1 Appendix II: Proof of the Theorem 2

Proof of Theorem 2

We prove Theorem 2 using the same four steps as in the proof of Theorem 1. Without loss of generality we prove it for the case of $\alpha = 1,\ \beta = 2$. For expositional convenience we use the subscript 3 to denote the nuisance directions.

Step I. Decompositions

The closed form of the LIVE estimator for $\phi_{12}(y_1, y_2)$ is

$$\hat{\phi}_{12}(y_1, y_2)^T = e_1^T(\mathbf{Y}_{\_}^T\mathbf{K}\mathbf{Y}_{\_})^{-1}\mathbf{Y}_{\_}^T\mathbf{K}\mathbf{W}\tilde{\mathbf{R}} \tag{1}$$

where

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{Y}_{\_} = \begin{pmatrix} 1 & y_{d+1-1} - y_1 & y_{d+1-2} - y_2 \\ 1 & y_{d+2-1} - y_1 & y_{d+2-2} - y_2 \\ \vdots & \vdots & \vdots \\ 1 & y_{d+n-1} - y_1 & y_{d+n-2} - y_2 \end{pmatrix}, \quad \tilde{\mathbf{R}} = \begin{pmatrix} \tilde{r}_1 \\ \tilde{r}_2 \\ \vdots \\ \tilde{r}_n \end{pmatrix} = \begin{pmatrix} \tilde{y}_{d+1} & \tilde{\varepsilon}_{d+1}^2 \\ \tilde{y}_{d+2} & \tilde{\varepsilon}_{d+2}^2 \\ \vdots & \vdots \\ \tilde{y}_{d+n} & \tilde{\varepsilon}_{d+n}^2 \end{pmatrix}$$

$\mathbf{K} = diag\{K_h(y_{d+l-1} - y_1)K_h(y_{d+m-1} - y_1)\}$, $l, m = 1, \ldots, n$, $\mathbf{W} = diag\{\hat{W}_{d+k}\}$, $k = 1, \ldots, n$ and $\tilde{\varepsilon}_t^2 = (y_t - \tilde{m}(\mathbf{y}_t))^2$, $\tilde{y}_t = y_t - \sum_{\alpha=1}^d \hat{m}_\alpha(y_{t-\alpha})$ and $\tilde{\varepsilon}_t^2 = (y_t - \tilde{m}(\mathbf{y}_t))^2 - \sum_{\alpha=1}^d \hat{v}_\alpha(y_{t-\alpha})$. The preliminary density estimators used are

$$\hat{p}_{12}(y_1, y_2) = \frac{1}{ng^2}\sum_{t=d+1}^{d+n} L_g(y_{t-1} - y_1)L_g(y_{t-2} - y_2)$$

$$\hat{p}_3(\mathbf{y}_{\underline{12}}) = \frac{1}{ng^{d-2}}\sum_{t=d+1}^{d+n}\prod_{\alpha=3}^{d} L_g(y_{t-\alpha} - y_\alpha)$$

$$\hat{p}(\mathbf{y}) = \frac{1}{ng^d}\sum_{t=d+1}^{d+n}\prod_{\alpha=1}^{d} L_g(y_{t-\alpha} - y_\alpha)$$

$$\hat{W}_t = \frac{\hat{p}_{12}(y_{t-1}, y_{t-2})\hat{p}_3(\mathbf{y}_{t,\underline{12}})}{\hat{p}(\mathbf{y}_t)}$$

In exactly the same way as in Theorem 1, we can decompose the estimation error as

$$\hat{\phi}_{12}(y_1, y_2) - \phi_{12}(y_1, y_2) = [I_2 \otimes e_1^T Q_n^{-1}]\tau_n \tag{2}$$

where

$$Q_n = \frac{1}{n} D_h^{-1} \mathbf{Y}_{\text{-}}^T \mathbf{K} \mathbf{Y}_{\text{-}} D_h^{-1} \tag{3}$$

$$\tau_n = \frac{1}{n} [I_2 \otimes D_h^{-1} \mathbf{Y}_{\text{-}}^T \mathbf{K}] vec[\mathbf{W}\tilde{\mathbf{R}} - l\phi_{12}^T(y_1, y_2) - Y_{\text{-}} \bigtriangledown \phi_{12}^T(y_1, y_2)] \tag{4}$$

where $D_h = diag\{1,\ h,\ h\}$.

Step II. Approximations

τ_n from (4) can be represented as

$$\tau_n = \frac{1}{n} \sum_{k=d+1}^{d+n} \{K_h(y_{k-1} - y_1) K_h(y_{k-2} - y_2) e_k \otimes \left(1, \frac{y_{k-1} - y_1}{h}, \frac{y_{k-2} - y_2}{h}\right)^T \}$$

where

$$e_k = \hat{W}_k \tilde{r}_k - \phi_{12}(y_1, y_2) - (y_{k-1} - y_1) \frac{\partial \phi_{12}}{\partial y_{t-1}}(y_1, y_2) - (y_{k-2} - y_2) \frac{\partial \phi_{12}}{\partial y_{t-2}}(y_1, y_2)$$

Denote $K_h(y_{k-1} - y_1) K_h(y_{k-2} - y_2)$ by $KK_h(k)$, and $\left(1, \frac{y_{k-1}-y_1}{h}, \frac{y_{k-2}-y_2}{h}\right)^T$ by $X(k)$; then τ_n can be represented as

$$\tau_n = \frac{1}{n} \sum_{k=d+1}^{d+n} \{KK_h(k)[\hat{W}_k \tilde{r}_k - \phi_{12}(y_{k-1}, y_{k-2})] \otimes X(k)\} \tag{5}$$

$$+ \frac{1}{n} \sum_{k=d+1}^{d+n} \{KK_h(k) e_k^* \otimes X(k)\} \tag{6}$$

where

$$e_k^* = \phi_{12}(y_{k-1}, y_{k-2}) - \phi_{12}(y_1, y_2) - (y_{k-1} - y_1) \frac{\partial \phi_{12}}{\partial y_{t-1}}(y_1, y_2) - (y_{k-2} - y_2) \frac{\partial \phi_{12}}{\partial y_{t-2}}(y_1, y_2)$$

Define $r_k = r(\mathbf{y}_k) = \sum_{1 \le \alpha < \beta \le d} (m_{\alpha\beta}(y_{k-\alpha}, y_{k-\beta}),\ v_{\alpha\beta}(y_{k-\alpha}, y_{k-\beta}))^T = (r_k^m,\ r_k^v)^T$, Let us denote

$$\tilde{r}_k^* = \begin{pmatrix} \sum_{\alpha=1}^d [\hat{m}_\alpha(y_{k-\alpha}) - m_\alpha(y_{k-\alpha})] + v^{1/2}(\mathbf{y}_k)\varepsilon_k \\ \sum_{\alpha=1}^d [\hat{v}_\alpha(y_{k-\alpha}) - v_\alpha(y_{k-\alpha})] + v(\mathbf{y}_k)(\varepsilon_k^2 - 1) \end{pmatrix}$$

$$\begin{aligned} D_Y^2 &= \left(\frac{y_{k-1} - y_1}{h}\right)^2 \frac{\partial^2 \phi_{12}}{\partial y_{t-1}^2}(y_1, y_2) + \left(\frac{y_{k-2} - y_2}{h}\right)^2 \frac{\partial^2 \phi_{12}}{\partial y_{t-2}^2}(y_1, y_2) \\ &\quad + 2 \frac{(y_{k-1} - y_1)(y_{k-2} - y_2)}{h^2} \frac{\partial^2 \phi_{12}}{\partial y_{t-1} \partial y_{t-2}}(y_1, y_2) \\ &\equiv (D_Y^{2,m},\ D_Y^{2,v})^T \end{aligned}$$

The two components of τ_n can now be approximated in the same way as it was done for the two terms of (**??**) in the proof of Theorem 1; specifically, (5) becomes

$$\frac{1}{n}\sum_{k=d+1}^{d+n}\left\{KK_h(k)\left[\frac{p_{12}(y_{k-1},y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}r_k-\phi_{12}(y_{k-1},y_{k-2})\right]\otimes X(k)\right\}$$
$$+\frac{1}{n}\sum_{k=d+1}^{d+n}\left\{KK_h(k)\frac{p_{12}(y_{k-1},y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}\tilde{r}_k^*\otimes X(k)\right\}$$
$$+\frac{1}{n}\sum_{k=d+1}^{d+n}\left\{KK_h(k)\left[\frac{\hat{p}_{12}(y_{k-1},y_{k-2})\hat{p}_3(\mathbf{y}_{k,\underline{12}})}{\hat{p}(\mathbf{y}_k)}-\frac{p_{12}(y_{k-1},y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}\right]\tilde{r}_k\otimes X(k)\right\}$$
$$+\,o_p(1/\sqrt{nh^2})$$

while (6) becomes $\frac{h^2}{2n}\sum_{k=d+1}^{d+n}\{KK_h(k)D_Y^2\otimes X(k)\}+o_p(1/\sqrt{nh^2})$. Thus, the four components that need to be analyzed in order to obtain the asymptotic distribution and the bias of the estimator $\hat{\phi}_{12}(y_1,y_2)$ are

$$R_{1n}=\frac{h^2}{2n}\sum_{k=d+1}^{d+n}\{KK_h(k)D_Y^2\otimes X(k)\}$$
$$R_{2n}=\frac{1}{n}\sum_{k=d+1}^{d+n}\left\{KK_h(k)\left[\frac{\hat{p}_{12}(y_{k-1},y_{k-2})\hat{p}_3(\mathbf{y}_{k,\underline{12}})}{\hat{p}(\mathbf{y}_k)}-\frac{p_{12}(y_{k-1},y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}\right]\tilde{r}_k\otimes X(k)\right\}$$
$$R_{3n}=\frac{1}{n}\sum_{k=d+1}^{d+n}\{KK_h(k)\frac{p_{12}(y_{k-1},y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}\tilde{r}_k^*\otimes X(k)\}$$
$$R_{4n}=\frac{1}{n}\sum_{k=d+1}^{d+n}\left\{KK_h(k)\left[\frac{p_{12}(y_{k-1},y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}r_k-\phi_{12}(y_{k-1},y_{k-2})\right]\otimes X(k)\right\}$$

Step III. Asymptotic Bias

Applying ergodic theorem for α-mixing stationary processes to R_{1n} and using the Taylor expansion of $p_{12}(\cdot,\cdot)$, we obtain

$$
\begin{aligned}
R_{1n} &= \frac{h^2}{2n}\sum_{k=d+1}^{d+n}\{KK_h(k)D_Y^2 \otimes X(k)\} \\
&\xrightarrow{p} \frac{h^2}{2}E\left[K_h(y_{k-1}-y_1)K_h(y_{k-2}-y_2)\left(\frac{y_{k-1}-y_1}{h}\right)^2 \frac{\partial^2\phi_{12}}{\partial y_{t-1}^2}(y_1,y_2)\otimes\left(1,\frac{y_{k-1}-y_1}{h},\frac{y_{k-2}-y_2}{h}\right)^T\right] \\
&\quad + \frac{h^2}{2}E\left[K_h(y_{k-1}-y_1)K_h(y_{k-2}-y_2)\left(\frac{y_{k-2}-y_2}{h}\right)^2 \frac{\partial^2\phi_{12}}{\partial y_{t-2}^2}(y_1,y_2)\otimes\left(1,\frac{y_{k-1}-y_1}{h},\frac{y_{k-2}-y_2}{h}\right)^T\right] \\
&\quad + \frac{h^2}{2}E\left[K_h(y_{k-1}-y_1)K_h(y_{k-2}-y_2)\frac{2(y_{k-1}-y_1)(y_{k-2}-y_2)}{h^2}\frac{\partial^2\phi_{12}}{\partial y_{t-1}\partial y_{t-2}}(y_1,y_2)\right. \\
&\qquad\qquad \left.\otimes\left(1,\frac{y_{k-1}-y_1}{h},\frac{y_{k-2}-y_2}{h}\right)^T\right] \\
&= \frac{h^2}{2}p_{12}(y_1,y_2)\left[\frac{\partial^2\phi_{12}}{\partial y_{t-1}^2}(y_1,y_2)\otimes\left(\mu_K^2,\ \mu_K^3,0\right)^T + \frac{\partial^2\phi_{12}}{\partial y_{t-2}^2}(y_1,y_2)\otimes\left(\mu_K^2,\ 0,\ \mu_K^3\right)^T\right] + o(h^2)
\end{aligned}
$$

In the same way as before, we denote marginal expectations of estimated density functions $\hat{p}_{12}(\cdot)$, $\hat{p}_3(\cdot)$ and $\hat{p}(\cdot)$

$$
\begin{aligned}
\bar{p}_{12}(y_{k-1},y_{k-2}) &= \int L_g(z_1-y_{k-1})L_g(z_2-y_{k-2})p_{12}(z_1,z_2)dz_1dz_2 \\
\bar{p}_3(\mathbf{y}_{k,\underline{12}}) &= \int L_g(z_3-\mathbf{y}_{k,\underline{12}})p_3(z_3)dz_3 \\
\bar{p}_(\mathbf{y}_k) &= \int L_g(z_1-y_{k-1})L_g(z_2-y_{k-2})L_g(z_3-\mathbf{y}_{k,\underline{12}})p(z_1,z_2,z_3)dz_1dz_2dz_3
\end{aligned}
$$

Using the fact that

$$
\frac{1}{n}\sum_{k=d+1}^{d+n}\left\{KK_h(k)\left[\frac{\hat{p}_{12}(y_{k-1},y_{k-2})\hat{p}_3(\mathbf{y}_{k,\underline{12}})}{\hat{p}(\mathbf{y}_k)} - \frac{\bar{p}_{12}(y_{k-1},y_{k-2})\bar{p}_3(\mathbf{y}_{k,\underline{12}})}{\bar{p}(\mathbf{y}_k)}\right]\tilde{r}_k\otimes X(k)\right\} = o_p(1/\sqrt{nh^2})
$$

we can approximate R_{2n} as

$$
\begin{aligned}
R_{2n} &= \frac{1}{n}\sum_{k=d+1}^{d+n}\left\{KK_h(k)\left[\frac{\hat{p}_{12}(y_{k-1},y_{k-2})\hat{p}_3(\mathbf{y}_{k,\underline{12}})}{\hat{p}(\mathbf{y}_k)} - \frac{p_{12}(y_{k-1},y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}\right]\tilde{r}_k\otimes X(k)\right\} \\
&= \frac{1}{n}\sum_{k=d+1}^{d+n}\left\{KK_h(k)\left[\frac{\bar{p}_{12}(y_{k-1},y_{k-2})\bar{p}_3(\mathbf{y}_{k,\underline{12}})}{\bar{p}(\mathbf{y}_k)} - \frac{p_{12}(y_{k-1},y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}\right]\tilde{r}_k\otimes X(k)\right\} \\
&\quad + o_p(1/\sqrt{nh^2})
\end{aligned}
$$

Next, we use yet again the uniform convergence result from Masry (1996) in combination with ergodic theorem for α-mixing stationary processes from Fan and Yao (2003); it allows us to conclude that $KK_h(k)\left[\frac{\bar p_{12}(y_{k-1},y_{k-2})\bar p_3(\mathbf{y}_{k,\underline{12}})}{\bar p(\mathbf{y}_k)}-\frac{p_{12}(y_{k-1},y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}\right][\tilde m(\mathbf{y}_k)-m(\mathbf{y}_k)]^2\otimes X(k)=o_p(1/\sqrt{nh^2})$ and the left-hand side has finite expectation. Thus,

$$
\begin{aligned}
R_{2n} &= \frac{1}{n}\sum_{k=d+1}^{d+n}\{KK_h(k)[\frac{\bar p_{12}(y_{k-1},y_{k-2})\bar p_3(\mathbf{y}_{k,\underline{12}})}{\bar p(\mathbf{y}_k)}-\frac{p_{12}(y_{k-1},y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}]\tilde r_k\otimes X(k)\}+o_p(1/\sqrt{nh^2})\\
&\xrightarrow{p} E\left[KK_h(k)[\frac{\bar p_{12}(y_{k-1},y_{k-2})\bar p_3(\mathbf{y}_{k,\underline{12}})}{\bar p(\mathbf{y}_k)}-\frac{p_{12}(y_{k-1},y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}]\tilde r_k\otimes X(k)\right]\\
&= E\left[E\left[KK_h(k)[\frac{\bar p_{12}(y_{k-1},y_{k-2})\bar p_3(\mathbf{y}_{k,\underline{12}})}{\bar p(\mathbf{y}_k)}-\frac{p_{12}(y_{k-1},y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}]\tilde r_k\otimes X(k)|\mathcal{F}_k\right]\right]\\
&\simeq E\Bigg[KK_h(k)[\frac{\bar p_{12}(y_{k-1},y_{k-2})\bar p_3(\mathbf{y}_{k,\underline{12}})}{\bar p(\mathbf{y}_k)}-\frac{p_{12}(y_{k-1},y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}]\\
&\qquad\cdot\left(\sum_{\alpha=1}^{d}[m_\alpha(y_{k-\alpha})-\hat m_\alpha(y_{k-\alpha})]+r_k^m,\ \sum_{\alpha=1}^{d}[v_\alpha(y_{k-\alpha})-\hat v_\alpha(y_{k-\alpha})]+r_k^v\right)^T\otimes X(k)\Bigg]
\end{aligned}
$$

$$
\begin{aligned}
&\simeq -\int K_h(z_1-y_1)K_h(z_2-y_2)[\frac{\bar p_{12}(z_1,z_2)\bar p_3(z_3)}{p(z)}-\frac{p_{12}(z_1,z_2)p_3(z_3)\bar p(z)}{p^2(z)}]\left(\sum_{\alpha=1}^{d}b_\alpha^m(z_\alpha),\sum_{\alpha=1}^{d}b_\alpha^v(z_\alpha)\right)^T\\
&\qquad\otimes\left(1,\frac{z_1-y_1}{h},\frac{z_2-y_2}{h}\right)^T p(z)dz\\
&\quad+\int K_h(z_1-y_1)K_h(z_2-y_2)[\frac{\bar p_{12}(z_1,z_2)\bar p_3(z_3)}{p(z)}-\frac{p_{12}(z_1,z_2)p_3(z_3)\bar p(z)}{p^2(z)}]r(z)\\
&\qquad\otimes\left(1,\frac{z_1-y_1}{h},\frac{z_2-y_2}{h}\right)^T p(z)dz\\
&\simeq -\int K_h(z_1-y_1)K_h(z_2-y_2)\Big[p_{12}(z_1,z_2)[\bar p_3(z_3)-p_3(z_3)]+p_3(z_3)[\bar p_{12}(z_1,z_2)-p_{12}(z_1,z_2)]\\
&\qquad-\frac{p_{12}(z_1,z_2)p_3(z_3)}{p(z)}[\bar p(z)-p(z)]\Big]\left(\sum_{\alpha=1}^{d}b_\alpha^m(z_\alpha),\sum_{\alpha=1}^{d}b_\alpha^v(z_\alpha)\right)^T\otimes\left(1,\frac{z_1-y_1}{h},\frac{z_2-y_2}{h}\right)^T dz\\
&\quad+\int K_h(z_1-y_1)K_h(z_2-y_2)\Big[p_{12}(z_1,z_2)[\bar p_3(z_3)-p_3(z_3)]+p_3(z_3)[\bar p_{12}(z_1,z_2)-p_{12}(z_1,z_2)]\\
&\qquad-\frac{p_{12}(z_1,z_2)p_3(z_3)}{p(z)}[\bar p(z)-p(z)]\Big]r(z)\otimes\left(1,\frac{z_1-y_1}{h},\frac{z_2-y_2}{h}\right)^T dz\\
&= \frac{g^2}{2}p_{12}(y_1,y_2)\times\\
&\qquad\times\int\left[p_3^{(2)}(z_3)+\frac{p_{12}^{(2)}(y_1,y_2)}{p_{12}(y_1,y_2)}p_3(z_3)-\frac{p_3(z_3)}{p(y_1,y_2,z_3)}p^{(2)}(y_1,y_2,z_3)\right]\Delta_{\underline{12}}(y_1,y_2,z_3)dz_3\otimes(\mu_L^2,0,0)^T
\end{aligned}
$$

where

$$\Delta_{\underline{12}}(y_1, y_2, z_3) = \begin{pmatrix} m_{12}(y_1,y_2) - b_1^m(y_1) - b_2^m(y_2) + \sum_{3\le\alpha<\beta\le d} m_{\alpha\beta}(z_\alpha, z_\beta) - \sum_{\alpha=3}^d b_\alpha^m(z_\alpha) \\ v_{12}(y_1,y_2) - b_1^v(y_1) - b_2^v(y_2) + \sum_{3\le\alpha<\beta\le d} v_{\alpha\beta}(z_\alpha, z_\beta) - \sum_{\alpha=3}^d b_\alpha^v(z_\alpha) \end{pmatrix}$$

The term R_{3n} is easy to analyze; again, we apply the ergodic theorem for α-mixing stationary processes on R_{3n} to obtain

$$\begin{aligned} R_{3n} &= \frac{1}{n}\sum_{k=d+1}^{d+n}\left\{KK_h(k)\frac{p_{12}(y_{k-1}, y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}\tilde{r}_k^* \otimes X(k)\right\} \\ &\xrightarrow{p} E\left[KK_h(k)\frac{p_{12}(y_{k-1}, y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}\tilde{r}_k^* \otimes X(k)\right] \\ &= p_{12}(y_1, y_2)\int p_3(z_3)\Delta_{12}(y_1, y_2, z_3)dz_3 \otimes (1,0,0)^T \end{aligned}$$

where

$$\Delta_{12}(y_1, y_2, z_3) = \begin{pmatrix} b_1^m(y_1) + b_2^m(y_2) + \sum_{\alpha=3}^d b_\alpha^m(z_\alpha) \\ b_1^v(y_1) + b_2^v(y_2) + \sum_{\alpha=3}^d b_\alpha^v(z_\alpha) \end{pmatrix}$$

Finally, it is easy to realize that the term R_{4n} does not contribute to the asymptotic bias since it converges in probability to zero due to the same ergodic theorem

$$\begin{aligned} R_{4n} &= \frac{1}{n}\sum_{k=d+1}^{d+n}\left\{KK_h(k)\left[\frac{p_{12}(y_{k-1}, y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}r_k - \phi_{12}(y_{k-1}, y_{k-2})\right] \otimes X(k)\right\} \\ &\xrightarrow{p} (0\ 0\ 0\ 0\ 0\ 0)^T \end{aligned}$$

The analysis of $[I_2 \otimes e_1^T Q_n^{-1}]$ is almost exactly the same in Theorem 1; it is easy to show that,

$$\begin{aligned} Q_n &\xrightarrow{p} \begin{pmatrix} p_{12}(y_1,y_2) & 0 & 0 \\ 0 & p_{12}(y_1,y_2)\mu_K^2 & 0 \\ 0 & 0 & p_{12}(y_1,y_2)\mu_K^2 \end{pmatrix} \\ Q_n^{-1} &\xrightarrow{p} \begin{pmatrix} \frac{1}{p_{12}(y_1,y_2)} & 0 & 0 \\ 0 & \frac{1}{p_{12}(y_1,y_2)\mu_K^2} & 0 \\ 0 & 0 & \frac{1}{p_{12}(y_1,y_2)\mu_K^2} \end{pmatrix} \\ e_1^T Q_n^{-1} &\xrightarrow{p} \frac{1}{p_{12}(y_1,y_2)}e_1^T \end{aligned}$$

Therefore, using Slutzky's theorem we conclude that the bias of our estimator is

$$
\begin{aligned}
&B_{12,n}(y_1,y_2) = [I_2 \otimes e_1^T Q_n^{-1}](R_{1n} + R_{2n} + R_{3n} + R_{4n}) \\
&\xrightarrow{p} \left[I_2 \otimes \begin{pmatrix} \frac{1}{p_{12}(y_1,y_2)} & 0 & 0 \end{pmatrix} \right] \frac{h^2}{2} p_{12}(y_1,y_2) \Big[\frac{\partial^2 \phi_{12}}{\partial y_{t-1}^2}(y_1,y_2) \otimes \left(\mu_K^2,\ \mu_K^3, 0\right)^T \\
&+ \frac{\partial^2 \phi_{12}}{\partial y_{t-2}^2}(y_1,y_2) \otimes \left(\mu_K^2,\ 0,\ \mu_K^3\right)^T \Big] \\
&\quad + \left[I_2 \otimes \begin{pmatrix} \frac{1}{p_{12}(y_1,y_2)} & 0 & 0 \end{pmatrix} \right] \frac{g^2}{2} p_{12}(y_1,y_2) \\
&\quad \cdot \int \left[p_3^{(2)}(z_3) + \frac{p_{12}^{(2)}(y_1,y_2)}{p_{12}(y_1,y_2)} p_3(z_3) - \frac{p_3(z_3)}{p(y_1,y_2,z_3)} p^{(2)}(y_1,y_2,z_3) \right] \Delta_{\underline{12}}(y_1,y_2,z_3) dz_3 \otimes (\mu_L^2, 0, 0)^T \\
&\quad + \left[I_2 \otimes \begin{pmatrix} \frac{1}{p_{12}(y_1,y_2)} & 0 & 0 \end{pmatrix} \right] p_{12}(y_1,y_2) \\
&\quad \cdot \int p_3(z_3) \Delta_{12}(y_1,y_2,z_3) dz_3 \otimes (1,0,0)^T \\
&= \frac{h^2}{2} \mu_K^2 \left[\frac{\partial^2 \phi_{12}}{\partial y_{t-1}^2}(y_1,y_2) + \frac{\partial^2 \phi_{12}}{\partial y_{t-2}^2}(y_1,y_2) \right] \\
&\quad + \frac{g^2}{2} \mu_L^2 \int \left[p_3^{(2)}(z_3) + \frac{p_{12}^{(2)}(y_1,y_2)}{p_{12}(y_1,y_2)} p_3(z_3) - \frac{p_3(z_3)}{p(y_1,y_2,z_3)} p^{(2)}(y_1,y_2,z_3) \right] \Delta_{\underline{12}}(y_1,y_2,z_3) dz_3 \\
&\quad + \int p_3(z_3) \Delta_{12}(y_1,y_2,z_3) dz_3 \\
&\equiv B_{12}(y_1,y_2)
\end{aligned}
$$

Step IV. Asymptotic Normality

Finally, we derive the asymptotic distribution of $\sqrt{nh^2}[I_2 \otimes e_1^T Q_n^{-1}](R_{1n}+R_{2n}+R_{3n}+R_{4n})$. Using the same technics as the ones used in Theorem 1 to treat the terms T_{1n} and T_{2n}, it is easy to show that both $var(\sqrt{nh^2}R_{1n})$ and $var(\sqrt{nh^2}R_{2n})$ are negligible, with orders of $O(h^4)$ and $O\left(\frac{g^{2\nu}}{h^2}\right)$ respectively.

To analyze R_{3n} we expand it as

$$
\begin{aligned}
R_{3n} &= \frac{1}{n} \sum_{k=d+1}^{d+n} \{KK_h(k) \frac{p_{12}(y_{k-1}, y_{k-2}) p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)} \tilde{r}_k^* \otimes X(k)\} \\
&= \frac{1}{n} \sum_{k=d+1}^{d+n} \{KK_h(k) \frac{p_{12}(y_{k-1}, y_{k-2}) p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)} \tilde{r}_{k,1}^* \otimes X(k)\} \\
&\quad + \frac{1}{n} \sum_{k=d+1}^{d+n} \{KK_h(k) \frac{p_{12}(y_{k-1}, y_{k-2}) p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)} \tilde{r}_{k,2}^* \otimes X(k)\} \\
&\equiv R_{3n,1} + R_{3n,2}
\end{aligned}
$$

where

$$\tilde{r}^*_{k,1} = \begin{pmatrix} \sum_{\alpha=1}^{d}[\hat{m}_\alpha(y_{k-\alpha}) - m_\alpha(y_{k-\alpha})] \\ \sum_{\alpha=1}^{d}[\hat{v}_\alpha(y_{k-\alpha}) - v_\alpha(y_{k-\alpha})] \end{pmatrix}$$

$$\tilde{r}^*_{k,2} = \begin{pmatrix} v^{1/2}(\mathbf{y}_k)\varepsilon_k \\ v(\mathbf{y}_k)(\varepsilon_k^2 - 1) \end{pmatrix}$$

We claim that $var(\sqrt{nh^2}R_{3n,1})$ is negligible too. To see why, rewrite $R_{3n,1}$ as

$$R_{3n,1} = (x_{1n},\ x_{2n},\ x_{3n},\ x_{4n},\ x_{5n},\ x_{6n})^T$$

where

$$x_{1n} = \frac{1}{n}\sum_{k=d+1}^{d+n} K_h(y_{k-1} - y_1)K_h(y_{k-2} - y_2)\frac{p_{12}(y_{k-1}, y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}\sum_{\alpha=1}^{d}[\hat{m}_\alpha(y_{k-\alpha}) - m_\alpha(y_{k-\alpha})]$$

$$x_{2n} = \frac{1}{n}\sum_{k=d+1}^{d+n} K_h(y_{k-1} - y_1)K_h(y_{k-2} - y_2)\left(\frac{y_{k-1} - y_1}{h}\right)$$
$$\cdot\frac{p_{12}(y_{k-1}, y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}\sum_{\alpha=1}^{d}[\hat{m}_\alpha(y_{k-\alpha}) - m_\alpha(y_{k-\alpha})]$$

$$x_{3n} = \frac{1}{n}\sum_{k=d+1}^{d+n} K_h(y_{k-1} - y_1)K_h(y_{k-2} - y_2)\left(\frac{y_{k-2} - y_2}{h}\right)$$
$$\cdot\frac{p_{12}(y_{k-1}, y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}\sum_{\alpha=1}^{d}[\hat{m}_\alpha(y_{k-\alpha}) - m_\alpha(y_{k-\alpha})]$$

$$x_{4n} = \frac{1}{n}\sum_{k=d+1}^{d+n} K_h(y_{k-1} - y_1)K_h(y_{k-2} - y_2)\frac{p_{12}(y_{k-1}, y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}\sum_{\alpha=1}^{d}[\hat{v}_\alpha(y_{k-\alpha}) - v_\alpha(y_{k-\alpha})]$$

$$x_{5n} = \frac{1}{n}\sum_{k=d+1}^{d+n} K_h(y_{k-1} - y_1)K_h(y_{k-2} - y_2)\left(\frac{y_{k-1} - y_1}{h}\right)$$
$$\cdot\frac{p_{12}(y_{k-1}, y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}\sum_{\alpha=1}^{d}[\hat{v}_\alpha(y_{k-\alpha}) - v_\alpha(y_{k-\alpha})]$$

$$x_{6n} = \frac{1}{n}\sum_{k=d+1}^{d+n} K_h(y_{k-1} - y_1)K_h(y_{k-2} - y_2)\left(\frac{y_{k-2} - y_2}{h}\right)$$
$$\cdot\frac{p_{12}(y_{k-1}, y_{k-2})p_3(\mathbf{y}_{k,\underline{12}})}{p(\mathbf{y}_k)}\sum_{\alpha=1}^{d}[\hat{v}_\alpha(y_{k-\alpha}) - v_\alpha(y_{k-\alpha})]$$

We claim that $var(\sqrt{nh^2}x_{in})$, $i = 1, \ldots, 6$ are all negligible, with orders of $O(h^2)$. This is due to Davydov's lemma and the fact that $\hat{m}_\alpha(y_{k-\alpha}) - m_\alpha(y_{k-\alpha}) = O_p(h^2)$, $\hat{v}_\alpha(y_{k-\alpha}) -$

$v_\alpha(y_{k-\alpha}) = O_p(h^2)$ uniformly over all possible values of $y_{k-\alpha}$, with the latter following directly from Theorem 1. $var(\sqrt{nh^2}R_{3n,2})$ can be analyzed in exactly the same way as $var(\sqrt{nh}T_{3n})$ in Theorem 1; the outcome is that

$$\sqrt{nh^2}R_{3n,2} \xrightarrow{d} N(0, \Sigma_{1,R3}) \tag{7}$$

where

$$\Sigma_{1,R3} = \begin{pmatrix} \sigma^m_{1,s}(y_1, y_2) & * & * & \sigma^{mv}_{1,s}(y_1, y_2) & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ \sigma^{mv}_{1,s}(y_1, y_2) & * & * & \sigma^v_{1,s}(y_1, y_2) & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \end{pmatrix} \tag{8}$$

with

$$\sigma^m_{1,s}(y_1, y_2) =\| K \|_2^2 \, p^2_{12}(y_1, y_2) \int \frac{p_3^2(z_3)}{p(y_1, y_2, z_3)} v(y_1, y_2, z_3) dz_3 \tag{9}$$

$$\sigma^v_{1,s}(y_1, y_2) =\| K \|_2^2 \, p^2_{12}(y_1, y_2) \int \frac{p_3^2(z_3)}{p(y_1, y_2, z_3)} v^2(y_1, y_2, z_3)\kappa_4(y_1, y_2, z_3) dz_3 \tag{10}$$

$$\sigma^{mv}_{1,s}(y_1, y_2) =\| K \|_2^2 \, p^2_{12}(y_1, y_2) \int \frac{p_3^2(z_3)}{p(y_1, y_2, z_3)} v^{3/2}(y_1, y_2, z_3)\kappa_3(y_1, y_2, z_3) dz_3 \tag{11}$$

Also, $var(\sqrt{nh^2}R_{4n})$ can again be handled in the same way as $var(\sqrt{nh}T_{4n})$ in Theorem 1; the result is also an asymptotic normality for R_{4n} :

$$\sqrt{nh^2}R_{4n} \xrightarrow{d} N(0, \Sigma_{1,R4}) \tag{12}$$

where

$$\Sigma_{1,R4} = \begin{pmatrix} \sigma^m_{1,t}(y_1, y_2) & * & * & \sigma^{mv}_{1,t}(y_1, y_2) & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ \sigma^{mv}_{1,t}(y_1, y_2) & * & * & \sigma^v_{1,t}(y_1, y_2) & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \end{pmatrix} \tag{13}$$

with

$$\sigma_{1,t}^{m}(y_1,y_2) = \| K \|_2^2 \, p_{12}^2(y_1,y_2) \int \left[p_3(z_3) m(y_1,y_2,z_3) - p_{3|1,2}(z_3|y_1,y_2) m_{12}(y_1,y_2) \right]^2 dz_3 \tag{14}$$

$$\sigma_{1,t}^{v}(y_1,y_2) = \| K \|_2^2 \, p_{12}^2(y_1,y_2) \int \left[p_3(z_3) v(y_1,y_2,z_3) - p_{3|1,2}(z_3|y_1,y_2) v_{12}(y_1,y_2) \right]^2 dz_3 \tag{15}$$

$$\begin{aligned} \sigma_{1,t}^{mv}(y_1,y_2) = \| K \|_2^2 \, p_{12}^2(y_1,y_2) \int & \left[p_3(z_3) m(y_1,y_2,z_3) - p_{3|1,2}(z_3|y_1,y_2) m_{12}(y_1,y_2) \right] \\ & \cdot \left[p_3(z_3) v(y_1,y_2,z_3) - p_{3|1,2}(z_3|y_1,y_2) v_{12}(y_1,y_2) \right] dz_3 \end{aligned} \tag{16}$$

In both (8) and (13) those covariance matrix elements that are not important for the eventual result are replaced with $*$. It is also easy to conclude that the covariance between $R_{3n,2}$ and R_{4n} is 0 since $cov(R_{3n,2}, R_{4n}) = E[R_{3n,2} R_{4n}] = E[E[R_{3n,2} R_{4n} | \mathcal{F}_{n+d}]] = 0$.

Combining the four terms and using Slutzky's theorem, we can find that

$$\sqrt{nh^2}[I_2 \otimes e_1^T Q_n^{-1}](R_{1n} + R_{2n} + R_{3n} + R_{4n}) \xrightarrow{d} N(0, \Sigma_{12}^*(y_1,y_2))$$

and, as a consequence,

$$\sqrt{nh^2}[\hat{\phi}_{12}(y_1,y_2) - \phi_{12}(y_1,y_2) - B_{12}(y_1,y_2)] \xrightarrow{d} N[0, \Sigma_{12}^*(y_1,y_2)]$$

where

$$B_{12}(y_1,y_2) = \begin{pmatrix} b_{12}^m(y_1,y_2) \\ b_{12}^v(y_1,y_2) \end{pmatrix}, \quad \Sigma_{12}^*(y_1,y_2) = \begin{pmatrix} \sigma_{12}^m(y_1,y_2) & \sigma_{12}^{mv}(y_1,y_2) \\ \sigma_{12}^{mv}(y_1,y_2) & \sigma_{12}^v(y_1,y_2) \end{pmatrix}$$

with

$$B_{12}(y_1,y_2) = \frac{h^2}{2}\mu_K^2\left[\frac{\partial^2\phi_{12}}{\partial y_{t-1}^2}(y_1,y_2)+\frac{\partial^2\phi_{12}}{\partial y_{t-2}^2}(y_1,y_2)\right]$$
$$+\frac{g^2}{2}\mu_L^2\int\left[p_3^{(2)}(z_3)+\frac{p_{12}^{(2)}(y_1,y_2)}{p_{12}(y_1,y_2)}p_3(z_3)-\frac{p_3(z_3)}{p(y_1,y_2,z_3)}p^{(2)}(y_1,y_2,z_3)\right]\Delta_{\underline{12}}(y_1,y_2,z_3)dz_3$$
$$+\int p_3(z_3)\Delta_{12}(y_1,y_2,z_3)dz_3$$
$$\sigma_{12}^m(y_1,y_2) = \| K \|_2^2\int\left\{\frac{p_3^2(z_3)}{p(y_1,y_2,z_3)}v(y_1,y_2,z_3)\right.$$
$$\left.+\left[p_3(z_3)m(y_1,y_2,z_3)-p_{3|1,2}(z_3|y_1,y_2)m_{12}(y_1,y_2)\right]^2\right\}dz_3$$
$$\sigma_{12}^v(y_1,y_2) = \| K \|_2^2\int\left\{\frac{p_3^2(z_3)}{p(y_1,y_2,z_3)}v^2(y_1,y_2,z_3)\kappa_4(y_1,y_2,z_3)\right.$$
$$\left.+\left[p_3(z_3)v(y_1,y_2,z_3)-p_{3|1,2}(z_3|y_1,y_2)v_{12}(y_1,y_2)\right]^2\right\}dz_3$$
$$\sigma_{12}^{mv}(y_1,y_2) = \| K \|_2^2\int\left[p_3(z_3)m(y_1,y_2,z_3)-p_{3|1,2}(z_3|y_1,y_2)m_{12}(y_1,y_2)\right]$$
$$\times\left[p_3(z_3)v(y_1,y_2,z_3)-p_{3|1,2}(z_3|y_1,y_2)v_{12}(y_1,y_2)\right]$$
$$+\frac{p_3^2(z_3)}{p(y_1,y_2,z_3)}v^{3/2}(y_1,y_2,z_3)\kappa_3(y_1,y_2,z_3)dz_3$$

and

$$\kappa_3(y_1,y_2,z_3) \equiv E\left[\varepsilon_k^3|(y_{k-1},y_{k-2},\mathbf{y}_{k,3})=(y_1,y_2,z_3)\right],$$
$$\kappa_4(y_1,y_2,z_3) \equiv E\left[(\varepsilon_k^2-1)^2|(y_{k-1},y_{k-2},\mathbf{y}_{k,3})=(y_1,y_2,z_3)\right]$$