \infty$, and $w_2 := \int |y|^2 |k(y)|\, dy < \infty$. Define the function $\Delta g_h(x) := \int \frac{1}{h^d} k\left(\frac{y}{h}\right) [g(x-y) - g(x)]\, 1_{\mathcal{S}}(x-y)dy$, $x \in \mathcal{S}$, for any $h > 0$. Then*

$$\begin{aligned} \|\Delta g_h\|_{L^2(F)}^2 \leq\ & w_1 w_2 h^2 \|\nabla g\|_{L^2(F)}^2 \\ & + w_2 h \int |\nabla g(y)|^2 \left(\int_0^h \int |k(z)|\, |f(y+tz) - f(y)|\, dz dt\right) dy. \end{aligned}$$

Lemma B. 9 *Let $\{U_t = (U_{1,t}, U_{2,t}) : t = 1, ..., T\}$ be i.i.d. variables, with value in a convex set $\mathcal{S} = \mathcal{S}_1 \times \mathcal{S}_2$, $\mathcal{S}_i \subset \mathbb{R}^{d_i}$, $i = 1, 2$, and density f satisfying Assumptions B.1 (ii)-(iii). Let $\hat{f}$ denote the kernel estimator of f, with product kernel using K_1 on $\mathbb{R}^{d_1}$ and K_2 on $\mathbb{R}^{d_2}$ satisfying Assumption B.2, and bandwidths $h_{1T}, h_{2T} \to 0$. For given $x_1 \in \mathcal{S}_1$, let $\mathcal{G}$ denote the set of functions $\mathcal{G} = \left\{ g \in L_{x_1}^2(F) : E[g(U_2)|U_1 = x_1] = 0,\ \|g\|_{L_{x_1}^2(F^*)} < \infty,\ \|\nabla g\|_{L_{x_1}^2(F^*)} < \infty \right\}$, where $L_{x_1}^2(F)$ denotes the space of functions of U_2 which are square integrable w.r.t. density $f_{U_2|U_1}(.|x_1)$, and $\|.\|_{L_{x_1}^2(F^*)}$ denotes the L^2-norm w.r.t. the density $f_{x_1}^*$ defined by $f_{x_1}^*(u_2) := q(x_1, u_2) / \int q(x_1, u_2)\, du_2$, with q as in Assumption B.4. Further, for $g \in \mathcal{G}$ and $h > 0$ denote $\rho_{x_1}(g, h)^2 := \int g(u_2)^2 1(u_2 \in \partial\mathcal{S}_2(h)) f_{U_2|U_1}(u_2|x_1) du_2$, where $\partial\mathcal{S}_2(h) = \{u_2 \in \mathcal{S}_2 : dist(u_2, \mathcal{S}_2^c) \leq h\}$. Define $V_T(g, x_1) := \sqrt{Th_{1T}^{d_1}} \int g(u_2) \left[\hat{f}(x_1, u_2) - E\hat{f}(x_1, u_2)\right] du_2$, $T \in \mathbb{N}$, for $g \in \mathcal{G}$. Then*

$$\begin{aligned} E[V_T(g, x_1)V_T(e, x_1)] = \omega^2 f_{U_1}(x_1) Cov[g(U_2), e(U_2)|U_1 = x_1] + O\left(\|g\|_{L_{x_1}^2(F)} \rho_{x_1}(e, \kappa h_{2T})\right) \\ + O\Big(h_{2T} \|g\|_{L_{x_1}^2(F)} \|\nabla e\|_{L_{x_1}^2(F)} + h_{2T}^{3/2} \|g\|_{L_{x_1}^2(F)} \|\nabla e\|_{L_{x_1}^2(F^*)} \\ + (h_{1T} + h_{2T}) \|g\|_{L_{x_1}^2(F^*)} \|e\|_{L_{x_1}^2(F^*)} \\ + h_{2T}(h_{1T} + h_{2T}) \|g\|_{L_{x_1}^2(F^*)} \|\nabla e\|_{L_{x_1}^2(F^*)}\Big), \end{aligned}$$

uniformly in $g, e \in \mathcal{G}$, for $\omega^2 = \int K_1(u_1)^2 du_1$ and a constant $\kappa > 0$.

Lemma B. 10 *Under Assumptions B.1-B.4, B.8 (iii)-(iv) and 5, for any $x_1 \in \mathcal{X}_1$: (i) $E\left[Z_{x_1,j,T}^2\right] = \omega^2 f_{X_1}(x_1) + o(1)$, uniformly in $j \in \mathbb{N}$, where $\omega^2 = \int K(x_1)^2 dx_1$. (ii) $E\left[Z_{x_1,j,T} Z_{x_1,l,T}\right] = o(1)$, for any $j \neq l$, uniformly in $j, l \in \mathbb{N}$.*

Lemma B. 11 *Let $\{Z_j : j = 1, 2, ...\}$ be a sequence of zero mean r.v.'s, and let $(\alpha_{j,l})$, $j, l = 1, 2, \cdots$, be an array of positive numbers. Denote the correlation $\rho_{j,l} := corr\left(Z_j, Z_l\right)$. Then*

$$\left|\sum_{j,l=1}^{\infty} \alpha_{j,l} E\left[Z_j Z_l\right] - \sum_{j=1}^{\infty} \alpha_{j,j} E\left[Z_j^2\right]\right| \leq \left(\sum_{j=1}^{\infty} \alpha_{j,j} E\left[Z_j^2\right]\right) \left(\sum_{j,l=1,\ j\neq l}^{\infty} \rho_{jl}^2 \frac{\alpha_{j,l}^2}{\alpha_{j,j}\alpha_{l,l}}\right)^{1/2}.$$

Let us now conclude the proof of Lemma A.6. We apply Lemma B.11 to sequence $Z_j = Z_{x_1,j,T}$ in (TR.12) with $\alpha_{j,l} := \frac{1}{T h_{x_1,T}^{d_{X_1}}} \frac{\sqrt{\nu_{x_1,j}}}{\lambda_{x_1,T} + \nu_{x_1,j}} \frac{\sqrt{\nu_{x_1,l}}}{\lambda_{x_1,T} + \nu_{x_1,l}} \left\langle \phi_{x_1,j}, \phi_{x_1,l} \right\rangle_{L^2(\mathcal{X}_2)}$. It follows from (TR.13):

$$\begin{aligned} & E\left[\left\| \left(\lambda_{x_1,T} + A_{x_1}^* A_{x_1}\right)^{-1} A_{x_1}^* \hat{\psi}_{x_1} \right\|_{L^2(\mathcal{X}_2)}^2\right] \\ = & \left(\frac{1}{T h_{x_1,T}^{d_{X_1}}} \sum_{j=1}^{\infty} \frac{\nu_{x_1,j}}{\left(\lambda_{x_1,T} + \nu_{x_1,j}\right)^2} \left\|\phi_{x_1,j}\right\|_{L^2(\mathcal{X}_2)}^2 E\left[Z_{x_1,j,T}^2\right]\right) \left(1 + R_{1,x_1,T}\right), \end{aligned} \qquad \text{(TR.14)}$$

where $\left|R_{1,x_1,T}\right| \leq \left(\sum_{j,l=1,\ j\neq l}^{\infty} \rho_{x_1,jl,T}^2 \frac{\left\langle \phi_{x_1,j}, \phi_{x_1,l} \right\rangle_{L^2(\mathcal{X}_2)}^2}{\left\|\phi_{x_1,j}\right\|_{L^2(\mathcal{X}_2)}^2 \left\|\phi_{x_1,l}\right\|_{L^2(\mathcal{X}_2)}^2}\right)^{1/2}$, and $\rho_{x_1,jl,T} := corr\left(Z_{x_1,j,T}, Z_{x_1,l,T}\right)$. From Lemma B.10 (i) it follows $E\left[Z_{x_1,j,T}^2\right] \geq \omega^2 f_{X_1}(x_1)/2$, for all $j \in \mathbb{N}$, $x_1 \in \mathcal{X}_1$ and large T. Then, we get $\rho_{x_1,jl,T} = E\left[Z_{x_1,j,T} Z_{x_1,l,T}\right] / (E\left[Z_{x_1,j,T}^2\right]^{1/2} E\left[Z_{x_1,l,T}^2\right]^{1/2})$ $\leq \left[2/(\omega^2 f_{X_1}(x_1))\right] E\left[Z_{x_1,j,T} Z_{x_1,l,T}\right]$, for large T and any $x_1 \in \mathcal{X}_1$. Thus, from Lemma B.10 (ii) and Assumption B.8 (i) it follows $R_{1,x_1,T} = o(1)$ for any $x_1 \in \mathcal{X}_1$. Furthermore, from Lemma B.10 (i):

$$\begin{aligned} & \frac{1}{T h_{x_1,T}^{d_{X_1}}} \sum_{j=1}^{\infty} \frac{\nu_{x_1,j}}{\left(\lambda_{x_1,T} + \nu_{x_1,j}\right)^2} \left\|\phi_{x_1,j}\right\|_{L^2(\mathcal{X}_2)}^2 E\left[Z_{x_1,j,T}^2\right] \\ = & \left(\frac{\omega^2 f_{X_1}(x_1)}{T h_{x_1,T}^{d_{X_1}}} \sum_{j=1}^{\infty} \frac{\nu_{x_1,j}}{\left(\lambda_{x_1,T} + \nu_{x_1,j}\right)^2} \left\|\phi_{x_1,j}\right\|_{L^2(\mathcal{X}_2)}^2\right) \left(1 + R_{2,x_1,T}\right), \end{aligned} \qquad \text{(TR.15)}$$

with $R_{2,x_1,T} = o(1)$ for any $x_1 \in \mathcal{X}_1$. From (TR.14) and (TR.15), the conclusion follows.

7 Proof of Lemma A.7

We have $\mathcal{B}^e_{x_1,T} = \left(\lambda_{x_1,T} + A^*_{x_1}A_{x_1}\right)^{-1} A^*_{x_1}\zeta_{x_1}$, where $\zeta_{x_1}(z_1) = \int (y - \varphi_{x_1,0}(x_2)) \frac{E\left[\hat{f}_{W,Z}(w,z)\right] - f_{W,Z}(w,z)}{f_Z(z)} dw.$

Lemma B. 12 *Under Assumptions B.1, B.2, B.4 and B.6:* $\zeta_{x_1} = h^m_{x_1,T}\Xi_{x_1} + \Gamma_{x_1}$, *where* Γ_{x_1} *is such that* $\|\Gamma_{x_1}\|_{L^2_{x_1}(\mathcal{Z}_1)} = O(h_T h^{m-1}_{x_1,T} + h^m_T)$.

From Lemma B.12 and by using $\| \left(\lambda_{x_1,T} + A^*_{x_1}A_{x_1}\right)^{-1} A^*_{x_1}\|_{\mathcal{L}(L^2_{x_1}(\mathcal{Z}_1),L^2(\mathcal{X}_2))} = O(1/\sqrt{\lambda_{x_1,T}})$, we get $\left\|\left(\lambda_{x_1,T} + A^*_{x_1}A_{x_1}\right)^{-1} A^*_{x_1}\Gamma_{x_1}\right\|_{L^2(\mathcal{X}_2)} = O\left(\frac{h_T h^{m-1}_{x_1,T} + h^m_T}{\sqrt{\lambda_{x_1,T}}}\right)$. The conclusion follows from definition $\|\mathcal{B}^r_{x_1,T} + h^m_{x_1,T}\left(\lambda_{x_1,T} + A^*_{x_1}A_{x_1}\right)^{-1} A^*_{x_1}\Xi_{x_1}\|_{L^2(\mathcal{X}_2)} = b(\lambda_{x_1,T}, h_{x_1,T})$ and the condition $\frac{h_T h^{m-1}_{x_1,T} + h^m_T}{\sqrt{\lambda_{x_1,T}}} = o\left(b(\lambda_{x_1,T}, h_{x_1,T})\right)$.

8 Proof of results in Section 6 and Appendix 4

8.1 Characterization of operator $\mathcal{D}$ in the Gaussian example

For $\varphi_1, \varphi_2 \in H^1(\mathcal{X}_2)$:

$$\begin{aligned}
\langle \varphi_1, \varphi_2 \rangle_{H^1(\mathcal{X}_2)} &= \int \varphi_1(x_2)\varphi_2(x_2)\phi(x_2)dx_2 + \int \nabla\varphi_1(x_2)\nabla\varphi_2(x_2)\phi(x_2)dx_2 \\
&= \int \varphi_1(x_2)\varphi_2(x_2)\phi(x_2)dx_2 - \int \varphi_1(x_2)\nabla\left(\nabla\varphi_2(x_2)\phi(x_2)\right)dx_2 \\
&= \langle \varphi_1, \mathcal{D}\varphi_2 \rangle_{L^2(\mathcal{X}_2)}
\end{aligned}$$

where the operator $\mathcal{D}$ is defined by $\mathcal{D} = 1 - \nabla^2 - (\nabla \log \phi)\nabla = 1 - \nabla^2 + x_2\nabla$. Thus $A^*A = \mathcal{D}^{-1}\tilde{A}A$.

8.2 Proof of Lemma A.8

When $\alpha_4 > 2\alpha_2$, the quantity $\sum_{j=1}^{\infty} \frac{a_j}{(\lambda + \nu_j)^2}$ converges to $\sum_{j=1}^{\infty} a_j\nu_j^{-2} < \infty$ as $\lambda \to 0$. Let us now consider the case $\alpha_4 < 2\alpha_2$. Without loss of generality, let $\nu_j = j^{-\alpha_1}e^{-\alpha_2 j}$ and $a_j = j^{-\alpha_3}e^{-\alpha_4 j}$. Then:

$$\sum_{j=1}^{\infty} \frac{a_j}{(\lambda + \nu_j)^2} = \sum_{j=1}^{\infty} \frac{a_j\nu_j^{-2}}{(\lambda w_j + 1)^2} = \sum_{j=1}^{\infty} \frac{w_j^{\delta}}{(\lambda w_j + 1)^2} j^{\rho}$$

where $w_j := 1/\nu_j$, $\rho := \frac{\alpha_1\alpha_4}{\alpha_2} - \alpha_3$ and $\delta := \frac{2\alpha_2 - \alpha_4}{\alpha_2} \in (0, 2)$.

Define:

$$J(\lambda) := \lambda^{\delta} n_{\lambda}^{-\rho} \sum_{j=1}^{\infty} \frac{a_j}{(\lambda+\nu_j)^2} = \sum_{j=1}^{\infty} \frac{(\lambda w_j)^{\delta}}{(\lambda w_j + 1)^2} \left(\frac{j}{n_{\lambda}}\right)^{\rho}.$$

The conclusion follows if we show that $J(\lambda) \asymp 1$. We give the proof when $\rho \geq 0$ (similar arguments apply when $\rho < 0$). We split $J(\lambda)$ as

$$\begin{aligned} J(\lambda) &= \sum_{j=1}^{N_1(\lambda)} \frac{(\lambda w_j)^{\delta}}{(\lambda w_j + 1)^2} \left(\frac{j}{n_{\lambda}}\right)^{\rho} + \sum_{j=N_1(\lambda)+1}^{N_2(\lambda)} \frac{(\lambda w_j)^{\delta}}{(\lambda w_j + 1)^2} \left(\frac{j}{n_{\lambda}}\right)^{\rho} + \sum_{j=N_2(\lambda)+1}^{\infty} \frac{(\lambda w_j)^{\delta}}{(\lambda w_j + 1)^2} \left(\frac{j}{n_{\lambda}}\right)^{\rho} \\ &=: J_1(\lambda) + J_2(\lambda) + J_3(\lambda), \end{aligned}$$

where $N_1(\lambda), N_2(\lambda) \in \mathbb{N}$ are such that $N_1(\lambda) < n_{\lambda} < N_2(\lambda)$, $\frac{\nu_{n_{\lambda}}}{\nu_{N_1(\lambda)}} = o(1)$, $\left(\frac{N_2(\lambda)}{n_{\lambda}}\right)^{\rho} \left(\frac{\nu_{N_2(\lambda)}}{\nu_{n_{\lambda}}}\right)^{2-\delta} = o(1)$ and $r_{\lambda} := \max\left\{\frac{n_{\lambda} - N_1(\lambda)}{n_{\lambda}}, \frac{N_2(\lambda) - n_{\lambda}}{n_{\lambda}}\right\} = o(1)$. First we show that $J_i(\lambda) = o(1)$, for $i = 1, 3$. We have

$$\begin{aligned} J_1(\lambda) &\leq \sum_{j=1}^{N_1(\lambda)} \frac{(\lambda w_j)^{\delta}}{(\lambda w_j + 1)^2} \leq \lambda^{\delta} \sum_{j=1}^{N_1(\lambda)} w_j^{\delta} \asymp \frac{1}{w_{n_{\lambda}}^{\delta}} w_{N_1(\lambda)}^{\delta} = \left(\frac{\nu_{n_{\lambda}}}{\nu_{N_1(\lambda)}}\right)^{\delta} \\ &= o(1), \end{aligned} \tag{TR.16}$$

where we used $\sum_{j=1}^{n} w_j^{\delta} = O\left(w_n^{\delta}\right)$ as $n \to \infty$. Similarly,

$$J_3(\lambda) = \sum_{j=N_2(\lambda)+1}^{\infty} \frac{(\lambda w_j)^{\delta}}{(\lambda w_j + 1)^2} \left(\frac{j}{n_{\lambda}}\right)^{\rho} \leq \frac{\lambda^{-2+\delta}}{n_{\lambda}^{\rho}} \sum_{j=N_2(\lambda)+1}^{\infty} \nu_j^{2-\delta} j^{\rho} \tag{TR.17}$$

$$\leq \frac{\lambda^{-2+\delta}}{n_{\lambda}^{\rho}} \sum_{j=N_2(\lambda)+1}^{\infty} e^{-\tau j} j^{m}, \tag{TR.18}$$

where $\tau = (2-\delta)\alpha_2 > 0$ and $m = \lceil \rho - \alpha_1(2-\delta) \rceil \leq \rho$. Now, by using:

$$\sum_{j=n}^{\infty} e^{-\tau j} j^{m} = \left(-\frac{d}{d\tau}\right)^{m} \sum_{j=n}^{\infty} e^{-\tau j} = \left(-\frac{d}{d\tau}\right)^{m} \frac{e^{-\tau n}}{1 - e^{-\tau}} = O\left(n^{m} e^{-\tau n}\right),$$

as $n \to \infty$, we get:

$$J_3(\lambda) = O\left(\frac{\lambda^{-2+\delta}}{n_{\lambda}^{\rho}} N_2(\lambda)^{\rho} \nu_{N_2(\lambda)}^{2-\delta}\right) = O\left(\left(\frac{N_2(\lambda)}{n_{\lambda}}\right)^{\rho} \left(\frac{\nu_{N_2(\lambda)}}{\nu_{n_{\lambda}}}\right)^{2-\delta}\right) = o(1).$$

Second we can write $J_2(\lambda)$ as

$$J_2(\lambda) = \sum_{j=N_1(\lambda)+1}^{N_2(\lambda)} \frac{(\lambda w_j)^{\delta}}{(\lambda w_j + 1)^2} + \sum_{j=N_1(\lambda)+1}^{N_2(\lambda)} \frac{(\lambda w_j)^{\delta}}{(\lambda w_j + 1)^2} \left[\left(1 + \frac{j - n_{\lambda}}{n_{\lambda}}\right)^{\rho} - 1\right], \tag{TR.19}$$

and bound the second term in the RHS. For $N_1(\lambda)+1 \leq j \leq N_2(\lambda)$, the variable $x := \dfrac{j-n_\lambda}{n_\lambda}$ is such that $|x| \leq r_\lambda$. Thus, $\left|\left(1+\dfrac{j-n_\lambda}{n_\lambda}\right)^\rho - 1\right| = |(1+x)^\rho - 1| = O(r_\lambda)$, since the function $x \longmapsto (1+x)^\rho$ has bounded derivative around 0. We deduce that the second term in the RHS of (TR.19) is $o\left(\sum_{j=N_1(\lambda)+1}^{N_2(\lambda)} \dfrac{(\lambda w_j)^\delta}{(\lambda w_j+1)^2}\right)$. Hence we get that the sum is such that $J(\lambda) = \left(\sum_{j=N_1(\lambda)+1}^{N_2(\lambda)} \dfrac{(\lambda w_j)^\delta}{(\lambda w_j+1)^2}\right)[1+o(1)] + o(1)$. Moreover, note that

$$\begin{aligned}
\sum_{j=N_1(\lambda)+1}^{N_2(\lambda)} \frac{(\lambda w_j)^\delta}{(\lambda w_j+1)^2} &= \sum_{j=1}^{\infty} \frac{(\lambda w_j)^\delta}{(\lambda w_j+1)^2} - \sum_{j=1}^{N_1(\lambda)} \frac{(\lambda w_j)^\delta}{(\lambda w_j+1)^2} - \sum_{j=N_2(\lambda)+1}^{\infty} \frac{(\lambda w_j)^\delta}{(\lambda w_j+1)^2} \\
&= \sum_{j=1}^{\infty} \frac{(\lambda w_j)^\delta}{(\lambda w_j+1)^2} + o(1),
\end{aligned}$$

from similar arguments as above and as in (TR.16), (TR.17). Thus, we have proved that

$$J(\lambda) = \sum_{j=1}^{\infty} \frac{(\lambda w_j)^\delta}{(\lambda w_j+1)^2}[1+o(1)] + o(1). \tag{TR.20}$$

The conclusion follows if we show that:

$$\sum_{j=1}^{\infty} \frac{(\lambda w_j)^\delta}{(\lambda w_j+1)^2} \asymp 1.$$

This follows by using that:

$$\sum_{j=1}^{\infty} \frac{(\lambda w_j)^\delta}{(\lambda w_j+1)^2} = \sum_{j=1}^{n_\lambda-1} \frac{(\lambda w_j)^\delta}{(\lambda w_j+1)^2} + \frac{(\lambda w_{n_\lambda})^\delta}{(\lambda w_{n_\lambda}+1)^2} + \sum_{j=n_\lambda+1}^{\infty} \frac{(\lambda w_j)^\delta}{(\lambda w_j+1)^2},$$

and:

$$\frac{(\lambda w_{n_\lambda})^\delta}{(\lambda w_{n_\lambda}+1)^2} \asymp 1,$$

$$0 < \sum_{j=1}^{n_\lambda-1} \frac{(\lambda w_j)^\delta}{(\lambda w_j+1)^2} \leq \sum_{j=1}^{n_\lambda-1} (\lambda w_j)^\delta = O\left((\lambda w_{n_\lambda})^\delta\right) = O(1),$$

$$0 < \sum_{j=n_\lambda+1}^{\infty} \frac{(\lambda w_j)^\delta}{(\lambda w_j+1)^2} \leq \lambda^{-2+\delta} \sum_{j=n_\lambda+1}^{\infty} \nu_j^{2-\delta} = \left(\frac{\nu_{n_\lambda}}{\lambda}\right)^{2-\delta} = O(1).$$

8.3 Asymptotic expansion of $b_{x_1}\left(\lambda_{x_1,T}, h_{x_1,T}\right)^2$

In this subsection, we show that:

$$b_{x_1}\left(\lambda_{x_1,T}, h_{x_1,T}\right)^2 \asymp \lambda_{x_1,T}^{2\delta} n_{\lambda_{x_1,T}}^{2\alpha_1\delta-\beta} + h_{x_1,T}^{2m} \lambda_{x_1,T}^{2\rho-1} n_{\lambda_{x_1,T}}^{2\alpha_1\rho-\beta}. \tag{TR.21}$$

For this purpose, write:

$$\begin{aligned} b_{x_1}\left(\lambda_{x_1,T}, h_{x_1,T}\right)^2 &= \lambda_{x_1,T}^2 \sum_{j=1}^{\infty} \frac{d_{x_1,j}^2}{\left(\lambda_{x_1,T}+\nu_j\right)^2} \left\|\phi_j\right\|_{L^2(\mathcal{X}_2)}^2 + h_{x_1,T}^{2m} \sum_{j=1}^{\infty} \frac{\nu_j \xi_{x_1,j}^2}{\left(\lambda_{x_1,T}+\nu_j\right)^2} \left\|\phi_j\right\|_{L^2(\mathcal{X}_2)}^2 \\ &\quad -2\lambda_{x_1,T} h_{x_1,T}^m \sum_{j=1}^{\infty} \frac{d_{x_1,j}\sqrt{\nu_j}\xi_{x_1,j}}{\left(\lambda_{x_1,T}+\nu_j\right)^2} \left\|\phi_j\right\|_{L^2(\mathcal{X}_2)}^2 =: B_{1,T} + B_{2,T} - 2B_{3,T}. \end{aligned}$$

From Lemma A.8, we have:

$$B_{1,T} \asymp \lambda_{x_1,T}^{2\delta} n_{\lambda_{x_1,T}}^{2\alpha_1\delta-\beta}, \quad B_{2,T} \asymp h_{x_1,T}^{2m} \lambda_{x_1,T}^{2\rho-1} n_{\lambda_{x_1,T}}^{2\alpha_1\rho-\beta}.$$

Moreover, by the Cauchy-Schwarz inequality, $|B_{3,T}| \leq B_{1,T}^{1/2} B_{2,T}^{1/2}$. We distinguish three cases. (i) When $\lambda_{x_1,T}$ and $h_{x_1,T}$ are such that $B_{2,T}/B_{1,T} = o(1)$, then it follows that:

$$\begin{aligned} b_{x_1}\left(\lambda_{x_1,T}, h_{x_1,T}\right)^2 &= B_{1,T}\left(1 + O(\sqrt{B_{2,T}/B_{1,T}})\right) + B_{2,T} = B_{1,T}(1+o(1)) + B_{2,T} \\ &\asymp \lambda_{x_1,T}^{2\delta} n_{\lambda_{x_1,T}}^{2\alpha_1\delta-\beta} + h_{x_1,T}^{2m} \lambda_{x_1,T}^{2\rho-1} n_{\lambda_{x_1,T}}^{2\alpha_1\rho-\beta}, \end{aligned}$$

which yields (TR.21). (ii) When $B_{1,T}/B_{2,T} = o(1)$ a similar argument shows that (TR.21) holds.

(iii) Let us now consider the case where $B_{1,T} \asymp B_{2,T}$. We have:

$$\begin{aligned} b_{x_1}\left(\lambda_{x_1,T}, h_{x_1,T}\right)^2 &= \lambda_{x_1,T}^2 \sum_{j=1}^{\infty} \frac{d_{x_1,j}^2}{\left(\lambda_{x_1,T}+\nu_j\right)^2} \left(1 - \frac{h_{x_1,T}^m}{\lambda_{x_1,T}} \frac{\sqrt{\nu_j}\xi_{x_1,j}}{d_{x_1,j}}\right)^2 \left\|\phi_j\right\|_{L^2(\mathcal{X}_2)}^2 \\ &\geq \lambda_{x_1,T}^2 \sum_{l=1}^{\infty} \frac{d_{x_1,j_l}^2}{\left(\lambda_{x_1,T}+\nu_{j_l}\right)^2} \left\|\phi_{j_l}\right\|_{L^2(\mathcal{X}_2)}^2, \end{aligned}$$

where j_l, $l \in \mathbb{N}$, denotes the sequence of indices such that $\xi_{x_1,j} d_{x_1,j} \leq 0$ (see Condition (b) in Proposition 4). Now, let $N := \sup_{l\in\mathbb{N}}(j_{l+1} - j_l) < \infty$. Then, for any $\lambda_{x_1,T}$, there exists l such that $\left|j_l - n_{\lambda_{x_1,T}}\right| \leq N$. Then, we get:

$$b_{x_1}\left(\lambda_{x_1,T}, h_{x_1,T}\right)^2 \geq C\lambda_{x_1,T}^2 \frac{d_{x_1,n_{\lambda_{x_1,T}}}^2}{\left(\lambda_{x_1,T}+\nu_{n_{\lambda_{x_1,T}}}\right)^2} \left\|\phi_{n_{\lambda_{x_1,T}}}\right\|_{L^2(\mathcal{X}_2)}^2,$$

for some constant C. By the arguments in the proof of Lemma A.8, we know that:

$$\lambda_{x_1,T}^2 \frac{d_{x_1,n_{\lambda_{x_1,T}}}^2}{\left(\lambda_{x_1,T}+\nu_{n_{\lambda_{x_1,T}}}\right)^2} \left\|\phi_{n_{\lambda_{x_1,T}}}\right\|_{L^2(\mathcal{X}_2)}^2 \asymp \lambda_{x_1,T}^{2\delta} n_{\lambda_{x_1,T}}^{2\alpha_1\delta-\beta} \asymp B_{1,T}.$$

Since $B_{1,T} \asymp B_{2,T}$, we get $b_{x_1}\left(\lambda_{x_1,T}, h_{x_1,T}\right)^2 \geq C\left(B_{1,T}+B_{2,T}\right)$. Since $b_{x_1}\left(\lambda_{x_1,T}, h_{x_1,T}\right)^2 \leq B_{1,T} + B_{2,T} + 2B_{1,T}^{1/2} B_{2,T}^{1/2} \asymp B_{1,T} + B_{2,T}$, (TR.21) follows.

9 Proofs of Lemmas B and C

As explained in Appendix A.1, the estimator of the density $f_{Y,X_2,Z}$ is $\hat{f}_{Y,X_2,Z} = \hat{f}_{Y,X_2,Z^*}/\hat{\omega}$, where $\hat{f}_{Y,X_2,Z^*}$ is the kernel estimator computed on the sample $\{(Y_t, X_{2,t}, Z_t^*), \, t = 1, ..., T^*\}$ and $\hat{\omega} = \displaystyle\int_{\mathcal{Z}} \hat{f}_{Z^*,\tau}(z)dz$, with $\hat{f}_{Z^*,\tau} = \max\left\{\hat{f}_{Z^*}, (\log T)^{-1}\right\}$. The trimmed normalization factor $\hat{\omega}$ is such that $\hat{\omega} \geq (\log T)^{-1}$ a.s., and converges in probability to $\omega = \int_{\mathcal{Z}} f_{Z^*}(z)dz > 0$ at a parametric rate. Similarly, the estimator of f_Z is $\hat{f}_Z = \hat{f}_{Z^*}/\hat{\omega}$, while the estimator of the conditional density $f_{Y,X_2|Z}$ is $\hat{f}_{Y,X_2|Z} = \hat{f}_{Y,X_2,Z^*}/\hat{f}_{Z^*,\tau}$. Note that $\hat{f}_{Y,X_2|Z} \neq \hat{f}_{Y,X_2,Z}/\hat{f}_Z$ although both sides are consistent estimators of $f_{Y,X_2|Z}$. We prefer to work with the former because of trimming.

Estimator $\hat{f}_Z = \hat{f}_{Z^*}/\hat{\omega}$ admits the same asymptotic properties as the unfeasible estimator $\tilde{f}_Z = \hat{f}_{Z^*}/\omega$, both in probability and in mean square sense. Indeed, we have:

$$\hat{f}_Z(z) - f_Z(z) = \tilde{f}_Z(z) - f_Z(z) + \left(\frac{1}{\hat{\omega}} - \frac{1}{\omega}\right)\hat{f}_{Z^*}(z).$$

Then, we get:

$$E\left[\left(\hat{f}_Z(z) - f_Z(z)\right)^2\right]^{1/2} = E\left[\left(\tilde{f}_Z(z) - f_Z(z)\right)^2\right]^{1/2} + O\left(\frac{\log T}{\sqrt{T}}\right), \qquad \text{(TR.22)}$$

uniformly in $z \in \mathcal{Z}$, by using the triangular inequality, the Cauchy-Schwartz inequality and:

$$E\left[\left(\frac{1}{\hat{\omega}} - \frac{1}{\omega}\right)^2\right] \leq \frac{(\log T)^2}{\omega^2} E\left[(\hat{\omega} - \omega)^2\right] = O\left(\frac{(\log T)^2}{T}\right).$$

Equation (TR.22) means that we can derive the asymptotic properties of $\hat{f}_Z$ in mean square sense from those of kernel estimator $\hat{f}_{Z^*}$. We deduce from assumptions B.1 (i)-(ii) and B.2, and uniform convergence results similar to Hansen (2008):

$$\sup_{z_1 \in \mathcal{Z}_1} E\left[\left(\hat{f}_Z(z) - f_Z(z)\right)^2\right] = O\left(\frac{\log T}{T h_T^{d_{Z_1}} h_{x_1,T}^{d_{X_1}}} + h_T^{2m} + h_{x_1,T}^{2m}\right),$$

uniformly on $x_1 \in \mathcal{X}_1$. In particular, estimator $\hat{f}_Z$ does not feature a boundary bias on $\mathcal{Z}$. A similar result holds in probability.

By similar arguments, it is possible to show that the asymptotic properties of $\hat{f}_{Y,X_2,Z}$ can be deduced from those of $\hat{f}_{Y,X_2,Z^*}$ in a similar vein as above. In order to avoid a very lengthy exposition, in the proofs of Lemmas B.1-B.12 and C.1-C.2 we omit the normalization factor $\hat{\omega}$ in $\hat{f}_{Y,X_2,Z}$ and $\hat{f}_Z$, and write $T = T^*$, $Z = Z^*$. Moreover, we adopt a product kernel in the estimation of the density of (Y, X_2, Z) in $\mathbb{R}^d$. We use the generic notation K for both the d-dimensional product kernel and each of its components.

9.1 Proof of Lemma B.1

We have:

$$\left(\widetilde{\hat{A}}_{x_1}\hat{A}_{x_1}\right)\varphi(x_2) = \int\left(\int \hat{f}_{X_2|Z}(x_2|z)\hat{f}_{X_2|Z}(\xi_2|z)\hat{\Omega}_{x_1}(z_1)\,\hat{f}_{Z_1|X_1}(z_1|x_1)dz_1\right)\varphi(\xi_2)\,d\xi_2,$$

and:

$$\left(\tilde{A}_{x_1} A_{x_1}\right) \varphi(x_2) = \int \left(\int f_{X_2|Z}(x_2|z) f_{X_2|Z}(\xi_2|z) \Omega_{x_1,0}(z_1) f_{Z_1|X_1}(z_1|x_1) dz_1 \right) \varphi(\xi_2) d\xi_2.$$

Thus:

$$\left\| \widetilde{\widehat{A}}_{x_1} \widehat{A}_{x_1} - \tilde{A}_{x_1} A_{x_1} \right\|^2_{L^2(\mathcal{X}_2)} \leq \int \int \hat{\alpha}_{x_1}(x_2, \xi_2)^2 dx_2 d\xi_2, \tag{TR.23}$$

where:

$$\begin{aligned} \hat{\alpha}_{x_1}(x_2, \xi_2) &= \int \hat{f}_{X_2|Z}(x_2|z) \hat{f}_{X_2|Z}(\xi_2|z) \hat{\Omega}_{x_1}(z_1) \hat{f}_{Z_1|X_1}(z_1|x_1) dz_1 \\ &\quad - \int f_{X_2|Z}(x_2|z) f_{X_2|Z}(\xi_2|z) \Omega_{x_1,0}(z_1) f_{Z_1|X_1}(z_1|x_1) dz_1. \end{aligned}$$

Let us decompose:

$$\begin{aligned} \hat{\alpha}_{x_1}(x_2, \xi_2) &= \int \Delta \hat{f}_{X_2|Z}(x_2|z) f_{X_2|Z}(\xi_2|z) \Omega_{x_1,0}(z_1) f_{Z_1|X_1}(z_1|x_1) dz_1 \\ &\quad + \int f_{X_2|Z}(x_2|z) \Delta \hat{f}_{X_2|Z}(\xi_2|z) \Omega_{x_1,0}(z_1) f_{Z_1|X_1}(z_1|x_1) dz_1 \\ &\quad + \int f_{X_2|Z}(x_2|z) f_{X_2|Z}(\xi_2|z) \Delta \hat{\Omega}_{x_1}(z_1) f_{Z_1|X_1}(z_1|x_1) dz_1 \\ &\quad + \int f_{X_2|Z}(x_2|z) f_{X_2|Z}(\xi_2|z) \Omega_{x_1}(z_1) \Delta \hat{f}_{Z_1|X_1}(z_1|x_1) dz_1 \\ &\quad + I_2 \\ &=: I_{1,1} + I_{1,2} + I_{1,3} + I_{1,4} + I_2, \end{aligned} \tag{TR.24}$$

where I_2 contains higher-order terms. From results similar to Lemma C.1 and by using Assumptions B.1-B.3 (i) and B.6, we have $\sup_{x_2, \xi_2 \in \mathcal{X}_2} I_{1,i} = O_p\left(\sqrt{\frac{\log T}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{X_2}}}} + h_T^m + h_{x_1,T}^m \right)$, $i = 1, 2$, and $\sup_{x_2, \xi_2 \in \mathcal{X}_2} I_{1,4} = O_p\left(\sqrt{\frac{\log T}{T h_{x_1,T}^{d_{X_1}}}} + h_T^m + h_{x_1,T}^m \right)$, uniformly in $x_1 \in \mathcal{X}_1$. Moreover, from Assumption B.7 (ii) we have $\sup_{x_2, \xi_2 \in \mathcal{X}_2} I_{1,3} = O_p\left(\sqrt{\frac{\log T}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}}}} + h_T^m + h_{x_1,T}^m \right)$, uniformly in $x_1 \in \mathcal{X}_1$. Finally, from Assumptions B.1-B.3 (i), B.6 and B.7 (ii), and by using $\sup_{x_2 \in \mathcal{X}_2, z_1 \in \mathcal{Z}_1} \left| \Delta \hat{f}_{X_2|Z}(x_2|z) \right| = O_p\left(\sqrt{\frac{\log T}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{X_2}+d_{Z_1}}}} + h_{x_1,T}^m + h_T^m \right)$, $\sup_{z_1 \in \mathcal{Z}_1} \left| \Delta \hat{f}_{Z_1|X_1}(z_1|x_1) \right| = O_p\left(\sqrt{\frac{\log T}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}}}} + h_{x_1,T}^m + h_T^m \right)$ and $\frac{\log T}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}+d_{X_2}}} = O(1)$, uniformly in $x_1 \in \mathcal{X}_1$, we

get $\sup_{x_2,\xi_2\in\mathcal{X}_2} I_2 = O_p\left(\left(\left(\sqrt{\frac{\log T}{Th_{x_1,T}^{d_{X_1}}h_T^{d_{X_2}+d_{Z_1}}}}+h_{x_1,T}^m+h_T^m\right)\left(\sqrt{\frac{\log T}{Th_{x_1,T}^{d_{X_1}}h_T^{d_{Z_1}}}}+h_T^m+h_{x_1,T}^m\right)\right)\right)$
$= O_p\left(\sqrt{\frac{\log T}{Th_{x_1,T}^{d_{X_1}}h_T^{d_{Z_1}}}}+h_T^m+h_{x_1,T}^m\right)$, uniformly in $x_1\in\mathcal{X}_1$. The conclusion follows.

9.2 Proof of Lemma B.2

We have:

$$\begin{aligned}
&\left(\widetilde{\hat{A}}_{x_1}-\tilde{A}_{x_1}\right)\left(\hat{\psi}_{x_1}+\zeta_{x_1}\right)(x_2)\\
=\;&\frac{1}{f_{X_1}(x_1)}\int\int g_{x_1,0}(w)\hat{\Omega}_{x_1}(z_1)\,\hat{f}_{X_2|Z}(x_2|z)\frac{\hat{f}_{Z_1|X_1}(z_1|x_1)}{f_{Z_1|X_1}(z_1|x_1)}\left(\hat{f}_{W,Z}(w,z)-f_{W,Z}(w,z)\right)dwdz_1\\
&-\frac{1}{f_{X_1}(x_1)}\int\int g_{x_1,0}(w)\Omega_{x_1,0}(z_1)\,f_{X_2|Z}(x_2|z)\left(\hat{f}_{W,Z}(w,z)-f_{W,Z}(w,z)\right)dwdz_1,
\end{aligned}$$

where $g_{x_1,0}(w):=y-\varphi_{x_1,0}(x_2)$. Then:

$$\begin{aligned}
&\left(\widetilde{\hat{A}}_{x_1}-\tilde{A}_{x_1}\right)\left(\hat{\psi}_{x_1}+\zeta_{x_1}\right)(x_2)\\
=\;&\frac{1}{f_{X_1}(x_1)}\int\int g_{x_1,0}(w)\Delta\hat{\Omega}_{x_1}(z_1)\,f_{X_2|Z}(x_2|z)\Delta\hat{f}_{W,Z}(w,z)dwdz_1\\
&+\frac{1}{f_{X_1}(x_1)}\int\int g_{x_1,0}(w)\Omega_{x_1,0}(z_1)\,\frac{1}{f_Z(z)}\Delta\hat{f}_{X_2,Z}(x_2,z)\Delta\hat{f}_{W,Z}(w,z)dwdz_1\\
&-\frac{1}{f_{X_1}(x_1)^2}\Delta\hat{f}_{X_1}(x_1)\int\int g_{x_1,0}(w)\Omega_{x_1,0}(z_1)\,f_{X_2|Z}(x_2|z)\Delta\hat{f}_{W,Z}(w,z)dwdz_1\\
&+I_2\\
&=: I_{1,1}+I_{1,2}+I_{1,3}+I_2, \qquad \text{(TR.25)}
\end{aligned}$$

where I_2 is a higher-order term.

i) Bound of $I_{1,1}$. Write $I_{1,1}=\frac{1}{f_{X_1}(x_1)}\int\Delta\hat{\Omega}_{x_1}(z_1)\,f_{X_2|Z}(x_2|z)\hat{\Psi}_{x_1}(z_1)dz_1$, where $\hat{\Psi}_{x_1}(z_1)$ is defined in (TR.9). We have $I_{1,1}\le\frac{\int_{\mathcal{Z}_1}dz_1}{f_{X_1}(x_1)}\|f_{X_2|Z}\|_\infty\sup_{z_1\in\mathcal{Z}_1}|\Delta\hat{\Omega}_{x_1}(z_1)|\sup_{z_1\in\mathcal{Z}_1}|\hat{\Psi}_{x_1}(z_1)|$, uniformly in $x_1\in\mathcal{X}_1$, $x_2\in\mathcal{X}_2$. From Lemma C.1 (1), (2) and Assumptions B.3, B.7 (ii), we get:

$$I_{1,1}=O_p\left(\frac{\log T}{Th_{x_1,T}^{d_{X_1}}h_T^{d_{Z_1}}}+h_{x_1,T}^{2m}+h_T^{2m}\right),$$

uniformly in $x_1\in\mathcal{X}_1$, $x_2\in\mathcal{X}_2$.

ii) Bound $I_{1,2}$. Write $I_{1,2}=\frac{1}{f_{X_1}(x_1)}\int\Omega_{x_1,0}(z_1)\frac{1}{f_Z(z)}\Delta\hat{f}_{X_2,Z}(x_2,z)\hat{\Psi}_{x_1}(z_1)dz_1$. Thus, we have $I_{1,1}\le\frac{\int_{\mathcal{Z}_1}dz_1}{f_{X_1}(x_1)}\|\Omega_0/f_Z\|_\infty\sup_{z_1\in\mathcal{Z}_1}|\Delta\hat{f}_{X_2,Z}(x_2,z)|\sup_{z_1\in\mathcal{Z}_1}|\hat{\Psi}_{x_1}(z_1)|$, uniformly in

$x_1 \in \mathcal{X}_1$, $x_2 \in \mathcal{X}_2$. By using Lemma C.1 (1)-(2), Assumptions B.1-B.3 and B.6, and that $\sup_{z_1 \in \mathcal{Z}_1} |\Delta \hat{f}_{X_2,Z}(x_2,z)| = O_p\left(\sqrt{\frac{\log T}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}+d_{X_2}}}} + h_{x_1,T}^m + h_T^m\right)$, uniformly in $x_1 \in \mathcal{X}_1$, $x_2 \in \mathcal{X}_2$, we get:

$$I_{1,2} = O_p\left(\frac{\log T}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}+d_{X_2}}} + h_{x_1,T}^{2m} + h_T^{2m}\right),$$

uniformly in $x_1 \in \mathcal{X}_1$, $x_2 \in \mathcal{X}_2$.

iii) Bound of $I_{1,3}$. By similar arguments as in i) and ii) and using Lemma C.1 (3) we get:

$$I_{1,3} = O_p\left(\frac{\log T}{T h_{x_1,T}^{d_{X_1}}} + h_{x_1,T}^{2m} + h_T^{2m}\right),$$

uniformly in $x_1 \in \mathcal{X}_1$, $x_2 \in \mathcal{X}_2$. The conclusion follows.

9.3 Proof of Lemma B.3

We have:

$$\begin{aligned}
& -\widetilde{\hat{A}}_{x_1} \hat{q}_{x_1}(x_2) \\
= & \frac{1}{f_{X_1}(x_1)} \int\int g_{x_1,0}(w) \hat{\Omega}_{x_1}(z_1) \hat{f}_{X_2|Z}(x_2|z) \frac{\hat{f}_{Z_1|X_1}(z_1|x_1)}{f_{Z_1|X_1}(z_1|x_1)} \frac{\Delta \hat{f}_Z(z)}{\hat{f}_Z(z)} \Delta \hat{f}_{W,Z}(w,z) dw dz_1 \\
= & \frac{1}{f_{X_1}(x_1)} \int\int g_{x_1,0}(w) \Omega_{x_1,0}(z_1) f_{X_2|Z}(x_2|z) \frac{\Delta \hat{f}_Z(z)}{f_Z(z)} \Delta \hat{f}_{W,Z}(w,z) dw dz_1 \\
& + \frac{1}{f_{X_1}(x_1)} \int\int g_{x_1,0}(w) \Delta \hat{\Omega}_{x_1}(z_1) f_{X_2|Z}(x_2|z) \frac{\Delta \hat{f}_Z(z)}{f_Z(z)} \Delta \hat{f}_{W,Z}(w,z) dw dz_1 \\
& + I_3, \\
& =: I_1 + I_2 + I_3,
\end{aligned}$$

where I_3 is a term of higher order. Let us first consider I_1. Write:

$$I_1 = \frac{1}{f_{X_1}(x_1)} \int \Omega_{x_1,0}(z_1) f_{X_2|Z}(x_2|z) \frac{\Delta \hat{f}_Z(z)}{f_Z(z)} \hat{\Psi}_{x_1}(z_1) dz_1.$$

Thus, we have $I_1 \leq \frac{1}{f_{X_1}(x_1)} \|\Omega_0 f_{X_2|Z}/f_Z\|_\infty \sup_{z_1 \in \mathcal{Z}_1} |\Delta \hat{f}_Z(z)| \sup_{z_1 \in \mathcal{Z}_1} |\hat{\Psi}_{x_1}(z_1)|$, uniformly in $x_1 \in \mathcal{X}_1$, $x_2 \in \mathcal{X}_2$. From Lemma C.1 (1), (2) and Assumptions B.1-B.3, B.6 and B.7 (ii), we get:

$$I_1 = O_p\left(\frac{\log T}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}}} + h_T^{2m} + h_{x_1,T}^{2m}\right),$$

uniformly in $x_1 \in \mathcal{X}_1$, $x_2 \in \mathcal{X}_2$. Let us now consider I_2. Write:

$$I_2 = \frac{1}{f_{X_1}(x_1)} \int \Delta \hat{\Omega}_{x_1}(z_1) f_{X_2|Z}(x_2|z) \frac{1}{f_Z(z)} \Delta \hat{f}_Z(z) \hat{\Psi}_{x_1}(z_1) dz_1.$$

Thus, we have $I_2 \leq \frac{1}{f_{X_1}(x_1)} \|f_{X_2|Z}/f_Z\|_\infty \sup_{z_1 \in \mathcal{Z}_1} |\Delta\hat{\Omega}_{x_1}(z_1)| \sup_{z_1 \in \mathcal{Z}_1} |\Delta\hat{f}_Z(z)| \sup_{z_1 \in \mathcal{Z}_1} |\hat{\Psi}_{x_1}(z_1)|$, uniformly in $x_1 \in \mathcal{X}_1$, $x_2 \in \mathcal{X}_2$. From Lemma C.1 (1), (2) and Assumptions B.1-B.3, B.6 and B.7 (ii), we get:

$$I_2 = O_p\left(\left(\sqrt{\frac{\log T}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}}}} + h_T^m + h_{x_1,T}^m\right)^3\right),$$

uniformly in $x_1 \in \mathcal{X}_1$, $x_2 \in \mathcal{X}_2$. By using $\frac{\log T}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}}} = O(1)$, the conclusion follows.

9.4 Proof of Lemma B.4

By using (TR.23) and (TR.24), the conclusion follows if we prove that $\sup_{x_2,\xi_2 \in \mathcal{X}_2} E\left[I_{1,i}(x_2,\xi_2)^{2N}\right] = O\left(a_T^N\right)$, $i = 1, ..., 4$, and $\sup_{x_2,\xi_2 \in \mathcal{X}_2} E\left[I_2(x_2,\xi_2)^{2N}\right] = O\left(a_T^N\right)$ for any $N \in \mathbb{N}$ and $x_1 \in \mathcal{X}_1$.

(i) Bound of $E\left[I_{1,1}(x_2,\xi_2)^{2N}\right]$. The kernel estimator of f_Z in the denominator of $\hat{f}_{X_2|Z}$ is replaced by the trimmed estimator $\hat{f}_{Z,\tau} = \max\left\{\hat{f}_Z, \tau_T\right\}$, where the trimming sequence is $\tau_T = (\log T)^{-1}$. Write:

$$\begin{aligned} I_{1,1}(x_2,\xi_2) &= \frac{1}{f_{X_1}(x_1)} \int \Delta\hat{f}_{X_2,Z}(x_2,z) f_{X_2|Z}(\xi_2|z)\Omega_{x_1,0}(z_1)\,dz_1 \\ &\quad - \frac{1}{f_{X_1}(x_1)} \int \Delta\hat{f}_{Z,\tau}(z) f_{X_2|Z}(x_2|z) f_{X_2|Z}(\xi_2|z)\Omega_{x_1,0}(z_1)\,dz_1 \\ &\quad - \frac{1}{f_{X_1}(x_1)} \int \Delta\hat{f}_{X_2,Z}(x_2,z)\frac{\Delta\hat{f}_{Z,\tau}(z)}{\hat{f}_{Z,\tau}(z)} f_{X_2|Z}(\xi_2|z)\Omega_{x_1,0}(z_1)\,dz_1 \\ &\quad + \frac{1}{f_{X_1}(x_1)} \int \frac{\left[\Delta\hat{f}_{Z,\tau}(z)\right]^2}{\hat{f}_{Z,\tau}(z)} f_{X_2|Z}(x_2|z) f_{X_2|Z}(\xi_2|z)\Omega_{x_1,0}(z_1)\,dz_1 \\ &=: I_{1,1,1}(x_2,\xi_2) - I_{1,1,2}(x_2,\xi_2) - I_{1,1,3}(x_2,\xi_2) + I_{1,1,4}(x_2,\xi_2). \end{aligned}$$

Let us first consider $I_{1,1,1}(x_2,\xi_2)$. We have:

$$\begin{aligned} E\left[I_{1,1,1}(x_2,\xi_2)^{2N}\right] &= \frac{1}{f_{X_1}(x_1)^{2N}} \int \cdots \int \prod_l f_{X_2|Z}(\xi_2|z_{1,l},x_1)\Omega_{x_1,0}(z_{1,l}) \\ &\quad \cdot E\left[\prod_l \Delta\hat{f}_{X_2,Z}(x_2,z_{1,l},x_1)\right] \prod_l dz_{1,l}, \end{aligned}$$

where the product $\prod_l$ is over $l = 1, ..., 2N$. Write $\Delta\hat{f}_{X_2,Z} = \bar{f}_{X_2,Z} + b_{X_2,Z}$, where $\bar{f}_{X_2,Z} =$

$\hat{f}_{X_2,Z} - E\left[\hat{f}_{X_2,Z}\right]$ and $b_{X_2,Z} = E\left[\hat{f}_{X_2,Z}\right] - f_{X_2,Z}$. Then:

$$\begin{aligned}
& E\left[I_{1,1,1}\left(x_2,\xi_2\right)^{2N}\right] \\
= \; & \frac{1}{f_{X_1}\left(x_1\right)^{2N}} \sum_{J=0}^{2N} \binom{2N}{J} \left(\int E\left[\prod_{l=1}^{J} \bar{f}_{X_2,Z}(x_2, z_{1,l}, x_1)\right] \prod_{l=1}^{J} f_{X_2|Z}(\xi_2|z_{1,l}, x_1) \Omega_{x_1,0}\left(z_{1,l}\right) \prod_{l=1}^{J} dz_{1,l} \right) \\
& \cdot \left(\int b_{X_2,Z}(x_2, z) f_{X_2|Z}(\xi_2|z_1, x_1) \Omega_{x_1,0}\left(z_1\right) dz_1 \right)^{2N-J}.
\end{aligned}$$

From Assumptions B.1, B.2, B.3 (i), B.6 and 5, $\int b_{X_2,Z}(x_2, z) f_{X_2|Z}(\xi_2|z_1, x_1) \Omega_{x_1,0}\left(z_1\right) dz_1 = O\left(h_T^m + h_{x_1,T}^m\right)$, uniformly in $x_2, \xi_2 \in \mathcal{X}_2$. Furthermore, by writing $\bar{f}_{X_2,Z}(x_2, z) = \frac{1}{T}\sum_{t=1}^{T} \kappa_t(x_2, z)$ where $\kappa_t(x_2, z) = K_{h_T}\left(x_2 - X_{2,t}\right) K_{h_T}\left(z_1 - Z_{1,t}\right) K_{h_{x_1,T}}\left(x_1 - X_{1,t}\right) - E\left[K_{h_T}\left(x_2 - X_{2,t}\right) K_{h_T}\left(z_1 - Z_{1,t}\right) K_{h_{x_1,T}}\left(x_1 - X_{1,t}\right)\right]$, we can write $E\left[\prod_{l=1}^{J} \bar{f}_{X_2,Z}(x_2, z_{1,l}, x_1)\right] = \sum_{n=1}^{\lfloor J/2 \rfloor} \frac{D_{T,n}}{T^J} J_n\left(x_1, x_2, z_{1,1}, ..., z_{1,J}\right)$, where $J_n\left(x_1, x_2, z_{1,1}, ..., z_{1,J}\right)$ is a term splitted in a product of n expectations, $D_{T,n} := T\left(T-1\right)\cdots\left(T-n+1\right)$, and $\lfloor J/2 \rfloor$ denotes the largest integer which is smaller or equal to $J/2$. To derive the order of the term in $E\left[I_{1,1,1}\left(x_2,\xi_2\right)^{2N}\right]$ corresponding to $J_n\left(x_1, x_2, z_{1,1}, ..., z_{1,J}\right)$, note that all the powers $h_T^{-d_{Z_1}}$ can be eliminated by a change of variable, while a power $h_{x_1,T}^{-d_{X_1}} h_T^{-d_{X_2}}$ coming from variables X_1 and X_2 can be eliminated for each expectation term contained in $J_n\left(x_1, x_2, z_{1,1}, ..., z_{1,J}\right)$. Thus, $\int J_n\left(x_1, x_2, z_{1,1}, ..., z_{1,J}\right) \prod_{l=1}^{J} f_{X_2|Z}(\xi_2|z_{1,l}, x_1) \Omega_{x_1,0}\left(z_{1,l}\right) \prod_{l=1}^{J} dz_{1,l} = O\left(\frac{1}{h_{x_1,T}^{d_{X_1}(J-n)} h_T^{d_{X_2}(J-n)}}\right)$, uniformly in $x_2, \xi_2 \in \mathcal{X}_2$. This implies $\int E\left[\prod_{l=1}^{J} \bar{f}_{X_2,Z}(x_2, z_{1,l}, x_1)\right] \prod_{l=1}^{J} f_{X_2|Z}(\xi_2|z_{1,l}, x_1) \Omega_{x_1,0}\left(z_{1,l}\right) \prod_{l=1}^{J} dz_{1,l} = O\left(\sum_{n=1}^{\lfloor J/2 \rfloor} \frac{1}{\left(T h_{x_1,T}^{d_{X_1}} h_T^{d_{X_2}}\right)^{J-n}}\right) = O\left(\frac{1}{\left(T h_{x_1,T}^{d_{X_1}} h_T^{d_{X_2}}\right)^{\lceil J/2 \rceil}}\right)$, since $\frac{1}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{X_2}}} = o(1)$. We get $E\left[I_{1,1,1}\left(x_2,\xi_2\right)^{2N}\right] = O\left(\sum_{J=0}^{2N} \frac{1}{\left(T h_{x_1,T}^{d_{X_1}} h_T^{d_{X_2}}\right)^{\lceil J/2 \rceil}} \left(h_T^{2m} + h_{x_1,T}^{2m}\right)^{N-J/2}\right) = O\left(\sum_{k=0}^{N} \left(\frac{1}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{X_2}}}\right)^k \left(h_T^{2m} + h_{x_1,T}^{2m}\right)^{N-k}\right) = O\left(\left(\frac{1}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{X_2}}} + h_T^{2m} + h_{x_1,T}^{2m}\right)^N\right) = O\left(a_T^N\right)$, uniformly in $x_2, \xi_2 \in \mathcal{X}_2$, for any $x_1 \in \mathcal{X}_1$.

Let us now consider $I_{1,1,2}(x_2,\xi_2)$. Write:

$$\begin{aligned} I_{1,1,2}(x_2,\xi_2) &= \frac{1}{f_{X_1}(x_1)}\int \Delta\hat{f}_Z(z) f_{X_2|Z}(x_2|z) f_{X_2|Z}(\xi_2|z)\Omega_{x_1,0}(z_1)\,dz_1 \\ &\quad + \frac{1}{f_{X_1}(x_1)}\int \left[\hat{f}_{Z,\tau}(z)-\hat{f}_Z(z)\right] f_{X_2|Z}(x_2|z) f_{X_2|Z}(\xi_2|z)\Omega_{x_1,0}(z_1)\,dz_1 \\ &\quad =: I_{1,1,2,1}(x_2,\xi_2) + I_{1,1,2,2}(x_2,\xi_2). \end{aligned}$$

We have $E\left[I_{1,1,2,1}(x_2,\xi_2)^{2N}\right] = O\left(a_T^N\right)$, uniformly in $x_2,\xi_2 \in \mathcal{X}_2$, for any $x_1 \in \mathcal{X}_1$, by a similar argument as above. To control $I_{1,1,2,2}(x_2,\xi_2)$, we use that $\left|\hat{f}_{Z,\tau}(z)-\hat{f}_Z(z)\right| \leq Ch_{x_1,T}^{-d_{X_1}} h_T^{-d_{Z_1}} 1\left\{\inf_{z_1\in\mathcal{Z}_1}\hat{f}_Z(z) < \tau_T\right\}$, uniformly in $z_1 \in \mathcal{Z}_1$. Thus, from Assumptions B.3 (i) and B.6 we have $E\left[I_{1,1,2,2}(x_2,\xi_2)^{2N}\right] \leq Ch_{x_1,T}^{-2Nd_{X_1}} h_T^{-2Nd_{Z_1}} P\left[\inf_{z_1\in\mathcal{Z}_1}\hat{f}_Z(z) < \tau_T\right]$, uniformly in $x_2,\xi_2 \in \mathcal{X}_2$, for any $x_1 \in \mathcal{X}_1$. By setting $c := \inf_{z\in\mathcal{Z}} f_Z(z) > 0$, and using a covering argument for compact set $\mathcal{Z}_1$ similar to the proof of Theorem 2.2 in Bosq (1998), we have:

$$\begin{aligned} P\left[\inf_{z_1\in\mathcal{Z}_1}\hat{f}_Z(z) < \tau_T\right] &\leq P\left[\sup_{z_1\in\mathcal{Z}_1}\left|\hat{f}_Z(z)-f_Z(z)\right| \geq c/2\right] \leq P\left[\sup_{z_1\in\mathcal{Z}_1}\left|\hat{f}_Z(z)-E\left[\hat{f}_Z(z)\right]\right| \geq c/4\right] \\ &\leq n_T \sup_{z_1\in\mathcal{Z}_1} P\left[\left|\hat{f}_Z(z)-E\left[\hat{f}_Z(z)\right]\right| \geq c/4\right] + o(1), \end{aligned}$$

where the sequence n_T is such that $n_T = O\left(T^c\right)$ for some $c>0$ and $\frac{1}{h_{x_1,T}^{d_{X_1}} h_T^{d_Z+1} n_T^{1/d_{Z_1}}} = o\left((\log T)^{-1}\right)$. By using a large deviation approach based on Bernstein's inequality [e.g., Bosq (1998), Theorem 1.2 (2)], we can show that $P\left[\left|\hat{f}_Z(z)-E\left[\hat{f}_Z(z)\right]\right| \geq c/4\right] \leq 2\exp\left(-c_1 T h_{x_1,T}^{d_{X_1}} h_T^{d_Z}\right)$, uniformly in $z_1 \in \mathcal{Z}_1$, for a constant $c_1 > 0$. It follows $P\left[\inf_{z_1\in\mathcal{Z}_1}\hat{f}_Z(z) < \tau_T\right] = O\left(T^{-b}\right)$, for any $b>0$. Thus, we get $E\left[I_{1,1,2}(x_2,\xi_2)^{2N}\right] = O\left(a_T^N\right)$, uniformly in $x_2,\xi_2 \in \mathcal{X}_2$, for any $x_1 \in \mathcal{X}_1$.

Let us now consider $I_{1,1,3}(x_2,\xi_2)$. We have:

$$\begin{aligned} E\left[I_{1,1,3}(x_2,\xi_2)^{2N}\right] &= \frac{1}{f_{X_1}(x_1)^{2N}}\int\cdots\int \prod_l f_{X_2|Z}(\xi_2|z_{1,l},x_1)\Omega_{x_1,0}(z_{1,l}) \\ &\quad \cdot E\left[\prod_l \Delta\hat{f}_{X_2,Z}(x_2,z_{1,l},x_1)\frac{\Delta\hat{f}_{Z,\tau}(z_{1,l},x_1)}{\hat{f}_{Z,\tau}(z_{1,l},x_1)}\right]\prod_l dz_{1,l} \qquad \text{(TR.26)} \end{aligned}$$

Now we use $\hat{f}_{Z,\tau}(z) \geq \tau_T$ and the Cauchy-Schwarz inequality to get

$$\begin{aligned}
&\left| E\left[\prod_l \Delta\hat{f}_{X_2,Z}(x_2,z_{1,l},x_1)\frac{\Delta\hat{f}_{Z,\tau}(z_{1,l},x_1)}{\hat{f}_{Z,\tau}(z_{1,l},x_1)}\right]\right| \\
\leq\;& \tau_T^{-2N} E\left[\prod_l \Delta\hat{f}_{X_2,Z}(x_2,z_{1,l},x_1)^2\right]^{1/2} E\left[\prod_l \Delta\hat{f}_{Z,\tau}(z_{1,l},x_1)^2\right]^{1/2} \\
\leq\;& \tau_T^{-2N} \prod_l E\left[\Delta\hat{f}_{X_2,Z}(x_2,z_{1,l},x_1)^{2M}\right]^{1/(2M)} E\left[\Delta\hat{f}_{Z,\tau}(z_{1,l},x_1)^{2M}\right]^{1/(2M)},
\end{aligned}$$

where $M = 2^{2N-1}$. Moreover, from $\left|\hat{f}_{Z,\tau}(z) - \hat{f}_Z(z)\right| \leq C h_{x_1,T}^{-d_{X_1}} h_T^{-d_{Z_1}} 1\left\{\inf_{z_1\in\mathcal{Z}_1} \hat{f}_Z(z) < \tau_T\right\}$, uniformly in $z_1 \in \mathcal{Z}_1$, $P\left[\inf_{z_1\in\mathcal{Z}_1} \hat{f}_Z(z) < \tau_T\right] = O\left(T^{-b}\right)$, for any $b>0$, we get:

$$\begin{aligned}
&\left| E\left[\prod_l \Delta\hat{f}_{X_2,Z}(x_2,z_{1,l},x_1)\frac{\Delta\hat{f}_{Z,\tau}(z_{1,l},x_1)}{\hat{f}_{Z,\tau}(z_{1,l},x_1)}\right]\right| \\
\leq\;& \tau_T^{-2N} \prod_l E\left[\Delta\hat{f}_{X_2,Z}(x_2,z_{1,l},x_1)^{2M}\right]^{1/(2M)} E\left[\Delta\hat{f}_{Z}(z_{1,l},x_1)^{2M}\right]^{1/(2M)} \\
&+ C T^{-2Nb} \tau_T^{-2N} h_{x_1,T}^{-2Nd_{X_1}} h_T^{-2Nd_{Z_1}} \prod_l E\left[\Delta\hat{f}_{X_2,Z}(x_2,z_{1,l},x_1)^{2M}\right]^{1/(2M)} . \quad \text{(TR.27)}
\end{aligned}$$

Thus, from (TR.26), (TR.27) and Assumptions B.3 (i), B,6 (ii), we get:

$$\begin{aligned}
&E\left[I_{1,1,3}(x_2,\xi_2)^{2N}\right] \\
=\;& O\left(\tau_T^{-2N} \sup_{x_2\in\mathcal{X}_2, z_1\in\mathcal{Z}_1} E\left[\Delta\hat{f}_{X_2,Z}(x_2,z_1,x_1)^{2M}\right]^{N/M} \sup_{z_1\in\mathcal{Z}_1} E\left[\left|\Delta\hat{f}_Z(z_1,x_1)\right|^{2M}\right]^{N/M}\right) \\
&+ O\left(T^{-2Nb}\tau_T^{-2N} h_{x_1,T}^{-2Nd_{X_1}} h_T^{-2Nd_{Z_1}} \sup_{x_2\in\mathcal{X}_2, z_1\in\mathcal{Z}_1} E\left[\Delta\hat{f}_{X_2,Z}(x_2,z_1,x_1)^{2M}\right]^{N/M}\right).
\end{aligned}$$

By standard kernel arguments and Assumptions B.1-B.3, we have $\sup_{x_2\in\mathcal{X}_2, z_1\in\mathcal{Z}_1} E\left[\Delta\hat{f}_{X_2,Z}(x_2,z_1,x_1)^{2M}\right]^{1/M} = O\left(\frac{1}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{X_2}+d_{Z_1}}} + h_T^{2m} + h_{x_1,T}^{2m}\right)$ and $\sup_{z_1\in\mathcal{Z}_1} E\left[\left|\Delta\hat{f}_Z(z_1,x_1)\right|^{2M}\right]^{1/M} = O\left(\frac{1}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}}} + h_T^{2m} + h_{x_1,T}^{2m}\right)$, for any $x_1 \in \mathcal{X}_1$. Then, from condition $\frac{(\log T)^2}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}+d_{X_2}}} = O(1)$, we get:

$$E\left[I_{1,1,3}(x_2,\xi_2)^{2N}\right] = O\left(\left(\frac{1}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}}} + h_T^{2m} + h_{x_1,T}^{2m}\right)^N\right) = O\left(a_T^N\right),$$

uniformly in $x_2, \xi_2 \in \mathcal{X}_2$, for any $x_1 \in \mathcal{X}_1$. The bound for term $I_{1,1,4}(x_2, \xi_2)$ is similar, and we conclude that $E\left[I_{1,1}(x_2,\xi_2)^{2N}\right] = O\left(a_T^N\right)$, uniformly in $x_2, \xi_2 \in \mathcal{X}_2$, for any $x_1 \in \mathcal{X}_1$.

(ii) Bound for $E\left[I_{1,3}(x_2,\xi_2)^{2N}\right]$. Write:

$$
\begin{aligned}
E\left[I_{1,3}(x_2,\xi_2)^{2N}\right] &= \int \cdots \int \prod_l f_{X_2|Z}(x_2|z_{1,l}, x_1) f_{X_2|Z}(\xi_2|z_{1,l}, x_1) f_{Z_1|X_1}(z_{1,l}|x_1) \\
&\cdot E\left[\prod_l \Delta\hat{\Omega}_{x_1}(z_{1,l})\right] \prod_l dz_{1,l}.
\end{aligned}
$$

From Assumption B.7 (iii), $\sup_{z_{1,1},\ldots,z_{1,2N} \in \mathcal{Z}_1} \left| E\left[\prod_l \Delta\hat{\Omega}_{x_1}(z_{1,l})\right]\right| \leq \sup_{z_1 \in \mathcal{Z}_1} E\left[\left|\Delta\hat{\Omega}_{x_1}(z_1)\right|^{2N}\right] = O\left(\left(\frac{1}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}}} + h_T^{2m} + h_{x_1,T}^{2m}\right)^N\right)$, for any $x_1 \in \mathcal{X}_1$. Thus, from Assumptions B.3 (i), we get $E\left[I_{1,3}(x_2,\xi_2)^{2N}\right] = O\left(a_T^N\right)$, uniformly in $x_2, \xi_2 \in \mathcal{X}_2$, for any $x_1 \in \mathcal{X}_1$.

(iii) Bound of $E\left[I_{1,2}(x_2,\xi_2)^{2N}\right]$, $E\left[I_{1,4}(x_2,\xi_2)^{2N}\right]$ and $E\left[I_2(x_2,\xi_2)^{2N}\right]$. The bounds for these terms are derived by similar arguments as in (i) and (ii).

9.5 Proof of Lemma B.5

From the proof of Lemma A.6, we have $\left(\lambda_{x_1,T} + A_{x_1}^* A_{x_1}\right)^{-1} A_{x_1}^* \hat{\psi}_{x_1} = \sum_{j=1}^{\infty} c_{x_1,j,T} Z_{x_1,j,T} \phi_{x_1,j}$, where $c_{x_1,j,T} := \frac{1}{\sqrt{T h_{x_1,T}^{d_{X_1}}}} \frac{\sqrt{\nu_{x_1,j}}}{\lambda_{x_1,T} + \nu_{x_1,j}}$ and $Z_{x_1,j,T}$ are defined in (TR.12). For expository purpose, let us omit the index x_1 in $c_{x_1,j,T}$, $Z_{x_1,j,T}$ and $\phi_{x_1,j}$. By using $\left\|\left(\lambda_{x_1,T} + A_{x_1}^* A_{x_1}\right)^{-1} A_{x_1}^* \hat{\psi}_{x_1}\right\|^2 = \sum_{j,l=1}^{\infty} c_{j,T} c_{l,T} Z_{j,T} Z_{l,T} \left\langle \phi_j, \phi_l \right\rangle_{L^2(\mathcal{X}_2)}$, we get $E\left[\left\|\left(\lambda_{x_1,T} + A_{x_1}^* A_{x_1}\right)^{-1} A_{x_1}^* \hat{\psi}_{x_1}\right\|^4\right]$ $= \sum_{j,l,m,n=1}^{\infty} c_{j,T} c_{l,T} c_{m,T} c_{n,T} E\left[Z_{j,T} Z_{l,T} Z_{m,T} Z_{n,T}\right] \left\langle \phi_j, \phi_l \right\rangle_{L^2(\mathcal{X}_2)} \left\langle \phi_m, \phi_n \right\rangle_{L^2(\mathcal{X}_2)}$. Let us now bound the expectation terms. By applying twice the Cauchy-Schwarz inequality, $\left|E\left[Z_{j,T} Z_{l,T} Z_{m,T} Z_{n,T}\right]\right| \leq E\left[Z_{j,T}^2 Z_{l,T}^2\right]^{1/2} E\left[Z_{m,T}^2 Z_{n,T}^2\right]^{1/2} \leq \sup_{j \in \mathbb{N}} E\left[Z_{j,T}^4\right]$. By similar arguments as in Lemma B.10, $\bar{C} := \sup_{j\in\mathbb{N}} E\left[Z_{j,T}^4\right] < \infty$. Then, we get

$$E\left[\left\|\left(\lambda_{x_1,T} + A_{x_1}^* A_{x_1}\right)^{-1} A_{x_1}^* \hat{\psi}_{x_1}\right\|^4\right] \le \bar{C} \sum_{j,l,m,n=1}^{\infty} c_{j,T} c_{l,T} c_{m,T} c_{n,T} \left|\langle \phi_j, \phi_l \rangle_{L^2(\mathcal{X}_2)}\right| \left|\langle \phi_m, \phi_n \rangle_{L^2(\mathcal{X}_2)}\right|$$

$$\begin{aligned}
= \quad & \bar{C}\Bigg\{\sum_{j\in\mathbb{N}} c_{j,T}^4 \left\|\phi_j\right\|_{L^2(\mathcal{X}_2)}^4 + 4 \sum_{(j,l)\in\mathbb{D}^2} c_{j,T}^3 c_{l,T} \left\|\phi_j\right\|_{L^2(\mathcal{X}_2)}^2 \left|\langle \phi_j, \phi_l \rangle_{L^2(\mathcal{X}_2)}\right| \\
& + \sum_{(j,l)\in\mathbb{D}^2} c_{j,T}^2 c_{l,T}^2 \left(\left\|\phi_j\right\|_{L^2(\mathcal{X}_2)}^2 \left\|\phi_l\right\|_{L^2(\mathcal{X}_2)}^2 + 2\left|\langle \phi_j, \phi_l \rangle_{L^2(\mathcal{X}_2)}\right|^2\right) \\
& + 2 \sum_{(j,l,m)\in\mathbb{D}^3} c_{j,T}^2 c_{l,T} c_{m,T} \left\|\phi_j\right\|_{L^2(\mathcal{X}_2)}^2 \left|\langle \phi_l, \phi_m \rangle_{L^2(\mathcal{X}_2)}\right| \\
& + 4 \sum_{(j,l,m)\in\mathbb{D}^3} c_{j,T}^2 c_{l,T} c_{m,T} \left|\langle \phi_j, \phi_l \rangle_{L^2(\mathcal{X}_2)}\right| \left|\langle \phi_j, \phi_m \rangle_{L^2(\mathcal{X}_2)}\right| \\
& + \sum_{(j,l,m,n)\in\mathbb{D}^4} c_{j,T} c_{l,T} c_{m,T} c_{n,T} \left|\langle \phi_j, \phi_l \rangle_{L^2(\mathcal{X}_2)}\right| \left|\langle \phi_m, \phi_n \rangle_{L^2(\mathcal{X}_2)}\right| \Bigg\} \\
=: \quad & \bar{C}\left\{J_1 + 4J_2 + J_3 + 2J_4 + 4J_5 + J_6\right\},
\end{aligned}$$

where $\mathbb{D}^d$ denotes the set of d-tuples, which consist of d different natural numbers. Let us now bound separately the different terms.

(i) Bound of J_1. We have

$$J_1 \le \sum_{j,l=1}^{\infty} c_{j,T}^2 c_{l,T}^2 \left\|\phi_j\right\|_{L^2(\mathcal{X}_2)}^2 \left\|\phi_l\right\|_{L^2(\mathcal{X}_2)}^2 = \left(\sum_{j=1}^{\infty} c_{j,T}^2 \left\|\phi_j\right\|_{L^2(\mathcal{X}_2)}^2\right)^2 = q_{x_1,T}^4$$

where we denote $q_{x_1,T}^2 := \sum_{j=1}^{\infty} c_{j,T}^2 \left\|\phi_j\right\|_{L^2(\mathcal{X}_2)}^2 = \frac{1}{Th_{x_1,T}^{d_{X_1}}} \sum_{j=1}^{\infty} \frac{\nu_{x_1,j}}{\left(\lambda_{x_1,T} + \nu_{x_1,j}\right)^2} \left\|\phi_{x_1,j}\right\|_{L^2(\mathcal{X}_2)}^2$.

(ii) Bound of J_2. Using $c_{j,T}^2 \left\|\phi_j\right\|_{L^2(\mathcal{X}_2)}^2 \le q_{x_1,T}^2$, $j \in \mathbb{N}$, we get $J_2 \le q_{x_1,T}^2 \sum_{(j,l)\in\mathbb{D}^2} c_{j,T} c_{l,T} \left|\langle \phi_j, \phi_l \rangle_{L^2(\mathcal{X}_2)}\right|$. Let us consider the term

$$\sum_{(j,l)\in\mathbb{D}^2} c_{j,T} c_{l,T} \left|\langle \phi_j, \phi_l \rangle_{L^2(\mathcal{X}_2)}\right| = \sum_{j=1}^{\infty} c_{j,T} \left\|\phi_j\right\|_{L^2(\mathcal{X}_2)} \left(\sum_{l:l\neq j} c_{l,T} \left\|\phi_l\right\|_{L^2(\mathcal{X}_2)} \frac{\left|\langle \phi_j, \phi_l \rangle_{L^2(\mathcal{X}_2)}\right|}{\left\|\phi_j\right\|_{L^2(\mathcal{X}_2)} \left\|\phi_l\right\|_{L^2(\mathcal{X}_2)}}\right).$$

Using Cauchy-Schwarz inequality, for any j we have $\sum_{l:l\neq j} c_{l,T} \left\|\phi_l\right\|_{L^2(\mathcal{X}_2)} \frac{\left|\langle \phi_j, \phi_l \rangle_{L^2(\mathcal{X}_2)}\right|}{\left\|\phi_j\right\|_{L^2(\mathcal{X}_2)} \left\|\phi_l\right\|_{L^2(\mathcal{X}_2)}} \le$

$$\left(\sum_{l:l\neq j} c_{l,T}^2 \left\|\phi_l\right\|_{L^2(\mathcal{X}_2)}^2\right)^{1/2} \left(\sum_{l:l\neq j} \frac{\langle \phi_j, \phi_l \rangle_{L^2(\mathcal{X}_2)}^2}{\left\|\phi_j\right\|_{L^2(\mathcal{X}_2)}^2 \left\|\phi_l\right\|_{L^2(\mathcal{X}_2)}^2}\right)^{1/2} \le q_{x_1,T} \left(\sum_{l:l\neq j} \frac{\langle \phi_j, \phi_l \rangle_{L^2(\mathcal{X}_2)}^2}{\left\|\phi_j\right\|_{L^2(\mathcal{X}_2)}^2 \left\|\phi_l\right\|_{L^2(\mathcal{X}_2)}^2}\right)^{1/2}.$$

Thus, we get again by Cauchy-Schwarz inequality,

$$\begin{aligned}
&\sum_{(j,l)\in\mathbb{D}^2} c_{j,T}c_{l,T}\left|\left\langle \phi_j,\phi_l\right\rangle_{L^2(\mathcal{X}_2)}\right| \\
\leq\ & q_{x_1,T}\sum_{j=1}^{\infty} c_{j,T}\left\|\phi_j\right\|_{L^2(\mathcal{X}_2)}\left(\sum_{l:l\neq j}\frac{\left\langle \phi_j,\phi_l\right\rangle^2_{L^2(\mathcal{X}_2)}}{\left\|\phi_j\right\|^2_{L^2(\mathcal{X}_2)}\left\|\phi_l\right\|^2_{L^2(\mathcal{X}_2)}}\right)^{1/2} \\
\leq\ & q_{x_1,T}\left(\sum_{j=1}^{\infty} c^2_{j,T}\left\|\phi_j\right\|^2_{L^2(\mathcal{X}_2)}\right)^{1/2}\left(\sum_{j=1}^{\infty}\sum_{l:l\neq j}\frac{\left\langle \phi_j,\phi_l\right\rangle^2_{L^2(\mathcal{X}_2)}}{\left\|\phi_j\right\|^2_{L^2(\mathcal{X}_2)}\left\|\phi_l\right\|^2_{L^2(\mathcal{X}_2)}}\right)^{1/2} = q^2_{x_1,T}\bar{\rho},
\end{aligned} \tag{TR.28}$$

where $\bar{\rho} := \left(\sum_{j,l=1:l\neq j}^{\infty}\frac{\left\langle \phi_j,\phi_l\right\rangle^2_{L^2(\mathcal{X}_2)}}{\left\|\phi_j\right\|^2_{L^2(\mathcal{X}_2)}\left\|\phi_l\right\|^2_{L^2(\mathcal{X}_2)}}\right)^{1/2} < \infty$ by Assumption B.8 (i). We deduce $J_2 \leq q^4_{x_1,T}\bar{\rho}$.

(iii) Bound of J_3. We have $J_3 \leq 3\sum_{(j,l)\in\mathbb{D}^2} c^2_{j,T}c^2_{l,T}\left\|\phi_j\right\|^2_{L^2(\mathcal{X}_2)}\left\|\phi_l\right\|^2_{L^2(\mathcal{X}_2)} \leq 3q^4_{x_1,T}$.

(iv) Bound of J_4. We have $J_4 \leq \left(\sum_j c^2_{j,T}\left\|\phi_j\right\|^2_{L^2(\mathcal{X}_2)}\right)\sum_{(l,m)\in\mathbb{D}^2} c_{l,T}c_{m,T}\left|\langle\phi_l,\phi_m\rangle_{L^2(\mathcal{X}_2)}\right| \leq q^4_{x_1,T}\bar{\rho}$, using (TR.28).

(v) Bound of J_5. We have

$$\begin{aligned}
J_5 \leq\ & \sum_j c^2_{j,T}\left\|\phi_j\right\|^2_{L^2(\mathcal{X}_2)} \\
&\cdot\left(\sum_{l:l\neq j}\sum_{m:m\neq j} c_{l,T}\left\|\phi_l\right\|_{L^2(\mathcal{X}_2)}\frac{\left|\left\langle \phi_j,\phi_l\right\rangle_{L^2(\mathcal{X}_2)}\right|}{\left\|\phi_j\right\|_{L^2(\mathcal{X}_2)}\left\|\phi_l\right\|_{L^2(\mathcal{X}_2)}} c_{m,T}\left\|\phi_m\right\|_{L^2(\mathcal{X}_2)}\frac{\left|\left\langle \phi_j,\phi_m\right\rangle_{L^2(\mathcal{X}_2)}\right|}{\left\|\phi_j\right\|_{L^2(\mathcal{X}_2)}\left\|\phi_m\right\|_{L^2(\mathcal{X}_2)}}\right) \\
=\ & \sum_j c^2_{j,T}\left\|\phi_j\right\|^2_{L^2(\mathcal{X}_2)}\left(\sum_{l:l\neq j} c_{l,T}\left\|\phi_l\right\|_{L^2(\mathcal{X}_2)}\frac{\left|\left\langle \phi_j,\phi_l\right\rangle_{L^2(\mathcal{X}_2)}\right|}{\left\|\phi_j\right\|_{L^2(\mathcal{X}_2)}\left\|\phi_l\right\|_{L^2(\mathcal{X}_2)}}\right)^2.
\end{aligned}$$

Using $c^2_{j,T}\left\|\phi_j\right\|^2_{L^2(\mathcal{X}_2)} \leq q^2_{x_1,T}$ for all $j\in\mathbb{N}$, and $\sum_{l:l\neq j} c_{l,T}\left\|\phi_l\right\|_{L^2(\mathcal{X}_2)}\frac{\left|\left\langle \phi_j,\phi_l\right\rangle_{L^2(\mathcal{X}_2)}\right|}{\left\|\phi_j\right\|_{L^2(\mathcal{X}_2)}\left\|\phi_l\right\|_{L^2(\mathcal{X}_2)}} \leq q_{x_1,T}\left(\sum_{l:l\neq j}\frac{\left\langle \phi_j,\phi_l\right\rangle^2_{L^2(\mathcal{X}_2)}}{\left\|\phi_j\right\|^2_{L^2(\mathcal{X}_2)}\left\|\phi_l\right\|^2_{L^2(\mathcal{X}_2)}}\right)^{1/2}$, we get $J_5 \leq q^4_{x_1,T}\sum_j\sum_{l:l\neq j}\frac{\left\langle \phi_j,\phi_l\right\rangle^2_{L^2(\mathcal{X}_2)}}{\left\|\phi_j\right\|^2_{L^2(\mathcal{X}_2)}\left\|\phi_l\right\|^2_{L^2(\mathcal{X}_2)}} = q^4_{x_1,T}\bar{\rho}^2$.

(vi) Bound of J_6. Finally, $J_6 \leq \sum_{(j,l)\in\mathbb{D}^2,(m,n)\in\mathbb{D}^2} c_{j,T}c_{l,T}c_{m,T}c_{n,T}\left|\left\langle \phi_j,\phi_l\right\rangle_{L^2(\mathcal{X}_2)}\right|\left|\langle\phi_m,\phi_n\rangle_{L^2(\mathcal{X}_2)}\right| = \left(\sum_{(j,l)\in\mathbb{D}^2} c_{j,T}c_{l,T}\left|\left\langle \phi_j,\phi_l\right\rangle_{L^2(\mathcal{X}_2)}\right|\right)^2 \leq q^4_{x_1,T}\bar{\rho}^2$, using (TR.28).

To summarize, we have proved $E\left[\left\|\left(\lambda_{x_1,T}+A^*_{x_1}A_{x_1}\right)^{-1}A^*_{x_1}\hat{\psi}_{x_1}\right\|^4\right]\leq C^*q^4_{x_1,T}$, for a constant $C^*<\infty$. Since in Lemma A.6 we prove that $E\left[\left\|\left(\lambda_{x_1,T}+A^*_{x_1}A_{x_1}\right)^{-1}A^*_{x_1}\hat{\psi}_{x_1}\right\|^2\right]=q^2_{x_1,T}\left(1+o(1)\right)$, the conclusion follows.

9.6 Proof of Lemma B.6

By (TR.25) we have:

$$\begin{aligned}&E\left[\left\|\left(\widetilde{\hat{A}}_{x_1}-\tilde{A}_{x_1}\right)\left(\hat{\psi}_{x_1}+\zeta_{x_1}\right)\right\|^2_{L^2(\mathcal{X}_2)}\right]^{1/2}\\ \leq\ &E\left[\|I_{1,1}\|^2_{L^2(\mathcal{X}_2)}\right]^{1/2}+E\left[\|I_{1,2}\|^2_{L^2(\mathcal{X}_2)}\right]^{1/2}+E\left[\|I_{1,3}\|^2_{L^2(\mathcal{X}_2)}\right]^{1/2}+E\left[\|I_2\|^2_{L^2(\mathcal{X}_2)}\right]^{1/2}.\end{aligned}$$

We bound separately the four terms.

Let us first bound $E\left[\|I_{1,1}\|^2_{L^2(\mathcal{X}_2)}\right]$. We have

$$\begin{aligned}E\left[\|I_{1,1}\|^2_{L^2(\mathcal{X}_2)}\right]&=\frac{1}{f_X\left(x_1\right)^2}\int\int\int f_{X_2|Z}(x_2|z_1,x_1)f_{X_2|Z}(x_2|\tilde{z}_1,x_1)\\&\quad\cdot E\left[\Delta\hat{\Omega}_{x_1}\left(z_1\right)\Delta\hat{\Omega}_{x_1}\left(\tilde{z}_1\right)\hat{\Psi}_{x_1}\left(z_1\right)\hat{\Psi}_{x_1}\left(\tilde{z}_1\right)\right]dz_1d\tilde{z}_1dx_2,\end{aligned}$$

where $\hat{\Psi}_{x_1}\left(z_1\right)$ is defined in (TR.9). Thus, by Assumption B.3 (i) and the Cauchy-Schwarz inequality, we get $E\left[\|I_{1,1}\|^2_{L^2(\mathcal{X}_2)}\right]^{1/2}=O\left(\sup_{z_1\in\mathcal{Z}_1}E\left[\left|\Delta\hat{\Omega}_{x_1}\left(z_1\right)\right|^4\right]^{1/4}\sup_{z_1\in\mathcal{Z}_1}E\left[\left|\hat{\Psi}_{x_1}\left(z_1\right)\right|^4\right]^{1/4}\right)$.

LEMMA C.2: *Let Assumptions B.1-B.3 hold.* (1) *Uniformly in* $x_1\in\mathcal{X}_1$*:*

$$\sup_{z_1\in\mathcal{Z}}V\left[\hat{\Psi}_{x_1}(z_1)\right]=O\left(\frac{1}{Th^{d_{X_1}}_{x_1,T}h^{d_{Z_1}}_T}\right).$$

(2) *If in addition Assumption B.5 holds, then uniformly in* $x_1\in\mathcal{X}_1$*:*

$$\sup_{z_1\in\mathcal{Z}}E\left[\left(\hat{\Psi}_{x_1}(z_1)-E\left[\hat{\Psi}_{x_1}(z_1)\right]\right)^4\right]=O\left(\frac{1}{T^3h^{3d_{X_1}}_{x_1,T}h^{3d_{Z_1}}_T}+\frac{1}{T^2h^{2d_{X_1}}_{x_1,T}h^{2d_{Z_1}}_T}\right).$$

From Lemmas C.1 (1) and C.2 (2), Assumptions B.1-B.3, B.5 and B.7 (iii) and by using the condition $\dfrac{1}{Th^{d_{X_1}}_{x_1,T}h^{d_{Z_1}}_T}=O(1)$, we get:

$$E\left[\|I_{1,1}\|^2_{L^2(\mathcal{X}_2)}\right]^{1/2}=O\left(\frac{1}{Th^{d_{X_1}}_{x_1}h^{d_{Z_1}}_T}+h^{2m}_{x_1,T}+h^{2m}_T\right).$$

By similar arguments as above and as in the proof of Lemma B.2, we get:

$$E\left[\|I_{1,2}\|^2_{L^2(\mathcal{X}_2)}\right]^{1/2} = O\left(\frac{1}{Th^{d_{X_1}}_{x_1,T}h_T^{d_{Z_1}+d_{X_2}}} + h^{2m}_{x_1,T} + h_T^{2m}\right),$$
$$E\left[\|I_{1,3}\|^2_{L^2(\mathcal{X}_2)}\right]^{1/2} = O\left(\frac{1}{Th^{d_{X_1}}_{x_1}} + h^{2m}_{x_1,T} + h_T^{2m}\right).$$

Finally, by arguments similar as in the proof of Lemma B.4 to control the kernel density estimator in the denominator, we have:

$$E\left[\|I_2\|^2_{L^2(\mathcal{X}_2)}\right]^{1/2} = o\left(\frac{1}{Th^{d_{X_1}}_{x_1,T}h_T^{d_{Z_1}+d_{X_2}}} + h^{2m}_{x_1,T} + h_T^{2m}\right).$$

9.7 Proof of Lemma B.7

The proof is similar to that of Lemmas B.3, B.4 and B.6 and is therefore omitted.

9.8 Proof of Lemma B.8

We have

$$\Delta g_h(x) = \int k(z)\left[g(x-hz) - g(x)\right] 1_{\mathcal{S}}(x-hz)dz. \tag{TR.29}$$

For $x, x-hz \in \mathcal{S}$, write $g(x-hz) - g(x) = -\int_0^h \left(\nabla g\,(x-tz)\cdot z\right) dt$. Thus, we get

$$\begin{aligned} |g(x-hz)-g(x)| &\leq \int_0^h |\nabla g\,(x-tz)|\,|z|\,dt \\ &\leq |z|\left(\int_0^h |\nabla g\,(x-tz)|^2\,dt\right)^{1/2}\left(\int_0^h dt\right)^{1/2} \\ &= |z|\sqrt{h}\left(\int_0^h |\nabla g\,(x-tz)|^2\,dt\right)^{1/2}. \end{aligned}$$

We deduce from (TR.29) and the Cauchy-Schwarz inequality:

$$\begin{aligned} |\Delta g_h(x)| &\leq \sqrt{h}\int |k(z)|\,|z|\left(\int_0^h |\nabla g\,(x-tz)|^2\,dt\right)^{1/2} 1_{\mathcal{S}}(x-hz)dz \\ &\leq \sqrt{h}\left(\int |k(z)|\,|z|^2\,dz\right)^{1/2}\left(\int |k(z)|\left(\int_0^h |\nabla g\,(x-tz)|^2\,dt\right) 1_{\mathcal{S}}(x-hz)dz\right)^{1/2}. \end{aligned}$$

Thus, since $1_{\mathcal{S}}(x-hz) \leq 1_{\mathcal{S}}(x-tz)$ for any $x \in \mathcal{S}$ and $0 \leq t \leq h$ by convexity of $\mathcal{S}$, we get $\int |\Delta g_h(x)|^2 f(x)dx \leq hw_2 \int_0^h \int |k(z)| \left(\int |\nabla g\,(x-tz)|^2 1_{\mathcal{S}}(x-tz)f(x)dx\right) dzdt$. Now, $\int |\nabla g\,(x-tz)|^2 1_{\mathcal{S}}(x-tz)f(x)dx = \int |\nabla g\,(y)|^2 1_{\mathcal{S}}(y)f(y+tz)dy$

$\leq \int |\nabla g(y)|^2 f(y)dy + \int |\nabla g(y)|^2 1_{\mathcal{S}}(y) |f(y+tz) - f(y)| dy$. Then, $\int |\Delta g_h(x)|^2 f(x)dx$
$\leq h^2 w_1 w_2 \int |\nabla g(y)|^2 f(y)dy + hw_2 \int_0^h \int |k(z)| \int |\nabla g(y)|^2 1_{\mathcal{S}}(y) |f(y+tz) - f(y)| dydzdt.$

9.9 Proof of Lemma B.9

Let $x = (x_1, x_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ and $y = (y_1, y_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, with $y_1 = x_1$. Then we have

$$E[V_T(g, x_1)V_T(e, x_1)] = Th_{1T}^{d_1} \int \int g(x_2)e(y_2)E\left[\left(\hat{f}(x) - E\hat{f}(x)\right)\left(\hat{f}(y) - E\hat{f}(y)\right)\right] dx_2 dy_2.$$

Let us compute the cross-moment of kernel estimator in the RHS:

$$\begin{aligned} & E\left[\left(\hat{f}(x) - E\hat{f}(x)\right)\left(\hat{f}(y) - E\hat{f}(y)\right)\right] \\ = & \frac{1}{Th_{1T}^{2d_1} h_{2T}^{2d_2}} \left\{ E\left[K\left(\frac{U-x}{h_T}\right) K\left(\frac{U-y}{h_T}\right)\right] - E\left[K\left(\frac{U-x}{h_T}\right)\right] E\left[K\left(\frac{U-y}{h_T}\right)\right]\right\}, \end{aligned}$$

where $K\left(\frac{U-x}{h_T}\right) := K_1\left(\frac{U_1-x_1}{h_{1T}}\right) K_2\left(\frac{U_2-x_2}{h_{2T}}\right)$. Thus, we get

$$\begin{aligned} & E[V_T(g, x_1)V_T(e, x_1)] \\ = & \frac{1}{h_{1T}^{d_1} h_{2T}^{2d_2}} \int \int g(x_2)e(y_2)E\left[K\left(\frac{U-x}{h_T}\right) K\left(\frac{U-y}{h_T}\right)\right] dx_2 dy_2 \\ & - \left(\frac{1}{h_{1T}^{d_1/2} h_{2T}^{d_2}} \int g(x_2)E\left[K\left(\frac{U-x}{h_T}\right)\right] dx_2\right) \left(\frac{1}{h_{1T}^{d_1/2} h_{2T}^{d_2}} \int e(x_2)E\left[K\left(\frac{U-x}{h_T}\right)\right] dx_2\right) \\ =: & A_T - B_T(g)B_T(e). \end{aligned} \quad \text{(TR.30)}$$

Let us derive the asymptotic expansions of these two terms.

i) Asymptotic expansion of B_T. Let us first consider the second term in (TR.30). We have

$$\begin{aligned} B_T(g) &= \int \int g(x_2) \frac{1}{h_{1T}^{d_1/2} h_{2T}^{d_2}} K_1\left(\frac{u_1-x_1}{h_{1T}}\right) K_2\left(\frac{u_2-x_2}{h_{2T}}\right) f(u_1, u_2) du_1 du_2 dx_2 \\ &= h_{1T}^{d_1/2} \int \int g(x_2)K(z)f(x + h_T z) dz dx_2 \\ &= h_{1T}^{d_1/2} \int \int g(x_2)K(z)[f(x + h_T z) - f(x)] dz dx_2, \end{aligned}$$

where $f(x + h_T z) := f(x_1 + h_{1T}z_1, x_2 + h_{2T}z_2)$ and $K(z) := K_1(z_1)K_2(z_2)$, since $E[g(U_2)|U_1 = x_1] = 0$. Then $|B_T(g)| \leq h_{1T}^{d_1/2} \int |g(x_2)| \left(\int |K(z)| |f(x + h_T z) - f(x)| dz\right) dx_2$. Since K has a bounded support (Assumption B.2), and by the mean-value Theorem, we get $\int |K(z)| |f(x + h_T z) - f(x)| dz \leq \left(h_{1T} \int |K_1(z_1)| |z_1| dz_1 + h_{2T} \int |K_2(z_2)| |z_2| dz_2\right) q(x)$, for large T, where q is defined in Assumption B.4. Then, by Cauchy-Schwarz inequality and Assumption B.4,

$$|B_T(g)| \leq c h_{1T}^{d_1/2} (h_{1T} + h_{2T}) \int |g(x_2)| \, q(x) dx_2 \leq c h_{1T}^{d_1/2} (h_{1T} + h_{2T}) \left(\int |g(x_2)|^2 q(x) dx_2 \right)^{1/2} \left(\int q(x) dx_2 \right)^{1/2},$$

for a constant c. Thus,

$$B_T(g) = O\left(h_{1T}^{d_1/2} (h_{1T} + h_{2T}) \|g\|_{L^2_{x_1}(F^*)} \right). \tag{TR.31}$$

ii) Asymptotic expansion of A_T. Let us now consider the first term in (TR.30). We have

$$\begin{aligned} \frac{1}{h_{1T}^{d_1} h_{2T}^{2d_2}} E\left[K\left(\frac{U-x}{h_T}\right) K\left(\frac{U-y}{h_T}\right)\right] &= \frac{1}{h_{1T}^{d_1} h_{2T}^{2d_2}} E\left[K_1\left(\frac{U_1-x_1}{h_{1T}}\right)^2 K_2\left(\frac{U_2-x_2}{h_{2T}}\right) K_2\left(\frac{U_2-y_2}{h_{2T}}\right)\right] \\ &= \frac{1}{h_{2T}^{d_2}} \int K_1(u_1)^2 K_2(u_2) K_2\left(u_2 + \frac{x_2-y_2}{h_{2T}}\right) f\left(x + h_T u\right) du. \end{aligned}$$

We get

$$\begin{aligned} A_T &= \frac{\omega^2}{h_{2T}^{d_2}} \int\int\int g(x_2) e(y_2) K_2(u_2) K_2\left(u_2 + \frac{x_2-y_2}{h_{2T}}\right) f(x) \, du_2 dx_2 dy_2 \\ &\quad + \frac{1}{h_{2T}^{d_2}} \int\int\int g(x_2) e(y_2) K_1(u_1)^2 K_2(u_2) K_2\left(u_2 + \frac{x_2-y_2}{h_{2T}}\right) \left[f\left(x + h_T u\right) - f(x)\right] du dx_2 dy_2 \\ &=: \widetilde{A}_{1T} + \widetilde{A}_{2T}. \end{aligned}$$

To rewrite these terms, we have

$$\begin{aligned} & \int \frac{1}{h_{2T}^{d_2}} K_2\left(u_2 + \frac{x_2-y_2}{h_{2T}}\right) e(y_2) dy_2 \\ =& \int K_2(u_2+z_2) \, e(x_2 - h_{2T} z_2) 1_{\mathcal{S}_2}(x_2 - h_{2T} z_2) dz_2 \\ =& \left(\int K_2(u_2+z_2) 1_{\mathcal{S}_2}(x_2 - h_{2T} z_2) dz_2 \right) e(x_2) \\ & + \int K_2(u_2+z_2) \left[e(x_2 - h_{2T} z_2) - e(x_2) \right] 1_{\mathcal{S}_2}(x_2 - h_{2T} z_2) dz_2. \end{aligned}$$

Thus, we get

$$\begin{aligned} & \widetilde{A}_{1T} \\ =& \ \omega^2 \int g(x_2) e(x_2) f(x) \left(\int\int K_2(u_2) K_2(u_2+z_2) 1_{\mathcal{S}_2}(x_2 - h_{2T} z_2) dz_2 du_2 \right) dx_2 \\ & + \omega^2 \int\int g(x_2) f(x) K_2(u_2) \int K_2(u_2+z_2) \left[e(x_2 - h_{2T} z_2) - e(x_2) \right] 1_{\mathcal{S}_2}(x_2 - h_{2T} z_2) dz_2 du_2 dx_2 \\ =& \ \omega^2 f_{U_1}(x_1) Cov\left[g(U_2), e(U_2) | U_1 = x_1 \right] \\ & - \omega^2 \int g(x_2) e(x_2) f(x) \left(\int\int K_2(u_2) K_2(u_2+z_2) 1_{\mathcal{S}_2^c}(x_2 - h_{2T} z_2) dz_2 du_2 \right) dx_2 \\ & + \omega^2 \int\int g(x_2) f(x) K_2(u_2) \left(\int K_2(u_2+z_2) \left[e(x_2 - h_{2T} z_2) - e(x_2) \right] 1_{\mathcal{S}_2}(x_2 - h_{2T} z_2) dz_2 \right) du_2 dx_2. \end{aligned}$$

Similarly

$$\begin{aligned}\widetilde{A}_{2T} &= \int g(x_2)e(x_2)\int\int K_1(u_1)^2K_2\left(u_2\right)K_2\left(u_2+z_2\right)1_{\mathcal{S}_2}(x_2-h_{2T}z_2)\left[f\left(x+h_Tu\right)-f(x)\right]dz_2dudx_2\\ &+\int\int g(x_2)K_1(u_1)^2K_2(u_2)\int K_2\left(u_2+z_2\right)\left[e\left(x_2-h_{2T}z_2\right)-e\left(x_2\right)\right]1_{\mathcal{S}_2}(x_2-h_{2T}z_2)dz_2\\ &\qquad \left[f\left(x+h_Tu\right)-f(x)\right]dudx_2.\end{aligned}$$

We conclude

$$\begin{aligned}& A_T\\ = \;& \omega^2 f_{U_1}(x_1)Cov\left[g(U_2),e(U_2)|U_1=x_1\right]\\ & -\omega^2\int g(x_2)e(x_2)f(x)\left(\int\int K_2\left(u_2\right)K_2\left(u_2+z_2\right)1_{\mathcal{S}_2^c}(x_2-h_{2T}z_2)dz_2du_2\right)dx_2\\ & +\omega^2\int\int g(x_2)f(x)K_2(u_2)\left(\int K_2\left(u_2+z_2\right)\left[e\left(x_2-h_{2T}z_2\right)-e\left(x_2\right)\right]1_{\mathcal{S}_2}(x_2-h_{2T}z_2)dz_2\right)du_2dx_2\\ & +\int g(x_2)e(x_2)\int\int K_1(u_1)^2K_2\left(u_2\right)K_2\left(u_2+z_2\right)1_{\mathcal{S}_2}(x_2-h_{2T}z_2)\left[f\left(x+h_Tu\right)-f(x)\right]dz_2dudx_2\\ & +\int\int g(x_2)K_1(u_1)^2K_2(u_2)\int K_2\left(u_2+z_2\right)\left[e\left(x_2-h_{2T}z_2\right)-e\left(x_2\right)\right]1_{\mathcal{S}_2}(x_2-h_{2T}z_2)dz_2\\ & \cdot\left[f\left(x+h_Tu\right)-f(x)\right]dudx_2.\\ & =:\omega^2 f_{U_1}(x_1)Cov\left[g(U_2),e(U_2)|U_1=x_1\right]+A_{1,T}+A_{2,T}+A_{3,T}+A_{4,T}. \qquad \text{(TR.32)}\end{aligned}$$

Let us now bound terms $A_{1,T}$-$A_{4,T}$, separately.

iii) Bound of $A_{1,T}$. We have

$$|A_{1,T}|\leq\omega^2\int|g(x_2)|\,|e(x_2)|\,f(x)\left(\int\overline{K}_2\left(z_2\right)1_{\mathcal{S}_2}(x_2)1_{\mathcal{S}_2^c}(x_2-h_{2T}z_2)dz_2\right)dx_2,$$

where $\overline{K}_2\left(z_2\right):=\int|K_2(u_2)K_2(u_2+z_2)|\,du_2$. By Assumption B.2, $\overline{K}_2$ has bounded support included in $B_\kappa(0)\subset\mathbb{R}^{d_2}$, $\kappa=2\sup_{z_2\in\,\text{supp}(K_2)}|z_2|$. Then, $\int\overline{K}_2\left(z_2\right)1_{\mathcal{S}_2}(x_2)1_{\mathcal{S}_2^c}(x_2-h_{2T}z_2)dz_2\leq c1\left(x_2\in\partial\mathcal{S}_2(\kappa h_{2T})\right)$, for large T, for a constant c. Thus,

$$|A_{1,T}|\leq c\int|g(x_2)|\,|e(x_2)|\,1\left(x_2\in\partial\mathcal{S}_2(\kappa h_{2T})\right)f(x)dx_2\leq cf_{U_1}(x_1)\left\|g\right\|_{L^2_{x_1}(F)}\rho_{x_1}(e,\kappa h_{2T}), \qquad \text{(TR.33)}$$

by Cauchy-Schwarz inequality.

iv) Bound of $A_{2,T}$. We have $A_{2,T}=\omega^2\int g(x_2)\left(\int k_2\left(z_2\right)\left[e\left(x_2-h_{2T}z_2\right)-e\left(x_2\right)\right]1_{\mathcal{S}_2}(x_2-h_{2T}z_2)dz_2\right)f(x)dx_2=\omega^2 f_{U_1}(x_1)\int g(x_2)\left[\Delta e_{h_{2T}}(x_2)\right]f_{U_2|U_1}\left(x_2|x_1\right)dx_2$, where $\Delta e_{h_{2T}}(x_2):=\int k_2\left(z_2\right)\left[e\left(x_2-h_{2T}z_2\right)-e\left(x_2\right)\right]1_{\mathcal{S}_2}(x_2-h_{2T}z_2)dz_2$ and $k_2(z_2):=\int K_2(u_2)K_2(u_2+z_2)du_2$. Note that $k_2(.)$ satisfies $\int k_2(z_2)dz_2=1$. By the Cauchy-Schwarz inequality, we have

$$|A_{2,T}|\leq\omega^2 f_{U_1}(x_1)\left\|g\right\|_{L^2_{x_1}(F)}\left\|\Delta e_{h_{2T}}\right\|_{L^2_{x_1}(F)}. \qquad \text{(TR.34)}$$

To bound the term $\|\Delta e_{h_{2T}}\|_{L^2_{x_1}(F)}$, we can apply Lemma B.8 to function e and density $f_{x_1}(.) := f_{U_2|U_1}(.|x_1)$ with

$$\begin{aligned} w_1 &= \int |k_2(z_2)|\, dz_2 \leq \int\int |K_2(u_2)K_2(u_2+z_2)|\, du_2 dz_2 = \left(\int |K_2(u_2)|\, du_2\right)^2, \\ w_2 &= \int |z_2|^2 |k_2(z_2)|\, dz_2 \leq \int\int |z_2|^2 |K_2(u_2)K_2(u_2+z_2)|\, du_2 dz_2, \end{aligned}$$

which are finite by Assumptions B.2 (i), (ii). We get

$$\begin{aligned} \|\Delta e_{h_T}\|^2_{L^2_{x_1}(F)} \leq\; & w_1 w_2 h^2_{2T} \|\nabla e\|^2_{L^2_{x_1}(F)} \\ & + w_2 h_{2T} \int |\nabla e(y_2)|^2 \left(\int_0^{h_{2T}} \int |k_2(z_2)|\, |f_{x_1}(y_2+tz_2) - f_{x_1}(y_2)|\, dz_2 dt\right) dy_2. \end{aligned}$$

Since k_2 is bounded and has bounded support, $\int_0^{h_{2T}} \int |k_2(z_2)|\, |f_{x_1}(y_2+tz_2) - f_{x_1}(y_2)|\, dz_2 dt \leq ch^2_{2T} q(x_1, y_2)$, for a constant c, where q is defined in Assumption B.4. Thus, $\|\Delta e_{h_{2T}}\|^2_{L^2_{x_1}(F)} = O\left(h^2_{2T} \|\nabla e\|^2_{L^2_{x_1}(F)} + h^3_{2T} \|\nabla e\|^2_{L^2_{x_1}(F^*)}\right)$. We conclude from (TR.34):

$$A_{2,T} = O\left(h_{2T} \|g\|_{L^2_{x_1}(F)} \|\nabla e\|_{L^2_{x_1}(F)} + h^{3/2}_{2T} \|g\|_{L^2_{x_1}(F)} \|\nabla e\|_{L^2_{x_1}(F^*)}\right). \qquad \text{(TR.35)}$$

v) Bound of $A_{3,T}$. We have

$$|A_{3,T}| \leq \left(\int |K_2(z_2)|\, dz_2\right) \int |g(x_2)|\, |e(x_2)| \left(\int K_1(u_1)^2 |K_2(u_2)|\, |f(x+h_T u) - f(x)|\, du\right) dx_2.$$

Again, by Assumptions B.2 (i), (ii) and B.4, $\int K_1(u_1)^2 |K_2(u_2)|\, |f(x+h_T u) - f(x)|\, du \leq c(h_{1T} + h_{2T})\, q(x)$, for a constant c and large T. Thus,

$$\begin{aligned} |A_{3,T}| &\leq \tilde{c}\,(h_{1T} + h_{2T}) \int |g(x_2)|\, |e(x_2)|\, q(x) dx_2 \\ &\leq \tilde{c}\,(h_{1T} + h_{2T}) \left(\int g(x_2)^2 q(x) dx_2\right)^{1/2} \left(\int e(x_2)^2 q(x) dx_2\right)^{1/2} \end{aligned}$$

from Cauchy-Schwarz inequality, and

$$A_{3,T} = O\left((h_{1T} + h_{2T}) \|g\|_{L^2_{x_1}(F^*)} \|e\|_{L^2_{x_1}(F^*)}\right). \qquad \text{(TR.36)}$$

vi) Bound of $A_{4,T}$. We have

$$A_{4,T} = \int K_1(u_1)^2 K_2(u_2) \left(\int g(x_2)\, [\Delta e_{u_2,h_{2T}}(x_2)]\, [f(x+h_T u) - f(x)]\, dx_2\right) du,$$

where $\Delta e_{u_2,h_{2T}}(x_2) = \int k_2(z_2;u_2)\left[e\left(x_2 - h_{2T}z_2\right) - e(x_2)\right] 1_{\mathcal{S}_2}(x_2 - h_{2T}z_2)dz_2$ and $k_2(z_2;u_2) = K_2(u_2+z_2)$. Since K is bounded and has bounded support,

$$\begin{aligned}
|A_{4,T}| &\leq c\int |K(u)| \left(\int |g(x_2)|\, |\Delta e_{h_{2T}}(x_2)|\, |f(x+h_T u) - f(x)|\, dx_2\right) du \\
&= c\int |g(x_2)|\, |\Delta e_{h_{2T}}(x_2)|\, \pi_T(x) dx_2
\end{aligned}$$

where $|\Delta e_{h_{2T}}(x_2)| := \int 1(|z_2| < c_1)\, |e\left(x_2 - h_{2T}z_2 x\right) - e(x_2)|\, 1_{\mathcal{S}_2}(x_2 - h_{2T}z_2)dz_2$ and $\pi_T(x) := \int |K(u)|\, |f\left(x + h_T u\right) - f(x)|\, du$, for constant c, c_1. By Cauchy-Schwarz inequality,

$$|A_{4,T}| \leq c\left(\int |g(x_2)|^2\, \pi_T(x) dx_2\right)^{1/2} \left(\int |\Delta e_{h_{2T}}(x_2)|^2\, \pi_T(x) dx_2\right)^{1/2}. \qquad \text{(TR.37)}$$

Since $\pi_T(x) \leq c_2\left(h_{1T} + h_{2T}\right) q(x)$, for a constant c_2, we get

$$\int |g(x_2)|^2\, \pi_T(x) dx = O\left(\left(h_{1T} + h_{2T}\right) \|g\|^2_{L^2_{x_1}(F^*)}\right). \qquad \text{(TR.38)}$$

To bound $\int |\Delta e_{h_{2T}}(x_2)|^2\, \pi_T(x) dx_2$, we apply the argument in the proof of Lemma B.8, with $k(z) = 1(|z| < c_1)$. Then,

$$\begin{aligned}
\int |\Delta e_{h_T}(x_2)|^2\, \pi_T(x) dx_2 &\leq c_3 h_{2T}^2 \int |\nabla e(y_2)|^2\, \pi_T(x_1,y_2) dy_2 \\
&\quad + c_4 h_{2T} \int |\nabla e(y)|^2 \left(\int_0^{h_{2T}} \int |k(z_2)|\, |\pi_T(x_1, y_2 + tz_2) - \pi_T(x_1,y_2)|\, dz_2 dt\right) dy_2 \\
&\leq c_5 h_{2T}^2 \int |\nabla e(y_2)|^2\, \pi_T(x_1,y_2) dy_2 \\
&\quad + c_6 h_{2T} \int |\nabla e(y_2)|^2 \left(\int_0^{h_{2T}} \int |k(z_2)|\, \pi_T(x_1, y_2 + tz_2) dz_2 dt\right) dy_2,
\end{aligned}$$

for some constants $c_3, ..., c_6$. Using

$$\begin{aligned}
&\int_0^{h_{2T}} \int |k(z_2)|\, \pi_T(x_1, y_2 + tz_2) dz_2 dt \\
= &\int_0^{h_{2T}} \int |k(z_2)| \int |K(u)|\, |f\left(y + tz + h_T u\right) - f(y+tz)|\, du dz dt \leq c_7 h_{2T}\left(h_{1T} + h_{2T}\right) q(y_2),
\end{aligned}$$

for large T and a constant c_7, we deduce

$$\int |\Delta e_{h_{2T}}(x_2)|^2\, \pi_T(x) dx_2 = O\left(h_{2T}^2\left(h_{1T} + h_{2T}\right) \|\nabla e\|^2_{L^2_{x_1}(F^*)}\right). \qquad \text{(TR.39)}$$

From (TR.37), (TR.38) and (TR.39) we get

$$A_{4,T} = O\left(h_{2T}\left(h_{1T} + h_{2T}\right) \|g\|_{L^2_{x_1}(F^*)} \|\nabla e\|_{L^2_{x_1}(F^*)}\right). \qquad \text{(TR.40)}$$

From (TR.30), (TR.31), (TR.32), (TR.33), (TR.35), (TR.36) and (TR.40), we derive the asymptotic expansion of $E\left[V_T(g)V_T(e)\right]$:

$$
\begin{aligned}
&E\left[V_T(g,x_1)V_T(e,x_1)\right] = \omega^2 f_{U_1}(x_1)Cov\left[g(U_2),e(U_2)|U_1=x_1\right] \\
&+O\left(h_{1T}^{d_1}\left(h_{1T}+h_{2T}\right)^2\|g\|_{L^2_{x_1}(F^*)}\|e\|_{L^2_{x_1}(F^*)}\right)+O\left(\|g\|_{L^2_{x_1}(F)}\rho_{x_1}(e,\kappa h_{2T})\right) \\
&+O\left(h_{2T}\|g\|_{L^2_{x_1}(F)}\|\nabla e\|_{L^2_{x_1}(F)}+h_{2T}^{3/2}\|g\|_{L^2_{x_1}(F)}\|\nabla e\|_{L^2_{x_1}(F^*)}\right) \\
&+O\left(\left(h_{1T}+h_{2T}\right)\|g\|_{L^2_{x_1}(F^*)}\|e\|_{L^2_{x_1}(F^*)}\right)+O\left(h_{2T}\left(h_{1T}+h_{2T}\right)\|g\|_{L^2_{x_1}(F^*)}\|\nabla e\|_{L^2_{x_1}(F^*)}\right).
\end{aligned}
$$

The conclusion follows.

9.10 Proof of Lemma B.10

(i) Let us apply Lemma B.9 with $U_1 = X_1$, $U_2 = (W,Z_1)$, and $g(u_2) = = \left(\psi_{x_1,j}\right)(z_1)\,\Omega_{x_1,0}(z_1)g_{\varphi_0}(w) =: g_{x_1,j}(u_2)$, for any $j\in\mathbb{N}$. We have $E\left[g_{x_1,j}\left(U_2\right)|U_1=x_1\right] = 0$ and

$$
\begin{aligned}
& V\left[g_{x_1,j}\left(U_2\right)|U_1=x_1\right] \\
= \; & \frac{1}{\nu_{x_1,j}}E\left[\left(A_{x_1}\phi_{x_1j}\right)(Z_1)^2\,\Omega_{x_1,0}(Z_1)^2E\left[g_{\varphi_0}(W)^2 \mid Z_1, X_1=x_1\right]|X_1=x_1\right] \\
= \; & \frac{1}{\nu_{x_1,j}}E\left[\left(A_{x_1}\phi_{x_1j}\right)(Z_1)^2\,\Omega_{x_1,0}(Z_1)|X_1=x_1\right] = \frac{1}{\nu_{x_1,j}}\langle\phi_{x_1j},A^*_{x_1}A_{x_1}\phi_{x_1j}\rangle_{H^l(\mathcal{X}_2)} = 1,
\end{aligned}
$$

where we have used Assumption 5. Thus, we get

$$
\begin{aligned}
E\left[Z^2_{x_1,j,T}\right] \; = \; & f_{X_1}(x_1)\omega^2+O\left(\rho_{x_1}\left(g_{x_1,j},\kappa h_T\right)\right) \\
& +O\left(h_T\|\nabla g_{x_1,j}\|_{L^2_{x_1}(F)}+\left(h_{x_1,T}+h_T\right)\|g_{x_1,j}\|^2_{L^2_{x_1}(F^*)}+h_T^{3/2}\|\nabla g_{x_1,j}\|_{L^2_{x_1}(F^*)}\right. \\
& \left.+h_T\left(h_{x_1,T}+h_T\right)\|g_{x_1,j}\|_{L^2_{x_1}(F^*)}\|\nabla g_{x_1,j}\|_{L^2_{x_1}(F^*)}\right).
\end{aligned}
$$

The conclusion follows if the terms $\|g_{x_1,j}\|_{L^2_{x_1}(F^*)}$, $\|\nabla g_{x_1,j}\|_{L^2_{x_1}(F)}$ and $\|\nabla g_{x_1,j}\|_{L^2_{x_1}(F^*)}$ are bounded and $\rho_{x_1}\left(g_{x_1,j},\kappa h_T\right)=o(1)$, uniformly in $j\in\mathbb{N}$. By using the Cauchy-Schwarz inequality and Assumption B.4, $\|g_{x_1,j}\|_{L^2_{x_1}(F^*)}$, $\|\nabla g_{x_1,j}\|_{L^2_{x_1}(F)}$ and $\|\nabla g_{x_1,j}\|_{L^2_{x_1}(F^*)}$ are uniformly bounded by Assumptions B.8 (iii)-(iv). Finally, the conclusion follows from $\rho_{x_1}\left(g_{x_1,j},\kappa h_T\right)^2 = E\left[g_{x_1,j}\left(U_2\right)^2 1\left(U_2\in\partial\mathcal{S}\left(\kappa h_T\right)\right)|X_1=x_1\right] \leq E\left[|g_{x_1,j}\left(U_2\right)|^4|X_1=x_1\right]^{1/2}$ $P\left(U_2\in\partial\mathcal{S}\left(\kappa h_T\right)|X_1=x_1\right)^{1/2}$ and Assumption B.8 (iii).

(ii) We apply Lemma B.9 with $g=g_{x_1,j}$ and $e=g_{x_1,l}$, for any $j\neq l\in\mathbb{N}$. Similarly to above, we have $Cov\left[g_{x_1,j}\left(U_2\right),g_{x_1,l}(U_2)|U_1=x_1\right] = \frac{1}{\sqrt{\nu_{x_1,j}\nu_{x_1,l}}}\langle\phi_{x_1j},A^*_{x_1}A_{x_1}\phi_{x_1l}\rangle_{H^l(\mathcal{X}_2)} = 0$, for $j\neq l$. Thus, we get $E\left[Z_{x_1,j,T}Z_{x_1,l,T}\right]=o(1)$, uniformly in $j\neq l$, using the bounds in point (i).

9.11 Proof of Lemma B.11

We have $\sum_{j,l=1}^{\infty}\alpha_{j,l}E\left[Z_jZ_l\right]-\sum_{j=1}^{\infty}\alpha_{j,j}E\left[Z_j^2\right]=\sum_{j,l=1,\ j\neq l}^{\infty}\alpha_{j,l}E\left[Z_jZ_l\right]=\sum_{j,l=1,\ j\neq l}^{\infty}\alpha_{j,l}\rho_{j,l}E\left[Z_j^2\right]^{1/2}E\left[Z_l^2\right]^{1/2}$.
Let us now bound the term in the RHS. We have

$$\begin{aligned}
&\left|\sum_{j,l=1,\ j\neq l}^{\infty}\alpha_{j,l}\rho_{j,l}E\left[Z_j^2\right]^{1/2}E\left[Z_l^2\right]^{1/2}\right| \\
\leq\ & \sum_{j=1}^{\infty}\left(\sqrt{\alpha_{j,j}}E\left[Z_j^2\right]^{1/2}\right)\left(\sum_{l=1:l\neq j}^{\infty}\frac{\alpha_{j,l}}{\sqrt{\alpha_{j,j}}\sqrt{\alpha_{l,l}}}\left|\rho_{j,l}\right|\sqrt{\alpha_{l,l}}E\left[Z_l^2\right]^{1/2}\right) \\
\leq\ & \left(\sum_{j=1}^{\infty}\alpha_{j,j}E\left[Z_j^2\right]\right)^{1/2}\left(\sum_{j=1}^{\infty}\left(\sum_{l=1:l\neq j}^{\infty}\frac{\alpha_{j,l}}{\sqrt{\alpha_{j,j}}\sqrt{\alpha_{l,l}}}\left|\rho_{j,l}\right|\sqrt{\alpha_{l,l}}E\left[Z_l^2\right]^{1/2}\right)^2\right)^{1/2},
\end{aligned} \tag{TR.41}$$

by Cauchy-Schwarz inequality. Moreover, again by Cauchy-Schwarz inequality:

$$\begin{aligned}
\sum_{l=1:l\neq j}^{\infty}\frac{\alpha_{j,l}}{\sqrt{\alpha_{j,j}}\sqrt{\alpha_{l,l}}}\left|\rho_{j,l}\right|\sqrt{\alpha_{l,l}}E\left[Z_l^2\right]^{1/2} &\leq \left(\sum_{l=1:l\neq j}^{\infty}\frac{\alpha_{j,l}^2}{\alpha_{j,j}\alpha_{l,l}}\rho_{j,l}^2\right)^{1/2}\left(\sum_{l=1:l\neq j}^{\infty}\alpha_{l,l}E\left[Z_l^2\right]\right)^{1/2} \\
&\leq \left(\sum_{l=1:l\neq j}^{\infty}\frac{\alpha_{j,l}^2}{\alpha_{j,j}\alpha_{l,l}}\rho_{j,l}^2\right)^{1/2}\left(\sum_{l=1}^{\infty}\alpha_{l,l}E\left[Z_l^2\right]\right)^{1/2}.
\end{aligned}$$

From (TR.41) we get

$$\left|\sum_{j,l=1,\ j\neq l}^{\infty}\alpha_{j,l}\rho_{j,l}E\left[Z_j^2\right]^{1/2}E\left[Z_j^2\right]^{1/2}\right|\leq\left(\sum_{j=1}^{\infty}\alpha_{j,j}E\left[Z_j^2\right]\right)\left(\sum_{j=1}^{\infty}\sum_{l=1:l\neq j}^{\infty}\frac{\alpha_{j,l}^2}{\alpha_{j,j}\alpha_{l,l}}\rho_{j,l}^2\right)^{1/2},$$

and the conclusion follows.

9.12 Proof of Lemma B.12

By a Taylor expansion of the kernel estimator bias we get:

$$\begin{aligned}
\zeta_{x_1}(z_1) &= \frac{1}{m!}\sum_{|\alpha_1|+|\alpha_2|=m}h_T^{|\alpha_1|}h_{x_1,T}^{|\alpha_2|}\int\int K(\eta)K(\zeta)K(\xi)(y-\varphi_{x_1,0}(x_2)) \\
&\quad \frac{\nabla^{\alpha_1}\nabla^{\alpha_2}f_{W,Z}(w+h_T\eta,z_1+h_T\zeta,x_1+h_{x_1,T}\xi)}{f_Z(z)}dwd\eta d\zeta d\xi,
\end{aligned}$$

where ∇^{α_1} and ∇^{α_2} denote gradient operators w.r.t. (Y,X_2,Z_1) and X_1, respectively. Thus, we get $\zeta_{x_1}=h_{x_1,T}^m\Xi_{x_1}+\Gamma_{x_1}$, where function Γ_{x_1} is such that $\sup_{z_1\in\mathcal{Z}_1}\left|\Gamma_{x_1}(z_1)\right|\leq C(h_Th_{x_1,T}^{m-1}+h_T^m)\left(\int\frac{q\left(w,z_1,x_1\right)^2}{f_{W,Z}(w,z)}dw\right)^{1/2}$ from Assumptions 5, B.4 and B.6. Then, the conclusion follows.

9.13 Proof of Lemma C.1

Let us write $\hat{\Psi}_{x_1}(z_1)$ as:

$$\hat{\Psi}_{x_1}(z_1) = \frac{1}{T}\sum_{t=1}^{T}\left(Y_t - V_{T,t}\right) K_h\left(Z_t - z\right), \qquad \text{(TR.42)}$$

where:

$$V_{T,t} = \int \varphi_{x_1,0}(X_{2,t} - h_T u)K(u)du,$$

and:

$$K_h\left(Z_t - z\right) := \frac{1}{h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}}} K\left(\frac{Z_{1,t} - z_1}{h_T}\right) K\left(\frac{X_{1,t} - x_1}{h_{x_1,T}}\right).$$

(1) The bias term is:

$$E\left[\hat{\Psi}_{x_1}(z_1)\right] = E\left[\left(Y - V_T\right) K_h\left(Z - z\right)\right].$$

By a change of variable:

$$\begin{aligned} E\left[\left(Y - V_T\right) K_h\left(Z - z\right)\right] &= E\left[E\left[Y - V_T|Z\right] K_h\left(Z - z\right)\right] \\ &= \int \mu(z_1 + h_T u_1, x_1 + h_{x_1,T} u_2)K(u)du \\ &\quad - \int \eta_T(z_1 + h_T u_1, x_1 + h_{x_1,T} u_2)K(u)du, \end{aligned}$$

where $\eta_T(z) = E\left[V_T|Z = z\right] f_Z(z)$. By standard bias expansions:

$$\begin{aligned} \int \mu(z_1 + h_T u_1, x_1 + h_{x_1,T} u_2)K(u)du &= \mu(z) + O\left(\|D^m \mu\|_\infty \left(h_T^m + h_{x_1,T}^m\right)\right), \\ \int \eta_T(z_1 + h_T u_1, x_1 + h_{x_1,T} u_2)K(u)du &= \eta_T(z) + O\left(\|D^m \eta_T\|_\infty \left(h_T^m + h_{x_1,T}^m\right)\right), \end{aligned}$$

uniformly in $z \in \mathcal{Z}$. Moreover:

$$\begin{aligned} \eta_T(z) &= \int\int \varphi_{x_1,0}(x_2 - h_T u) f_{X_2,Z}(x_2, z)K(u)dx_2 du \\ &= \int\int \varphi_{x_1,0}(x_2) f_{X_2,Z}(x_2 + h_T u, z)K(u)dx_2 du \\ &= E\left[\varphi_{x_1,0}\left(X_2\right)|Z = z\right] f_Z(z) + O\left(\|D^m f_{X_2,Z}\|_\infty h_T^m\right), \end{aligned}$$

uniformly in $z \in \mathcal{Z}$, and:

$$\begin{aligned} \|D^m \eta_T\|_\infty &\leq \|D^m f_{X_2,Z}\|_\infty \int \left|\varphi_{x_1,0}(x_2)\right| dx_2 \int |K(u)|\, du \\ &\leq \|D^m f_{X_2,Z}\|_\infty \left\|\varphi_{x_1,0}\right\|_{L^2(\mathcal{X}_2)} \int |K(u)|\, du < \infty, \end{aligned}$$

from Assumption B.3. Thus, by using $E\left[Y-\varphi_{x_1,0}\left(X_2\right)|Z=z\right]=0$, we deduce

$$\sup_{z_1\in\mathcal{Z}_1} E\left[\left(Y-V_T\right)K_h\left(Z-z\right)\right]=O\left(h_T^m+h_{x_1,T}^m\right),$$

uniformly in $x_1\in\mathcal{X}_1$, and Part (1) follows.

(2) The proof is similar to the proof of Theorem 2 in Hansen (2008). Let us first truncate the variables $W_{T,t}:=Y_t-V_{T,t}$ in (TR.42) and show that the truncation effect is negligible. Define $\tau_{x_1,T}=\left(\frac{Th_{x_1,T}^{d_{X_1}}h_T^{d_{Z_1}}}{\log T}\right)^{1/(2s)}$, for any $x_1\in\mathcal{X}_1$ and some $s>1$. Write:

$$\begin{aligned}\hat{\Psi}_{x_1}(z_1) &= \frac{1}{T}\sum_{t=1}^{T}W_{T,t}1\left\{|W_{T,t}|\le\tau_{x_1,T}\right\}K_h\left(Z_t-z\right)\\ &\quad+\frac{1}{T}\sum_{t=1}^{T}W_{T,t}1\left\{|W_{T,t}|>\tau_{x_1,T}\right\}K_h\left(Z_t-z\right)=:\tilde{\Psi}_{x_1}(z_1)+R_{x_1}(z_1).\end{aligned}$$

By the same argument as in Hansen (2008), p. 740, and Assumptions B.1-B.3, we have:

$$E\left|R_{x_1}(z_1)\right|\le C\tau_{x_1,T}^{-s}=O\left(\sqrt{\frac{\log T}{Th_{x_1,T}^{d_{X_1}}h_T^{d_{Z_1}}}}\right),$$

uniformly in $x_1\in\mathcal{X}_1$, $z_1\in\mathcal{Z}_1$. Thus, the conclusion follows if we show that $\sup_{z_1\in\mathcal{Z}_1}\tilde{\Psi}_{x_1}(z_1)=O_p\left(\sqrt{\frac{\log T}{Th_{x_1,T}^{d_{X_1}}h_T^{d_{Z_1}}}}\right)$ uniformly in $x_1\in\mathcal{X}_1$. For this purpose, define $\hat{\Psi}(z):=\sqrt{\frac{Th_{x_1,T}^{d_{X_1}}h_T^{d_{Z_1}}}{\log T}}\tilde{\Psi}_{x_1}(z_1)$.

Let us introduce a covering of $\mathcal{Z}$ by n_T hyperballs $B_{j,T}:=\left\{z:|z-z_{j,T}|\le\frac{C}{n_T^{1/d_Z}}\right\}$, $j=1,\cdots,n_T$, where C is a constant, $z_{j,T}\in\mathcal{Z}$, n_T is such that $n_T=O\left(T^c\right)$ for some $c>0$, and $\frac{\sqrt{T}}{\rho_T n_T^{1/d_Z}\sqrt{\log T}}=o\left(1\right)$, where $\rho_T:=\inf_{x_1\in\mathcal{X}_1}\left(\sqrt{h_{x_1,T}^{d_{X_1}}h_T^{d_{Z_1}}}\min\left\{h_{x_1,T},h_T\right\}\right)$. Then, from Assumption B.2 any by using that $\sup_{x_1\in\mathcal{X}_1}\left|\frac{\partial h_{x_1,T}}{\partial x_1}\right|=O(1)$, uniformly in $z\in B_{j,T}$ and $j=1,\cdots,n_T$ we have

$$\begin{aligned}\left|\hat{\Psi}(z)-\hat{\Psi}(z_{j,T})\right| &\le \frac{C_1}{\rho_T n_T^{1/d_Z}}\frac{1}{\sqrt{T\log T}}\sum_{t=1}^{T}|W_{T,t}|=o_p\left(1\right),\\ \left|E\left[\hat{\Psi}(z)\right]-E\left[\hat{\Psi}(z_{j,T})\right]\right| &\le \frac{C_1}{\rho_T n_T^{1/d_Z}}\frac{1}{\sqrt{T\log T}}\sum_{t=1}^{T}E\left|W_{T,t}\right|=o_p\left(1\right),\end{aligned}$$

for a constant C_1. It follows that:

$$\sup_{z\in\mathcal{Z}}\left|\hat{\Psi}(z)-E\left[\hat{\Psi}(z)\right]\right|\le\sup_{j=1,\cdots,n_T}\left|\hat{\Psi}(z_{j,T})-E\left[\hat{\Psi}(z_{j,T})\right]\right|+o_p\left(1\right).$$

Thus, the conclusion follows if we prove that the first term in the RHS is $o_p(1)$. For this purpose, let $\varepsilon > 0$ be given. We have:

$$\begin{aligned} & P\left[\sup_{j=1,\cdots,n_T}\left|\hat{\Psi}(z_{j,T}) - E\left[\hat{\Psi}(z_{j,T})\right]\right| \geq \varepsilon\right] \leq \sum_{j=1}^{n_T} P\left[\left|\hat{\Psi}(z_{j,T}) - E\left[\hat{\Psi}(z_{j,T})\right]\right| \geq \varepsilon\right] \\ \leq \quad & n_T \sup_{z\in\mathcal{Z}} P\left[\left|\hat{\Psi}(z) - E\left[\hat{\Psi}(z)\right]\right| \geq \varepsilon\right] \\ = \quad & n_T \sup_{z\in\mathcal{Z}} P\left[\left|\tilde{\Psi}_{x_1}(z_1) - E\left[\tilde{\Psi}_{x_1}(z_1)\right]\right| \geq \varepsilon\sqrt{\frac{\log T}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}}}}\right]. \end{aligned} \tag{TR.43}$$

In order to bound the probability in the RHS, we apply a large deviation bound as in Hansen (2008), p. 741-742. Let us write $\hat{\Psi}_{x_1}(z_1) = \frac{1}{T}\sum_{t=1}^{T}\kappa_{t,T}(z)$, where $\kappa_{t,T}(z) = W_{T,t}1\{|W_{T,t}| \leq \tau_{x_1,T}\} K_h(Z_t - z) - E[W_{T,t}1\{|W_{T,t}| \leq \tau_{x_1,T}\} K_h(Z_t - z)]$. From Assumption B.2 we have $|\kappa_{t,T}(z)| \leq \frac{2\tau_{x_1,T}\|K\|_\infty}{h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}}}$, uniformly in $z \in \mathcal{Z}$. Let us now bound $E\left[|\kappa_{tT}(z)|^2\right]$. We have:

$$\begin{aligned} E\left[(Y - V_T)^2 K_h(Z - z)^2\right] &= E\left[E\left[(Y - V_T)^2 |Z\right] K_h(Z - z)^2\right] \\ &= \frac{1}{h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}}}\int \sigma_T^2(z_1 + h_T u_1, x_1 + h_{x_1,T}u_2) K(u)^2 du, \end{aligned}$$

where $\sigma_T^2(z) = E\left[(Y - V_T)^2 |Z = z\right] f_Z(z)$. Now:

$$\begin{aligned} E\left[V_T^2|Z = z\right] &= \int\int E\left[\varphi_{x_1,0}(X_2 - h_T u)\varphi_{x_1,0}(X_2 - h_T v)|Z = z\right] K(u)K(v)dudv \\ &\leq \left(\int |K(u)|\, du\right)\int E\left[\varphi_{x_1,0}(X_2 - h_T u)^2|Z = z\right] |K(u)|\, du, \end{aligned}$$

and then:

$$\begin{aligned} E\left[V_T^2|Z = z\right] f_Z(z) &\leq \left(\int |K(u)|\, du\right)\int\int \varphi_{x_1,0}(x_2 - h_T u)^2 f_{X_2,Z}(x_2, z)\, |K(u)|\, dudx_2 \\ &= \left(\int |K(u)|\, du\right)\int\int \varphi_{x_1,0}(x_2)^2 f_{X_2,Z}(x_2 + h_T u, z)\, |K(u)|\, dudx_2 \\ &\leq \left(\int |K(u)|\, du\right)^2 \|f_{X_2,Z}\|_\infty \left\|\varphi_{x_1,0}\right\|_{L^2(\mathcal{X}_2)}^2. \end{aligned}$$

Thus, we get $\left\|\sigma_T^2\right\|_\infty < \infty$ and:

$$E\left[(Y - V_T)^2 K_h(Z - z)^2\right] = O\left(\frac{1}{h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}}}\right), \tag{TR.44}$$

uniformly in $z \in \mathcal{Z}$. Hence, we get $\sup_{z\in\mathcal{Z}} E\left[|\kappa_{tT}(z)|^2\right] \leq c_2/h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}}$ for large T and a constant $c_2 > 0$. By following the argument in Hansen (2008), p. 741-742, and using Assumptions B.1-B.3, we get:

$$P\left[\left|\tilde{\Psi}_{x_1}(z_1) - E\left[\tilde{\Psi}_{x_1}(z_1)\right]\right| \geq \varepsilon\sqrt{\frac{\log T}{T h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}}}}\right] \leq 4\exp\left(-c_1\varepsilon^2 \log T\right),$$

uniformly in $z \in \mathcal{Z}_T$, for a constant $c_1 > 0$. It follows from (TR.43) that $P\left[\sup_{j=1,\cdots,n_T}\left|\hat{\Psi}(z_{j,T}) - E\left[\hat{\Psi}(z_{j,T})\right]\right| \geq \varepsilon\right] = O\left(n_T T^{-c_1\varepsilon^2}\right) = O\left(T^{c-c_1\varepsilon^2}\right)$. By choosing $\varepsilon^2 > c/c_1$, we get that $\lim\sup_{T\to\infty} P\left[\sup_{j=1,\cdots,n_T}\left|\hat{\Psi}(z_{j,T}) - E\left[\hat{\Psi}(z_{j,T})\right]\right| \geq \varepsilon\right] = 0$, and Part (ii) follows.

(iii) We can write:

$$\int g(z,x_2)\hat{\Psi}_{x_1}(z_1)dz_1 = \frac{1}{T}\sum_{t=1}^{T}\left(Y_t - V_{T,t}\right) G_{T,t} K_{h_{x_1,T}}\left(X_{1,t} - x_1\right),$$

where $G_{T,t} = \int g(Z_{1,t} + h_T u, x_1, x_2)K(u)du$. Since the variables $G_{T,t}$ are bounded, and set $\mathcal{X}_1 \times \mathcal{X}_2$ is compact, the conclusion follows by a similar argument as in Part (ii).

9.14 Proof of Lemma C.2

Part (1) follows from (TR.42) and (TR.44). Let us now prove Part (2). The fourth centered moment of $\hat{\Psi}_{x_1}(z_1)$ is given by:

$$E\left[\left(\hat{\Psi}_{x_1}(z_1) - E\left[\hat{\Psi}_{x_1}(z_1)\right]\right)^4\right] = \frac{1}{T^3}E\left[\kappa_T(z)^4\right] + \frac{3(T-1)}{T^3}E\left[\kappa_T(z)^2\right]^2,$$

where:

$$\kappa_T(z) = (Y - V_T)K_h(Z - z) - E\left[(Y - V_T)K_h(Z - z)\right].$$

By the proof of Lemma C.1 we know:

$$\sup_{z_1\in\mathcal{Z}_1} E\left[\kappa_T(z)^2\right] = O\left(\frac{1}{h_{x_1,T}^{d_{X_1}} h_T^{d_{Z_1}}}\right).$$

Moreover:

$$\begin{aligned} E\left[\kappa_T(z)^4\right] &= E\left[\left((Y - V_T)K_h(Z - z)\right)^4\right] - 8E\left[\left((Y - V_T)K_h(Z - z)\right)^3\right]E\left[(Y - V_T)K_h(Z - z)\right] \\ &\quad + 6E\left[\left((Y - V_T)K_h(Z - z)\right)^2\right]^2 + E\left[(Y - V_T)K_h(Z - z)\right]^4. \end{aligned}$$

Let us focus on the first term. We have:

$$E\left[\left(\left(Y-V_T\right)K_h\left(Z-z\right)\right)^4\right]=\frac{1}{h_{x_1,T}^{3d_{X_1}}h_T^{3d_{Z_1}}}\int\tau_{4,T}\left(z_1+h_Tu_1,x_1+h_{x_1,T}u_2\right)K(u)^4du,$$

where $\tau_{4,T}\left(z\right)=E\left[\left(Y-V_T\right)^4|Z=z\right]f_Z(z)$. By similar arguments as in the proof of Lemma C.1 we get:

$$E\left[V_T^4|Z=z\right]f_Z(z)\leq\left(\int\left|K(u)\right|du\right)^4\left\|f_{X_2,Z}\right\|_\infty\int\varphi_{x_1,0}(x_2)^4dx_2.$$

Thus, we get:

$$\sup_{z_1\in\mathcal{Z}_1}E\left[\left(\left(Y-V_T\right)K_h\left(Z-z\right)\right)^4\right]=O\left(\frac{1}{h_{x_1,T}^{3d_{X_1}}h_T^{3d_{Z_1}}}\right).$$

It follows:

$$\sup_{z_1\in\mathcal{Z}_1}E\left[\kappa_T\left(z\right)^4\right]=O\left(\frac{1}{h_{x_1,T}^{3d_{X_1}}h_T^{3d_{Z_1}}}\right),$$

and:

$$\sup_{z_1\in\mathcal{Z}_1}E\left[\left(\hat{\Psi}_{x_1}(z_1)-E\left[\hat{\Psi}_{x_1}(z_1)\right]\right)^4\right]=\left(\frac{1}{T^3h_{x_1,T}^{3d_{X_1}}h_T^{3d_{Z_1}}}+\frac{1}{T^2h_{x_1,T}^{2d_{X_1}}h_T^{2d_{Z_1}}}\right).$$

The conclusion follows.

10 Characterization of operator $\mathcal{E}$

In this Section we discuss the characterization of operator $\mathcal{E}$ when $l,d_{X_2}\geq 1$ and the Sobolev embedding condition $2l>d_{X_2}$ is satisfied. In order to simplify the notation, we focus on the case $l=d_{X_2}=2$, but the arguments can be extended to the more general case $2l>d_{X_2}$ (see also CGS for the case $d_{X_2}=1$).

We have:

$$\begin{aligned}\left\langle\phi,u\right\rangle_{H^l(\mathcal{X}_2)} &= \left\langle\phi,u\right\rangle_{L^2(\mathcal{X}_2)}+\left\langle\nabla_1\phi,\nabla_1u\right\rangle_{L^2(\mathcal{X}_2)}+\left\langle\nabla_2\phi,\nabla_2u\right\rangle_{L^2(\mathcal{X}_2)}\\ &\quad+\left\langle\nabla_1^2\phi,\nabla_1^2u\right\rangle_{L^2(\mathcal{X}_2)}+\left\langle\nabla_2^2\phi,\nabla_2^2u\right\rangle_{L^2(\mathcal{X}_2)}+2\left\langle\nabla_1\nabla_2\phi,\nabla_1\nabla_2u\right\rangle_{L^2(\mathcal{X}_2)},\end{aligned}$$

where ∇_1 and ∇_2 denote the gradients w.r.t. the components ξ_1 and ξ_2 of $x_2=(\xi_1,\xi_2)$. By using partial integration, we have:

$$\left\langle\nabla_1\phi,\nabla_1u\right\rangle_{L^2(\mathcal{X}_2)}=\int_0^1\phi\nabla_1u|_{\xi_1=0}^1\,d\xi_2-\left\langle\phi,\nabla_1^2u\right\rangle_{L^2(\mathcal{X}_2)},$$

$$\left\langle\nabla_2\phi,\nabla_2u\right\rangle_{L^2(\mathcal{X}_2)}=\int_0^1\phi\nabla_2u|_{\xi_2=0}^1\,d\xi_1-\left\langle\phi,\nabla_2^2u\right\rangle_{L^2(\mathcal{X}_2)},$$

$$\begin{aligned}\left\langle \nabla_1^2\phi, \nabla_1^2 u\right\rangle_{L^2(\mathcal{X}_2)} &= \int_0^1 \nabla_1\phi\nabla_1^2 u\big|_{\xi_1=0}^1 d\xi_2 - \left\langle \nabla_1\phi, \nabla_1^3 u\right\rangle_{L^2(\mathcal{X}_2)} \\ &= \int_0^1 \nabla_1\phi\nabla_1^2 u\big|_{\xi_1=0}^1 d\xi_2 - \int_0^1 \phi\nabla_1^3 u\big|_{\xi_1=0}^1 d\xi_2 + \left\langle \phi, \nabla_1^4 u\right\rangle_{L^2(\mathcal{X}_2)},\end{aligned}$$

$$\begin{aligned}\left\langle \nabla_2^2\phi, \nabla_2^2 u\right\rangle_{L^2(\mathcal{X}_2)} &= \int_0^1 \nabla_2\phi\nabla_2^2 u\big|_{\xi_2=0}^1 d\xi_1 - \left\langle \nabla_2\phi, \nabla_2^3 u\right\rangle_{L^2(\mathcal{X}_2)} \\ &= \int_0^1 \nabla_2\phi\nabla_2^2 u\big|_{\xi_2=0}^1 d\xi_1 - \int_0^1 \phi\nabla_2^3 u\big|_{\xi_2=0}^1 d\xi_1 + \left\langle \phi, \nabla_2^4 u\right\rangle_{L^2(\mathcal{X}_2)},\end{aligned}$$

and:

$$\begin{aligned}\langle \nabla_1\nabla_2\phi, \nabla_1\nabla_2 u\rangle_{L^2(\mathcal{X}_2)} &= \int_0^1 \nabla_2\phi\nabla_1\nabla_2 u|_{\xi_1=0}^1 d\xi_2 - \left\langle \nabla_2\phi, \nabla_1^2\nabla_2 u\right\rangle_{L^2(\mathcal{X}_2)} \\ &= \int_0^1 \nabla_2\phi\nabla_1\nabla_2 u|_{\xi_1=0}^1 d\xi_2 - \int_0^1 \phi\nabla_1^2\nabla_2 u\big|_{\xi_2=0}^1 d\xi_1 + \left\langle \phi, \nabla_1^2\nabla_2^2 u\right\rangle_{L^2(\mathcal{X}_2)} \\ &= (\phi\nabla_1\nabla_2 u)(1,1) - (\phi\nabla_1\nabla_2 u)(1,0) - (\phi\nabla_1\nabla_2 u)(0,1) + (\phi\nabla_1\nabla_2 u)(0,0) \\ &\quad - \int_0^1 \phi\nabla_1\nabla_2^2 u\big|_{\xi_1=0}^1 d\xi_2 - \int_0^1 \phi\nabla_1^2\nabla_2 u\big|_{\xi_2=0}^1 d\xi_1 + \left\langle \phi, \nabla_1^2\nabla_2^2 u\right\rangle_{L^2(\mathcal{X}_2)}.\end{aligned}$$

Thus, we get:

$$\begin{aligned}\langle \phi, u\rangle_{H^l(\mathcal{X}_2)} &= \langle \phi, \mathcal{D}u\rangle_{L^2(\mathcal{X}_2)} + \int_0^1 \phi\nabla_1 u|_{\xi_1=0}^1 d\xi_2 + \int_0^1 \phi\nabla_2 u|_{\xi_2=0}^1 d\xi_1 \\ &\quad + \int_0^1 \nabla_1\phi\nabla_1^2 u\big|_{\xi_1=0}^1 d\xi_2 - \int_0^1 \phi\nabla_1^3 u\big|_{\xi_1=0}^1 d\xi_2 \\ &\quad + \int_0^1 \nabla_2\phi\nabla_2^2 u\big|_{\xi_2=0}^1 d\xi_1 - \int_0^1 \phi\nabla_2^3 u\big|_{\xi_2=0}^1 d\xi_1 \\ &\quad +2\left[(\phi\nabla_1\nabla_2 u)(1,1) - (\phi\nabla_1\nabla_2 u)(1,0) - (\phi\nabla_1\nabla_2 u)(0,1) + (\phi\nabla_1\nabla_2 u)(0,0)\right] \\ &\quad -2\int_0^1 \phi\nabla_1\nabla_2^2 u\big|_{\xi_1=0}^1 d\xi_2 - 2\int_0^1 \phi\nabla_1^2\nabla_2 u\big|_{\xi_2=0}^1 d\xi_1.\end{aligned}$$

Since $u \in H_0^{2l}(\mathcal{X}_2)$, the boundary terms involving an odd-order derivative of u vanish. We get:

$$\langle \phi, \mathcal{D}u\rangle_{L^2(\mathcal{X}_2)} = \langle \phi, u\rangle_{H^l(\mathcal{X}_2)} + \bar{\mathcal{T}}_u(\phi),$$

where:

$$\begin{aligned}\bar{\mathcal{T}}_u(\phi) &= -\int_0^1 \nabla_1\phi(1,\xi_2)\nabla_1^2 u(1,\xi_2)\, d\xi_2 + \int_0^1 \nabla_1\phi(0,\xi_2)\nabla_1^2 u(0,\xi_2)\, d\xi_2 \\ &\quad -\int_0^1 \nabla_2\phi(\xi_1,1)\nabla_2^2 u(\xi_1,1)\, d\xi_1 + \int_0^1 \nabla_2\phi(\xi_1,0)\nabla_2^2 u(\xi_1,0)\, d\xi_1 \\ &\quad -2\left[(\phi\nabla_1\nabla_2 u)(1,1) - (\phi\nabla_1\nabla_2 u)(1,0) - (\phi\nabla_1\nabla_2 u)(0,1) + (\phi\nabla_1\nabla_2 u)(0,0)\right].\end{aligned} \tag{TR.45}$$

Now, we have to show that $\bar{\mathcal{T}}_u(\phi)$ can be written as $\bar{\mathcal{T}}_u(\phi) = \langle \phi, \tau_u \rangle_{H^l(\mathcal{X}_2)}$ where $\tau_u \in H^l(\mathcal{X}_2)$. Then, $\mathcal{E}(u) = u + \tau_u$. We consider separately the different components.

For any $a \in \mathcal{X}_2$, the mapping $\phi \mapsto \phi(a)$ in $H^l(\mathcal{X}_2)$ is a continuous linear functional if $d_{X_2} < 2l$ because of the Sobolev embedding theorem. Thus, by the Rietz representation theorem, there exists $\delta_a \in H^l(\mathcal{X}_2)$ such that:

$$\phi(a) = \langle \delta_a, \phi \rangle_{H^l(\mathcal{X}_2)}, \quad \phi \in H^l(\mathcal{X}_2).$$

It follows that we can rewrite the last row in (TR.45) as:

$$\begin{aligned} & -2\left[(\phi\nabla_1\nabla_2 u)(1,1) - (\phi\nabla_1\nabla_2 u)(1,0) - (\phi\nabla_1\nabla_2 u)(0,1) + (\phi\nabla_1\nabla_2 u)(0,0)\right] \\ = \;& 2\left\langle -\nabla_1\nabla_2 u(1,1)\,\delta_{(1,1)} + \nabla_1\nabla_2 u(1,0)\,\delta_{(1,0)} + \nabla_1\nabla_2 u(0,1)\,\delta_{(0,1)} - \nabla_1\nabla_2 u(0,0)\,\delta_{(0,0)}, \phi\right\rangle_{H^l(\mathcal{X}_2)}. \end{aligned}$$

Let us now consider the terms in the first row of the RHS in (TR.45). For any $a \in [0,1]$, function $\xi_2 \to \nabla_1^2 u(a,\xi_2)$ is in $L^2[0,1]$. Indeed, by the Sobolev embedding theorem of one-dimensional spaces:

$$\left|\nabla_1^2 u(a,\xi_2)\right| \le C\left\|\nabla_1^2 u(.,\xi_2)\right\|_{H^1(0,1)},$$

where C is independent of ξ_2, and

$$\int_0^1 \left|\nabla_1^2 u(a,\xi_2)\right|^2 d\xi_2 \le C\int_0^1 \left\|\nabla_1^2 u(.,\xi_2)\right\|_{H^1(0,1)}^2 d\xi_2 \le C\|u\|_{H^{2l}(\mathcal{X}_2)}^2 < \infty.$$

Hence, we can write

$$\int_0^1 \nabla_1\phi(a,\xi_2)\,\nabla_1^2 u(a,\xi_2)\,d\xi_2 = \sum_{j=1}^{\infty}\beta_{1,a,j}\int_0^1 \nabla_1\phi(a,\xi_2)\,\tilde{\chi}_j(\xi_2)\,d\xi_2,$$

where $\beta_{1,a,j} = \int_0^1 \nabla_1^2 u(a,\xi_2)\,\tilde{\chi}_j(\xi_2)\,d\xi_2$ and the $\tilde{\chi}_j$ are the elements of the basis of $L^2[0,1]$ introduced in Appendix 2 of the paper. Define:

$$\mathcal{T}_{1,a,j}(\phi) = \int_0^1 \nabla_1\phi(a,\xi_2)\,\tilde{\chi}_j(\xi_2)\,d\xi_2,$$

for $a \in [0,1]$. Let us first prove that $\mathcal{T}_{1,a,j}$ defines a continuous linear function on $H^l(\mathcal{X}_2)$. We have by Cauchy-Schwartz inequality:

$$\mathcal{T}_{1,a,j}(\phi) \le \left(\int_0^1 |\nabla_1\phi(a,\xi_2)|^2 d\xi_2\right)^{1/2}\left(\int_0^1 \left|\tilde{\chi}_j(\xi_2)\right|^2 d\xi_2\right)^{1/2} = \left(\int_0^1 |\nabla_1\phi(a,\xi_2)|^2 d\xi_2\right)^{1/2}.$$

Now, for given $\xi_2 \in [0,1]$, by the embedding of one-dimensional spaces:

$$|\nabla_1\phi(a,\xi_2)| \le C\|\nabla_1\phi(.,\xi_2)\|_{H^1(0,1)},$$

where C is independent of ξ_2, and

$$\int_0^1 |\nabla_1\phi(a,\xi_2)|^2 d\xi_2 \le C\int_0^1 \|\nabla_1\phi(.,\xi_2)\|_{H^1(0,1)}^2 d\xi_2 \le C\|\phi\|_{H^l(\mathcal{X}_2)}^2.$$

Hence, we get $\mathcal{T}_{1,a,j}(\phi) \leq C \|\phi\|^2_{H^l(\mathcal{X}_2)}$ and $\mathcal{T}_{1,a,j}$ is a continuous linear functional on $H^l(\mathcal{X}_2)$. Thus, by the Rietz representation theorem, there exists $\delta_{1,a,j} \in H^l(\mathcal{X}_2)$ such that:

$$\mathcal{T}_{1,a,j}(\phi) = \langle \delta_{1,a,j}, \phi \rangle_{H^l(\mathcal{X}_2)}, \quad \phi \in H^l(\mathcal{X}_2).$$

It follows that we can rewrite the terms in the first row in the RHS in (TR.45) as

$$\begin{aligned} & -\int_0^1 \nabla_1 \phi(1,\xi_2) \nabla_1^2 u(1,\xi_2) \, d\xi_2 + \int_0^1 \nabla_1 \phi(0,\xi_2) \nabla_1^2 u(0,\xi_2) \, d\xi_2 \\ = & -\left\langle \sum_{j=1}^{\infty} \left(\beta_{1,1,j}\delta_{1,1,j} - \beta_{1,0,j}\delta_{1,0,j}\right), \phi \right\rangle_{H^l(\mathcal{X}_2)}. \end{aligned}$$

In a similar way, it is possible to define functions $\delta_{2,a,j} \in H^l(\mathcal{X}_2)$ that are the Rietz representants of the continuous linear functionals

$$\mathcal{T}_{2,a,j}(\phi) = \int_0^1 \nabla_2 \phi(\xi_1, a) \tilde{\chi}_j(\xi_1) \, d\xi_1,$$

and we can rewrite the terms in the second row in the RHS in (TR.45) as:

$$\begin{aligned} & -\int_0^1 \nabla_2 \phi(\xi_1,1) \nabla_2^2 u(\xi_1,1) \, d\xi_1 + \int_0^1 \nabla_2 \phi(\xi_1,0) \nabla_2^2 u(\xi_1,0) \, d\xi_1 \\ = & -\left\langle \sum_{j=1}^{\infty} \left(\beta_{2,1,j}\delta_{2,1,j} - \beta_{2,0,j}\delta_{2,0,j}\right), \phi \right\rangle_{H^l(\mathcal{X}_2)}. \end{aligned}$$

By putting everything together, we get:

$$\begin{aligned} \mathcal{E}(u) = \; & u - 2\left[\nabla_1\nabla_2 u(1,1)\,\delta_{(1,1)} - \nabla_1\nabla_2 u(1,0)\,\delta_{(1,0)} - \nabla_1\nabla_2 u(0,1)\,\delta_{(0,1)} + \nabla_1\nabla_2 u(0,0)\,\delta_{(0,0)}\right] \\ & - \sum_{j=1}^{\infty} \Big(\left\langle \nabla_1^2 u(1,.), \tilde{\chi}_j \right\rangle_{L^2[0,1]} \delta_{1,1,j} - \left\langle \nabla_1^2 u(0,.), \tilde{\chi}_j \right\rangle_{L^2[0,1]} \delta_{1,0,j} \\ & + \left\langle \nabla_2^2 u(.,1), \tilde{\chi}_j \right\rangle_{L^2[0,1]} \delta_{2,1,j} - \left\langle \nabla_2^2 u(.,0), \tilde{\chi}_j \right\rangle_{L^2[0,1]} \delta_{2,0,j} \Big). \end{aligned}$$