Simulated Nonparametric Estimation of Dynamic Models: Unpublished Appendix

Filippo Altissimo
European Central Bank, and CEPR

Antonio Mele
London School of Economics

May 5, 2008

Abstract

This appendix is the unabridged version of that appearing in Altissimo and Mele (2008) (hereafter Al-M). It provides:

1. Proofs of all the lemmas (in Section A).
2. Details of computations omitted in Al-M (in Sections B, C, D).
3. Details related to the SNE for continuous-time models (in Section E).
4. Examples relating identifiability to bandwidth choice, primitive conditions ensuring modulus of continuity assumptions in Al-M, and one example that illustrates analytically some properties of the Neyman Chi-Square measure of distance for dynamic models (in Section F).

Note that for the purpose of simplifying the exposition of this appendix, we provide the proofs for the J-SNE first, and for the CD-SNE after.

Contents

Remark on the notation. To simplify the notation of the present appendix, we shall denote the probability limits in eqs. (14) and (15) of Al-M as follows:

$$m_2\left(z,v;\theta\right)\equiv\pi_2^*\left(z,v;\theta,\bar{\lambda}\right),\quad m_1\left(z|\,v;\theta\right)\equiv\pi_1^*\left(v;\theta,\bar{\lambda}\right),\quad m\left(z|\,v;\theta\right)\equiv\pi^*\left(z|\,v;\theta,\bar{\lambda}\right),$$

and set, in eq. (15) of Al-M, $L^{\mathrm{CD}}\left(\theta\right)\equiv L^{\mathrm{CD}}(\theta,\bar{\lambda})$, and in Assumption 11 of Al-M, $L^{\mathrm{J}}\left(\theta\right)\equiv L^{\mathrm{J}}(\theta,\bar{\lambda})$.

A. Lemmas

This section collects the proofs of all the lemmas stated in the abridged appendix of Al-M. It also provides two remarks that explain how Assumptions 9 and 10 in Al-M imply that: (i) the bandwidth conditions in Lemma C1-C3 are satisfied, and (ii) all the suprema in Lemmas N1-N5 go to zero in probability. In Section A.1, we collect the proofs of the lemmas we use to prove consistency of our estimators. In Section A.2, we provide the proofs of the lemmas used to show asymptotic normality. In Section A.3, we discuss the bandwidth conditions in Assumptions 9 and 10 of Al-M.

A.1 Lemmas needed to prove consistency results

Lemma C0. *Let Assumptions 1-(a), 2-3 hold and for each t, let $x_t \equiv (z_t, v_t)$, as in the main text. We have:*

(a) $\sup_{x\in\mathbb{R}^q} |\pi_{2T}(x) - m_2(x;\theta_0)| = O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q}\right)$.

(b) $\sup_{x\in\mathbb{R}^q} |\pi_{2T}(x) - \pi_2(x;\theta_0)| = O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q}\right) + O_p(\lambda_T^r)$.

Proof. Part (a) and Part (b) of this lemma are special cases of Lemma A-2 (p. 588) and Theorem 1 (p. 568) in Andrews (1995). ■

Lemma C1. *Let Assumptions 1-(a), 2-3 hold, and for given $\theta \in \Theta$, set $\mathcal{B}_T \equiv \{v \in \mathbb{R}^{q-q^*} : \pi_{1T}(v) > \delta_T$ and $\pi_{1T}^i(v;\theta) > \delta_T$, $i = 1, \cdots, S\}$, where $\delta_T \to 0$ and $T^{\frac{1}{2}}\lambda_T^q\delta_T \to \infty$. We have:*

(a) *Let $\lambda_T \to \bar{\lambda}$, where $0 \le \bar{\lambda} < \infty$. Then, for all $\theta \in \Theta$,*

$$\sup_{(z,v)\in\mathbb{R}^{q^*}\times\mathcal{B}_T} \left|\frac{\pi_{2T}(z,v)}{\pi_{1T}(v)} - m(z|v;\theta_0)\right| \xrightarrow{p} 0; \tag{A1a}$$

$$\sup_{(z,v)\in\mathbb{R}^{q^*}\times\mathcal{B}_T} \left|\frac{\pi_{2T}^i(z,v;\theta)}{\pi_{1T}^i(v;\theta)} - m(z|v;\theta)\right| \xrightarrow{p} 0, \text{ for } i = 1, \cdots, S. \tag{A1b}$$

(b) *Let $\lambda_T \to 0$ and $\delta_T\lambda_T^{-r} \to \infty$. Then, for all $\theta \in \Theta$,*

$$\sup_{(z,v)\in\mathbb{R}^{q^*}\times\mathcal{B}_T} \left|\frac{\pi_{2T}(z,v)}{\pi_{1T}(v)} - \pi(z|v;\theta_0)\right| \xrightarrow{p} 0; \tag{A2a}$$

$$\sup_{(z,v)\in\mathbb{R}^{q^*}\times\mathcal{B}_T} \left|\frac{\pi_{2T}^i(z,v;\theta)}{\pi_{1T}^i(v;\theta)} - \pi(z|v;\theta)\right| \xrightarrow{p} 0, \text{ for } i = 1, \cdots, S. \tag{A2b}$$

Proof. We only provide the proof of the convergence results in (A1a) and (A2a), since the proof of (A1b) and (A2b) follows by nearly identical arguments. Moreover, in the proof of this and subsequent lemmas, we shall make a repeated use of the identity: $\frac{a}{b} - \frac{\bar{a}}{\bar{b}} = \frac{1}{b}(a-\bar{a}) - \frac{a}{b\bar{b}}\left(b-\bar{b}\right)$, where $a, b, \bar{a}$ and $\bar{b}$ are any four strictly positive functions.

(Part (a)) Let $\mathcal{B}_{1T}(\varepsilon) \equiv \{v \in \mathbb{R}^{q-q^*} : m_1(v;\theta_0) \geq \varepsilon\delta_T\}$, and for each $\theta \in \Theta$, $\mathcal{B}_{2T}(\varepsilon) \equiv \{v \in \mathbb{R}^{q-q^*} : \pi_{1T}(v) > \varepsilon\delta_T$ and $\pi^i_{1T}(v;\theta) \geq \varepsilon\delta_T$, $i = 1,\cdots,S\}$. Finally, let $\hat{\mathcal{B}}_T \equiv \hat{\mathcal{B}}_T(\varepsilon) \equiv \mathcal{B}_{1T}(\varepsilon) \cap \mathcal{B}_{2T}(\varepsilon)$, for some $\varepsilon > 0$. We have,

$$\begin{aligned}
&\sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T} |\pi_T(z|v) - m(z|v;\theta_0)| \\
&\leq \sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T} \left[\frac{1}{\pi_{1T}(v)} |\pi_{2T}(z,v) - m_2(z,v;\theta_0)|\right] \\
&+ \sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T} \left[\frac{m(z|v;\theta_0)}{\pi_{1T}(v)} |\pi_{1T}(v) - m_1(v;\theta_0)|\right] \\
&\leq \delta_T^{-1}\varepsilon^{-1} \sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T} [|\pi_{2T}(z,v) - m_2(z,v;\theta_0)|] + \delta_T^{-1}\varepsilon^{-1}c_0 \sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T} [|\pi_{1T}(v) - m_1(v;\theta_0)|] \\
&= O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q}\delta_T^{-1}\right) + O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\delta_T^{-1}\right), \qquad \text{(A3a)}
\end{aligned}$$

where $c_0 \equiv \sup_{(z,v)\in\mathbb{R}^{q^*}\times\mathbb{R}^{q-q^*}} m(z|v;\theta_0)$, and the last equality follows by Lemma C0-(a). We are left to show that eq. (A3a) holds when $\hat{\mathcal{B}}_T$ is replaced with the trimming set $\mathcal{B}_T = \mathcal{B}_{2T}(1)$. The argument is nearly identical to Andrews (1995, proof of Thm. 1, p. 588). We have $\hat{\mathcal{B}}_T \supseteq \hat{\mathcal{B}}^*_T \equiv \mathcal{B}_{1T}(\varepsilon) \cap \mathcal{B}_{2T}(2\varepsilon)$ and so eq. (A3a) holds with $\hat{\mathcal{B}}^*_T$ replacing $\hat{\mathcal{B}}_T$. Moreover, by Lemma C0-(a), and one argument similar to Andrews (1995, p. 588), $\mathcal{B}_{2T}(2\varepsilon) \subseteq \mathcal{B}_{1T}(\varepsilon)$ with probability (wp) 1 as $T \to \infty$. Therefore, eq. (A3a) holds with $\hat{\mathcal{B}}_T$ replaced by $\mathcal{B}_{2T}(2\varepsilon)$ wp 1 as $T \to \infty$, and the result follows by setting $\varepsilon = \frac{1}{2}$.

The proof of Part (b) is nearly identical. Define trimming sets $\mathcal{B}_T$, $\hat{\mathcal{B}}_T$, $\mathcal{B}_{1T}$ and $\mathcal{B}_{2T}$ as before, with the exception that the function $m_1(v;\theta_0)$ in $\mathcal{B}_{1T}$ is replaced $\pi_1(v;\theta_0)$. By Lemma C0-(b) and the same arguments leading to (A3a),

$$\begin{aligned}
\sup_{(z,v)\in\mathcal{B}_T} \left|\frac{\pi_{2T}(z,v)}{\pi_{1T}(v)} - \pi(z|v;\theta_0)\right| &= O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q}\delta_T^{-1}\right) + O_p\left(\delta_T^{-1}\lambda_T^r\right) \\
&+ O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\delta_T^{-1}\right) + O_p\left(\delta_T^{-1}\lambda_T^r\right). \qquad \text{(A3b)}
\end{aligned}$$

Part (b) then follows by replacing $\hat{\mathcal{B}}_T$ with the trimming set $\mathcal{B}_T$, exactly as in the proof of Part (a). ■

Lemma C2. *Let Assumptions 1-(a), 2-3 hold, and set* $\mathcal{A}_T \equiv \left\{(z,v) \in \mathbb{R}^{q^*} \times \mathbb{R}^{q-q^*} : \pi_{2T}(z,v) > \alpha_T\right\}$, *where* $\alpha_T \to 0$, $T^{\frac{1}{2}}\lambda_T^q \alpha_T^3 \to \infty$ *and* $\lambda_T^{q^*}\alpha_T \to 0$. *We have:*

(a) *Let* $\lambda_T \to \bar{\lambda}$, *where* $0 \leq \bar{\lambda} < \infty$; *then,*

$$\sup_{(z,v)\in\mathcal{A}_T} \left[\frac{1}{m_2(z,v;\theta_0)} \left|\frac{\pi_{1T}(v)}{\pi_T(z|v)} - \frac{m_1(v;\theta_0)}{m(z|v;\theta_0)}\right|\right] \xrightarrow{p} 0.$$

(b) *Let* $\lambda_T \to 0$ *and* $\alpha_T^3\lambda_T^{-r} \to \infty$; *then,*

$$\sup_{(z,v)\in\mathcal{A}_T} \left[\frac{1}{\pi_2(z,v;\theta_0)} \left|\frac{\pi_{1T}(v)}{\pi_T(z|v)} - \frac{\pi_1(v;\theta_0)}{\pi(z|v;\theta_0)}\right|\right] \xrightarrow{p} 0.$$

Proof. (Part (a)) Let $\mathcal{A}_{1T}(\varepsilon) \equiv \{(z,v) \in \mathbb{R}^{q^*} \times \mathbb{R}^{q-q^*} : m_2(z,v;\theta_0) \geq \varepsilon\alpha_T\}$, $\mathcal{A}_{2T}(\varepsilon) \equiv \{(z,v) \in \mathbb{R}^{q^*} \times \mathbb{R}^{q-q^*} : \pi_{2T}(z,v) \geq \varepsilon\alpha_T\}$ and $\hat{\mathcal{A}}_T \equiv \hat{\mathcal{A}}_T(\varepsilon) \equiv \mathcal{A}_{1T}(\varepsilon) \cap \mathcal{A}_{2T}(\varepsilon)$ for some $\varepsilon > 0$. We have:

$$\begin{aligned}\sup_{(z,v)\in\hat{\mathcal{A}}_T} \left|\frac{\pi_{1T}(v)}{\pi_T(z|v)} - \frac{m_1(v;\theta_0)}{m(z|v;\theta_0)}\right| \leq & \sup_{(z,v)\in\hat{\mathcal{A}}_T} \left[\frac{1}{m(z|v;\theta_0)} |\pi_{1T}(v) - m_1(v;\theta_0)|\right] \\ & + \sup_{(z,v)\in\hat{\mathcal{A}}_T} \left[\frac{\pi_{1T}(v)}{\pi_T(z|v)\, m(z|v;\theta_0)} |\pi_T(z|v) - m(z|v;\theta_0)|\right],\end{aligned}$$

and for all $(z,v) \in \hat{\mathcal{A}}_T$,

$$\begin{aligned}& \frac{\pi_{1T}(v)}{m(z|v;\theta_0)} |\pi_T(z|v) - m(z|v;\theta_0)| \\ \leq\ & \frac{\pi_{1T}(v)}{m_2(z,v;\theta_0)} |\pi_{2T}(z,v) - m_2(z,v;\theta_0)| + \frac{\pi_{2T}(z,v)}{m_2(z,v;\theta_0)} |\pi_{1T}(v) - m_1(v;\theta_0)|.\end{aligned}$$

Hence,

$$
\begin{aligned}
&\sup_{(z,v)\in\hat{\mathcal{A}}_T}\left[\frac{1}{m_2(z,v;\theta_0)}\left|\frac{\pi_{1T}(v)}{\pi_T(z|v)}-\frac{m_1(v;\theta_0)}{m(z|v;\theta_0)}\right|\right] \\
&\leq \sup_{(z,v)\in\hat{\mathcal{A}}_T}\left[\frac{m_1(v;\theta_0)+\pi_{1T}(v)}{m_2(z,v;\theta_0)^2}\left|\pi_{1T}(v)-m_1(v;\theta_0)\right|\right] \\
&+\sup_{(z,v)\in\hat{\mathcal{A}}_T}\left[\frac{\pi_{1T}(v)^2}{\pi_{2T}(z,v)\,m_2(z,v;\theta_0)^2}\left|\pi_{2T}(z,v)-m_2(z,v;\theta_0)\right|\right] \\
&\leq \sup_{(z,v)\in\hat{\mathcal{A}}_T}\left[2\frac{m_1(v;\theta_0)}{m_2(z,v;\theta_0)^2}\left|\pi_{1T}(v)-m_1(v;\theta_0)\right|\right] \\
&+\sup_{(z,v)\in\hat{\mathcal{A}}_T}\left[\frac{1}{m_2(z,v;\theta_0)^2}\left|\pi_{1T}(v)-m_1(v;\theta_0)\right|^2\right] \\
&+\sup_{(z,v)\in\hat{\mathcal{A}}_T}\left[\frac{m_1(v;\theta_0)^2}{\pi_{2T}(z,v)\,m_2(z,v;\theta_0)^2}\left|\pi_{2T}(z,v)-m_2(z,v;\theta_0)\right|\right] \\
&+\sup_{(z,v)\in\hat{\mathcal{A}}_T}\left[2\frac{m_1(v;\theta_0)}{\pi_{2T}(z,v)\,m_2(z,v;\theta_0)^2}\left|\pi_{1T}(v)-m_1(v;\theta_0)\right|\left|\pi_{2T}(z,v)-m_2(z,v;\theta_0)\right|\right] \\
&+\sup_{(z,v)\in\hat{\mathcal{A}}_T}\left[\frac{1}{\pi_{2T}(z,v)\,m_2(z,v;\theta_0)^2}\left|\pi_{1T}(v)-m_1(v;\theta_0)\right|^2\left|\pi_{2T}(z,v)-m_2(z,v;\theta_0)\right|\right] \\
&= O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\alpha_T^{-2}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q}\alpha_T^{-3}\right), \qquad \text{(A4)}
\end{aligned}
$$

where we have made use of Lemma C0-(a). Now suppose that $T^{\frac{1}{2}}\lambda_T^q\alpha_T^3=(T^{\frac{1}{2}}\lambda_T^{q-q^*}\alpha_T^2)(\lambda_T^{q^*}\alpha_T)\to\infty$. Since $\lambda_T^{q^*}\alpha_T\to 0$, then $T^{\frac{1}{2}}\lambda_T^{q-q^*}\alpha_T^2\to\infty$. We are left to show that eq. (A4) holds when substituting $\hat{\mathcal{A}}_T$ with the trimming set $\mathcal{A}_T=\mathcal{A}_{2T}(1)$. The argument is as in Lemma C1. We have $\hat{\mathcal{A}}_T\supseteq\hat{\mathcal{A}}_T^*\equiv\mathcal{A}_{1T}(\varepsilon)\cap\mathcal{A}_{2T}(2\varepsilon)$ and so eq. (A4) holds with $\hat{\mathcal{A}}_T^*$ replacing $\hat{\mathcal{A}}_T$. Moreover, $\mathcal{A}_{2T}(2\varepsilon)\subseteq\mathcal{A}_{1T}(\varepsilon)$ wp 1 as $T\to\infty$. Therefore, eq. (A4) holds with $\hat{\mathcal{A}}_T$ replaced by $\mathcal{A}_{2T}(2\varepsilon)$ wp 1 as $T\to\infty$, and the result follows by setting $\varepsilon=\frac{1}{2}$.

As for the proof of Part (b), define trimming sets $\mathcal{A}_T$, $\hat{\mathcal{A}}_T$, $\mathcal{A}_{1T}$ and $\mathcal{A}_{2T}$ as before, with the exception that $m_2(z,v;\theta_0)$ in $\mathcal{A}_{1T}$ is replaced with $\pi_2(z,v;\theta_0)$. By Lemma C0-(b) and the same arguments leading to (A4),

$$
\begin{aligned}
\sup_{(z,v)\in\hat{\mathcal{A}}_T}\left|\frac{\pi_{1T}(v)}{\pi_T(z|v)}-\frac{\pi_1(v;\theta_0)}{\pi(z|v;\theta_0)}\right| &= O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\alpha_T^{-1}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q}\alpha_T^{-2}\right) \\
&\quad +O_p\left(\alpha_T^{-1}\lambda_T^r\right)+O_p\left(\alpha_T^{-2}\lambda_T^r\right).
\end{aligned}
$$

Part (b) then follows by replacing $\hat{\mathcal{A}}_T$ with the trimming set $\mathcal{A}_T$, exactly as we did in Lemma C1 and in Part (a) of this lemma. ■

Lemma C3. *Let Assumptions 1-(a), 2-3 hold, and let* $\alpha_T \to 0$, $\delta_T \to 0$, $T^{\frac{1}{2}}\lambda_T^q \alpha_T^2 \delta_T \to \infty$ *and* $T^{\frac{1}{2}}\lambda_T^{q-q^*}\alpha_T^2\delta_T^2 \to \infty$. *We have:*

(a) *Let* $\lambda_T \to \bar{\lambda}$, *where* $0 \leq \bar{\lambda} < \infty$; *then, for each* $i = 1, \cdots, S$, *and* $\theta \in \Theta$,

$$\sup_{(z,v)\in\mathcal{A}_T\cap\mathcal{B}_T}\left[\frac{1}{m_2(z,v;\theta_0)\,m(z|\,v;\theta_0)}\left|\frac{\pi_{2T}^i(z,v;\theta)}{\pi_{1T}^i(v;\theta)} - m(z|\,v;\theta)\right|\right] \xrightarrow{p} 0.$$

(b) *Let* $\lambda_T \to 0$, $\alpha_T^2\delta_T\lambda_T^{-r} \to \infty$ *and* $\alpha_T^2\delta_T^2\lambda_T^{-r} \to \infty$; *then, for each* $i = 1, \cdots, S$, *and* $\theta \in \Theta$,

$$\sup_{(z,v)\in\mathcal{A}_T\cap\mathcal{B}_T}\left[\frac{1}{m_2(z,v;\theta_0)\,m(z|\,v;\theta_0)}\left|\frac{\pi_{2T}^i(z,v;\theta)}{\pi_{1T}^i(v;\theta)} - \pi(z|\,v;\theta)\right|\right] \xrightarrow{p} 0.$$

Proof. (Part (a)) The proof is only sketched as it is nearly identical to proofs in Lemmas C1-(a) and C2-(a). For each $i = 1, \cdots, S$, and $\theta \in \Theta$,

$$\begin{aligned}
&\sup_{(z,v)\in\hat{\mathcal{A}}_T\cap\hat{\mathcal{B}}_T}\left[\frac{m_1(v;\theta_0)}{m_2(z,v;\theta_0)^2}\left|\frac{\pi_{2T}^i(z,v;\theta)}{\pi_{1T}^i(v;\theta)} - m(z|\,v;\theta)\right|\right]\\
&\leq \sup_{(z,v)\in\hat{\mathcal{A}}_T\cap\hat{\mathcal{B}}_T}\left[\frac{m_1(v;\theta_0)}{m_2(z,v;\theta_0)^2\,m_1(v;\theta)}\left|\pi_{2T}^i(z,v;\theta) - m_2(z,v;\theta)\right|\right]\\
&+ \sup_{(z,v)\in\hat{\mathcal{A}}_T\cap\hat{\mathcal{B}}_T}\left[\frac{\pi_{2T}^i(z,v;\theta)\,m_1(v;\theta_0)}{m_2(z,v;\theta_0)^2\,m_1(v;\theta)\,\pi_{1T}^i(v;\theta)}\left|\pi_{1T}^i(v;\theta) - m_1(v;\theta)\right|\right]\\
&= O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q}\alpha_T^{-2}\delta_T^{-1}\right) + O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\alpha_T^{-2}\delta_T^{-2}\right),
\end{aligned}$$

where the last equality holds by Lemma C0-(a). Conclude as in the previous Lemmas C1 and C2. Part (b) is nearly identical given Lemma C0-(b). ■

A.2 Lemmas needed to prove asymptotic normality results

In the following lemmas, α_T and δ_T denote the same sequences introduced in the previous Lemmas C1 and C2.

Lemma N1. *Let Assumptions 1, 2-3 hold, let* $x \equiv [z \;\; v]$, *as in the main text, and let Assumption 7-(b) holds. Then, for all* $\theta \in \Theta$ *and* $j = 1, \cdots, n$,

(i) $$\sup_{x\in\mathbb{R}^q}\left|\nabla_{\theta_j}\pi_{2T,S}(x;\theta) - \nabla_{\theta_j}\pi_2(x;\theta)\right| = O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q-1}\right) + O_p\left(\lambda_T^r\right).$$

(ii) $$\begin{aligned}&\sup_{(z,v)\in\mathbb{R}^{q^*}\times\mathcal{B}_T}\left|\nabla_{\theta_j}\pi_{T,S}(z|\,v;\theta) - \nabla_{\theta_j}\pi(z|\,v;\theta)\right|\\
&= O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q-1}\delta_T^{-2}\right) + O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)-1}\delta_T^{-2}\right) + O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\delta_T^{-3}\right) + O_p\left(\lambda_T^r\delta_T^{-3}\right).\end{aligned}$$

Proof. (Part (a)). We have,

$$\sup_{x\in\mathbb{R}^q}\left|\nabla_{\theta_j}\pi_{2T,S}(x;\theta)-\nabla_{\theta_j}\pi_2(x;\theta)\right|$$
$$\le\frac{1}{S}\sum_{i=1}^{S}\sup_{x\in\mathbb{R}^q}\left|\nabla_{\theta_j}\pi_{2T}^i(x;\theta)-E[\nabla_{\theta_j}\pi_{2T}^i(x;\theta)]\right|+\frac{1}{S}\sum_{i=1}^{S}\sup_{x\in\mathbb{R}^q}\left|\nabla_{\theta_j}\pi_2(x;\theta)-E[\nabla_{\theta_j}\pi_{2T}^i(x;\theta)]\right|.$$

For each $i=1,\cdots,S$, and $\theta\in\Theta$,

$$\sup_{x\in\mathbb{R}^q}\left|\nabla_{\theta_j}\pi_{2T}^i(x;\theta)-E[\nabla_{\theta_j}\pi_{2T}^i(x;\theta)]\right|$$
$$=\sup_{x\in\mathbb{R}^q}\left|\frac{1}{T\lambda_T^{q+1}}\sum_{t=1+l}^{T}\frac{\partial x_t^i(\theta)}{\partial\theta_j}\cdot K_q'\left(\frac{x_t^i(\theta)-x}{\lambda_T}\right)-\frac{1}{\lambda_T^{q+1}}E\left[\frac{\partial x_\infty(\theta)}{\partial\theta_j}\cdot K_q'\left(\frac{x_\infty(\theta)-x}{\lambda_T}\right)\right]\right|$$
$$=O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q-1}\right),$$

where $x_\infty(\theta)$ is a value of $x=[z\ v]$ drawn from $\pi_2(z,v;\theta)$, and the second equality follows by Lemma A-2 (p. 588) in Andrews (1995) and the mixing condition in Assumption 7-(a). As for the bias term,

$$E[\nabla_{\theta_j}\pi_{2T}^i(x;\theta)]=\nabla_{\theta_j}E[\pi_{2T}^i(x;\theta)]$$
$$=\nabla_{\theta_j}\pi_2(x;\theta)+\frac{\lambda_T^r}{r!}\int\frac{\partial}{\partial\theta_j}\sum_{i_1,\cdots,i_r=1}^{q}\frac{\partial}{\partial x_{i_1}}\cdots\frac{\partial}{\partial x_{i_r}}\pi_2(x+\lambda_T^*z;\theta)\,z_{i_1}\cdots z_{i_r}K_q(z)\,dz,$$

where the first equality follows by dominated convergence, and $\lambda_T^*\in(0,\lambda_T)$. Part (a) of this lemma follows by uniform boundedness of $\partial^{r+1}\pi_2(x;\theta)/\partial\theta\partial x^r$.

We show that Part (b) of this lemma holds with the trimming set $\hat{\mathcal{B}}_T$. The extension to the set $\mathcal{B}_T$ follows by the same arguments in Lemma C1. We have,

$$\sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\left|\nabla_{\theta_j}\pi_{T,S}(z|v;\theta)-\nabla_{\theta_j}\pi(z|v;\theta)\right|\le\frac{1}{S}\sum_{i=1}^{S}\sup_{(z,v)\in\mathbb{R}^{q^*}\times B_T}\left|\nabla_{\theta_j}\pi_T^i(z|v;\theta)-\nabla_{\theta_j}\pi(z|v;\theta)\right|,$$

and

$$\sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\left|\nabla_{\theta_j}\pi_T^i(z|v;\theta)-\nabla_{\theta_j}\pi(z|v;\theta)\right|$$
$$\le\sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\left[\frac{1}{\pi_{1T}^i(v;\theta)}\left|\nabla_{\theta_j}\pi_{2T}^i(z,v;\theta)-\nabla_{\theta_j}\pi_2(z,v;\theta)\right|\right]$$
$$+\sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\left[\frac{\pi_{2T}^i(z,v;\theta)}{\pi_{1T}^i(v;\theta)^2}\left|\nabla_{\theta_j}\pi_{1T}^i(v;\theta)-\nabla_{\theta_j}\pi_1(v;\theta)\right|\right]$$
$$+\sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\left[\left|\nabla_{\theta_j}\pi_2(z,v;\theta)\right|\left|\frac{1}{\pi_{1T}^i(v;\theta)}-\frac{1}{\pi_1(v;\theta)}\right|+\left|\nabla_{\theta_j}\pi_1(v;\theta)\right|\left|\frac{\pi_{2T}^i(z,v;\theta)}{\pi_{1T}^i(v;\theta)^2}-\frac{\pi_2(z,v;\theta)}{\pi_1(v;\theta)^2}\right|\right]$$
$$\equiv T_1+T_2+T_3.$$

By Part (a) of this lemma, and boundedness of $\nabla_{\theta_j}\pi_2(z,v;\theta)$ (Assumption 1-(b)), we have that $T_1 = O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q-1}\delta_T^{-1}\right) + O_p\left(\lambda_T^r\delta_T^{-1}\right)$ and $T_2 = O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)-1}\delta_T^{-2}\right) + O_p\left(\lambda_T^r\delta_T^{-2}\right)$. As regards the T_3 term we have, by Lemma C0,

$$\begin{aligned}
&\sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\left|\frac{1}{\pi_{1T}^i(v;\theta)}-\frac{1}{\pi_1(v;\theta)}\right| \\
&\leq \sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\left[\frac{1}{\pi_{1T}^i(v;\theta)\,\pi_1(v;\theta)}\left|\pi_{1T}^i(v;\theta)-\pi_1(v;\theta)\right|\right] = O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\delta_T^{-2}\right)+O_p\left(\lambda_T^r\delta_T^{-2}\right),
\end{aligned}$$

and,

$$\begin{aligned}
&\sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\left|\frac{\pi_{2T}^i(z,v;\theta)}{\pi_{1T}^i(v;\theta)^2}-\frac{\pi_2(z,v;\theta)}{\pi_1(v;\theta)^2}\right| \\
&\leq \sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\left[\frac{1}{\pi_{1T}^i(v;\theta)^2}\left|\pi_{2T}^i(z,v;\theta)-\pi_2(z,v;\theta)\right|\right]+\sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\left|\frac{1}{\pi_{1T}^i(v;\theta)^2}-\frac{1}{\pi_1(v;\theta)^2}\right|\pi_2(z,v;\theta) \\
&\leq \sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\left[\frac{1}{\pi_{1T}^i(v;\theta)^2}\left|\pi_{2T}^i(z,v;\theta)-\pi_2(z,v;\theta)\right|\right]+\delta_T^{-3}\sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\left|\pi_{1T}^i(v;\theta)-\pi_1(v;\theta)\right|\pi_2(z,v;\theta) \\
&= O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q}\delta_T^{-2}\right)+O_p\left(\lambda_T^r\delta_T^{-2}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\delta_T^{-3}\right)+O_p\left(\lambda_T^r\delta_T^{-3}\right),
\end{aligned}$$

where the last equality follows by Lemma C0, and the second inequality holds because

$$\begin{aligned}
&\sup_{v\in\hat{\mathcal{B}}_T}\left|\frac{1}{\pi_{1T}^i(v;\theta)^2}-\frac{1}{\pi_1(v;\theta)^2}\right| \\
=\;&\sup_{v\in\hat{\mathcal{B}}_T}\left[\frac{1}{\pi_{1T}^i(v;\theta)\,\pi_1(v;\theta)}\left|\frac{\pi_1(v;\theta)}{\pi_{1T}^i(v;\theta)}-\frac{\pi_{1T}^i(v;\theta)}{\pi_1(v;\theta)}\right|\right] \\
\leq\;&\delta_T^{-2}\cdot\sup_{v\in\hat{\mathcal{B}}_T}\left[\frac{1}{\pi_1(v;\theta)}+\frac{1}{\pi_{1T}^i(v;\theta)}\right]\left|\pi_{1T}^i(v;\theta)-\pi_1(v;\theta)\right|
\end{aligned}$$

Hence, by boundedness of $\nabla_{\theta_j}\pi_2(z,v;\theta)$ and $\nabla_{\theta_j}\pi_1(v;\theta)$ (Assumption 1-(b)), we have that $T_3 = O_p\left(\lambda_T^r\delta_T^{-3}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q}\delta_T^{-2}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\delta_T^{-3}\right)$. ■

Lemma N2. *Let the assumptions in Lemma N1 and Assumption 7-(b) hold. Then, for all* $j=1,\cdots,n$,

$$\begin{aligned}
&\sup_{(z,v)\in\mathbb{R}^{q^*}\times\mathcal{B}_T}\left|\frac{\nabla_{\theta_j}\pi_{T,S}(z|\,v;\theta_0)w_T(z,v)}{\pi_{1T}(v)}-\frac{\nabla_{\theta_j}\pi(z|\,v;\theta_0)w(z,v)}{\pi_1(v;\theta_0)}\right| \\
=\;&O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q-1}\delta_T^{-3}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)-1}\delta_T^{-3}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\delta_T^{-4}\right)+O_p\left(\lambda_T^r\delta_T^{-4}\right).
\end{aligned}$$

Proof. We proceed as in the proof of Lemma N1, and demonstrate the result with the trimming set $\hat{\mathcal{B}}_T$. This is without loss of generality. We have,

$$\begin{aligned}
&\sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\left|\frac{\nabla_{\theta_j}\pi_{T,S}(z|\,v;\theta_0)w_T(z,v)}{\pi_{1T}(v)}-\frac{\nabla_{\theta_j}\pi(z|\,v;\theta_0)w(z,v)}{\pi_1(v;\theta_0)}\right| \\
&\le \sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\left[\frac{w_T(z,v)}{\pi_{1T}(v)}\left|\nabla_{\theta_j}\pi_{T,S}(z|\,v;\theta_0)-\nabla_{\theta_j}\pi(z|\,v;\theta_0)\right|\right] \\
&+\sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\frac{\left|\nabla_{\theta_j}\pi(z|\,v;\theta_0)\right|}{\pi_1(v;\theta_0)}\left|w_T(z,v)-w(z,v)\right| \\
&+\sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\frac{\left|\nabla_{\theta_j}\pi(z|\,v;\theta_0)\right|}{\pi_{1T}(v)\,\pi_1(v;\theta_0)}\left|\pi_{1T}(v)-\pi_1(v;\theta_0)\right| \\
&= O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q-1}\delta_T^{-3}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)-1}\delta_T^{-3}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\delta_T^{-4}\right)+O_p\left(\lambda_T^r\delta_T^{-4}\right) \\
&+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q}\delta_T^{-1}\right)+O_p\left(\lambda_T^r\delta_T^{-1}\right) \\
&+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q}\delta_T^{-2}\right)+O_p\left(\lambda_T^r\delta_T^{-2}\right),
\end{aligned}$$

by Lemma C0, Lemma N1-(ii), Assumptions 4 and 7, and boundedness of $|\nabla_\theta\pi(z|\,v;\theta_0)|$ (Assumption 1-(b)). ■

Lemma N3. *Let the assumptions in Lemma N2 hold. Then, for all* $i=1,\cdots,S$, *and* $j=1,\cdots,n$,

$$\begin{aligned}
&\sup_{(z,v)\in\mathbb{R}^{q^*}\times\mathcal{B}_T}\left|\frac{\nabla_{\theta_j}\pi_{T,S}(z|\,v;\theta_0)E\left[\pi_{2T}(z,v)\right]w_T(z,v)}{\pi_{1T}^i(v;\theta_0)\cdot\pi_{1T}(v)}-\frac{\nabla_{\theta_j}\pi(z|\,v;\theta_0)\pi_2(z,v;\theta_0)\,w(z,v)}{\pi_1(v;\theta_0)^2}\right| \\
=\;& O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q-1}\delta_T^{-4}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)-1}\delta_T^{-4}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\delta_T^{-5}\right)+O_p\left(\lambda_T^r\delta_T^{-5}\right).
\end{aligned}$$

Proof. As in the previous two lemmas, we demonstrate the result with the trimming set $\hat{\mathcal{B}}_T$

(w.l.o.g.). We have,

$$
\begin{aligned}
& \sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\left|\frac{\nabla_{\theta_j}\pi_{T,S}(z|\,v;\theta_0)E\left[\pi_{2T}\left(z,v\right)\right]w_T(z,v)}{\pi_{1T}^i\left(v;\theta_0\right)\cdot\pi_{1T}\left(v\right)}-\frac{\nabla_{\theta_j}\pi(z|\,v;\theta_0)\pi_2\left(z,v;\theta_0\right)w(z,v)}{\pi_1\left(v;\theta_0\right)^2}\right| \\
\leq & \sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\frac{E\left[\pi_{2T}\left(z,v\right)\right]w_T(z,v)}{\pi_{1T}^i\left(v;\theta_0\right)\pi_1\left(v;\theta_0\right)}\cdot\left|\nabla_{\theta_j}\pi_{T,S}(z|\,v;\theta_0)-\nabla_{\theta_j}\pi(z|\,v;\theta_0)\right| \\
& +\sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\frac{\left|\nabla_{\theta_j}\pi\left(z|\,v;\theta_0\right)\right|w_T(z,v)}{\pi_{1T}^i\left(v;\theta_0\right)\pi_1\left(v;\theta_0\right)}\cdot\left|E\left[\pi_{2T}\left(z,v\right)\right]-\pi_2\left(z,v;\theta_0\right)\right| \\
& +\sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\frac{\left|\nabla_{\theta_j}\pi\left(z|\,v;\theta_0\right)\right|\pi_2\left(z,v;\theta_0\right)}{\pi_{1T}^i\left(v;\theta_0\right)\pi_1\left(v;\theta_0\right)}\cdot\left|w_T(z,v)-w\left(z,v\right)\right| \\
& +\sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\frac{\left|\nabla_{\theta_j}\pi\left(z|\,v;\theta_0\right)\right|\pi_2\left(z,v;\theta_0\right)w\left(z,v\right)}{\pi_{1T}^i\left(v;\theta_0\right)\pi_1\left(v;\theta_0\right)^2}\cdot\left|\pi_{1T}^i\left(v;\theta_0\right)-\pi_1\left(v;\theta_0\right)\right| \\
& +\sup_{(z,v)\in\mathbb{R}^{q^*}\times\hat{\mathcal{B}}_T}\left[\frac{\left|\nabla_{\theta_j}\pi(z|\,v;\theta_0)\right|\pi_2\left(z,v;\theta_0\right)w(z,v)}{\pi_{1T}\left(v\right)\pi_1\left(v;\theta_0\right)\pi_T^i\left(v;\theta_0\right)}\cdot\left|\pi_{1T}\left(v\right)-\pi_1\left(v;\theta_0\right)\right|\right] \\
= & O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q-1}\delta_T^{-4}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)-1}\delta_T^{-4}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\delta_T^{-5}\right)+O_p\left(\lambda_T^r\delta_T^{-5}\right) \\
& +O_p\left(\lambda_T^r\delta_T^{-2}\right) \\
& +O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q}\delta_T^{-2}\right)+O_p\left(\lambda_T^r\delta_T^{-2}\right) \\
& +O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\delta_T^{-3}\right)+O_p\left(\lambda_T^r\delta_T^{-3}\right) \\
& +O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\delta_T^{-3}\right)+O_p\left(\lambda_T^r\delta_T^{-3}\right),
\end{aligned}
$$

by Lemma C0, Lemma N1-(ii), Assumption 7-(b), and boundedness of $\left|\nabla_{\theta_j}\pi(z|\,v;\theta_0)\right|$ (Assumption 1-(b)). ■

Lemma N4. *Let the assumptions in Lemma N1 hold. Let $v\mapsto\xi_{1T}\left(v\right)$ $(v\in\mathbb{R}^{q-q^*})$ be a sequence of real, bounded functions satisfying $\sup_{v\in\mathbb{R}^{q-q^*}}\left|\xi_{1T}\left(v\right)-\xi_1\left(v\right)\right|=O_p(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)})+O_p\left(\lambda_T^r\right)$, for some bounded function ξ_1. Then, for all $\theta\in\Theta$ and $j=1,\cdots,n$,*

$$
\begin{aligned}
& \sup_{(z,v)\in\mathcal{A}_T\cap\mathcal{B}_T}\left|\frac{\nabla_{\theta_j}\pi_{T,S}(z|\,v;\theta_0)\pi_{1T}\left(v\right)\xi_{1T}\left(v\right)}{\pi_{2T}\left(z,v\right)}-\frac{\nabla_{\theta_j}\pi\left(z|\,v;\theta_0\right)\xi_1\left(v\right)}{\pi\left(z|\,v;\theta_0\right)}\right| \\
& =O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q-1}\alpha_T^{-1}\delta_T^{-2}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)-1}\alpha_T^{-1}\delta_T^{-2}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\alpha_T^{-1}\delta_T^{-3}\right) \\
& +O_p\left(\lambda_T^r\alpha_T^{-1}\delta_T^{-3}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\alpha_T^{-1}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q}\alpha_T^{-2}\right)+O_p\left(\lambda_T^r\alpha_T^{-2}\right).
\end{aligned}
$$

Proof. Similarly as in the previous lemmas, we demonstrate the result with the trimming set $\hat{\mathcal{A}}_T\cap\hat{\mathcal{B}}_T$ (w.l.o.g.). We obviously have that at all points of continuity, $1/\pi\left(z|\,v;\theta_0\right)=$

$\pi(v;\theta_0)/\,\pi(z,v;\theta_0)$, and,

$$\begin{aligned}
&\sup_{(z,v)\in\hat{\mathcal{A}}_T\cap\hat{\mathcal{B}}_T}\left|\frac{\nabla_{\theta_j}\pi_{T,S}(z|\,v;\theta_0)\pi_{1T}(v)\,\xi_{1T}(v)}{\pi_{2T}(z,v)}-\frac{\nabla_{\theta_j}\pi(z|\,v;\theta_0)\pi_1(v;\theta_0)\,\xi_1(v)}{\pi_2(z,v;\theta_0)}\right| \\
\leq\;& \sup_{(z,v)\in\hat{\mathcal{A}}_T\cap\hat{\mathcal{B}}_T}\frac{\pi_{1T}(v)\,\xi_{1T}(v)}{\pi_{2T}(z,v)}\left|\nabla_{\theta_j}\pi_{T,S}(z|\,v;\theta_0)-\nabla_{\theta_j}\pi(z|\,v;\theta_0)\right| \\
&+\sup_{(z,v)\in\hat{\mathcal{A}}_T\cap\hat{\mathcal{B}}_T}\left|\nabla_{\theta_j}\pi(z|\,v;\theta_0)\right|\left|\frac{\pi_{1T}(v)\,\xi_{1T}(v)}{\pi_{2T}(z,v)}-\frac{\pi_1(v;\theta_0)\,\xi_1(v)}{\pi_2(z,v;\theta_0)}\right| \\
\equiv\;& S_{1T}+S_{2T}.
\end{aligned}$$

By Lemma N1-(ii),

$$\begin{aligned}
S_{1T} \;=\;& O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q-1}\delta_T^{-2}\alpha_T^{-1}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)-1}\delta_T^{-2}\alpha_T^{-1}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\delta_T^{-3}\alpha_T^{-1}\right) \\
&+O_p\left(\lambda_T^{r}\delta_T^{-3}\alpha_T^{-1}\right).
\end{aligned}$$

Moreover,

$$\begin{aligned}
S_{2T} \;\leq\;& \sup_{(z,v)\in\hat{\mathcal{A}}_T\cap\hat{\mathcal{B}}_T}\frac{\left|\nabla_{\theta_j}\pi(z|\,v;\theta_0)\right|\xi_{1T}(v)}{\pi_{2T}(z,v)}\left|\pi_{1T}(v)-\pi_1(v;\theta_0)\right| \\
&+\sup_{(z,v)\in\hat{\mathcal{A}}_T\cap\hat{\mathcal{B}}_T}\frac{\left|\nabla_{\theta_j}\pi(z|\,v;\theta_0)\right|\pi_1(v;\theta_0)}{\pi_{2T}(z,v)}\left|\xi_{1T}(v)-\xi_1(v)\right| \\
&+\sup_{(z,v)\in\hat{\mathcal{A}}_T\cap\hat{\mathcal{B}}_T}\frac{\left|\nabla_{\theta_j}\pi(z|\,v;\theta_0)\right|\pi_1(v;\theta_0)\,\xi_1(v)}{\pi_{2T}(z,v)\,\pi_2(z,v;\theta_0)}\left|\pi_{2T}(z,v)-\pi_2(z,v;\theta_0)\right| \\
\equiv\;& S_{21T}+S_{22T}+S_{23T}.
\end{aligned}$$

By Lemma C0 and Assumption 1-(b) on $\nabla_{\theta_j}\pi(z|\,v;\theta_0)$, and the assumption that $\xi_{1T}(v)$ is bounded, $S_{21T}=O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\alpha_T^{-1}\right)+O_p\left(\lambda_T^{r}\alpha_T^{-1}\right)$. By the assumption on the two functions $\xi_{1T}(v)$ and $\xi_1(v)$, we have that the S_{22T} term behaves as the S_{21T} term. Finally, by Lemma C0 and Assumption 7 on $\nabla_{\theta_j}\pi(z|\,v;\theta_0)$, $S_{23T}=O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q}\alpha_T^{-2}\right)+O_p\left(\lambda_T^{r}\alpha_T^{-2}\right)$. The result follows by boundedness of $\left|\nabla_{\theta_j}\pi(z|\,v;\theta_0)\right|$. ■

Lemma N5. *Let the assumptions in Lemma N1 hold, and let $\xi_{1T}(v)$ be the sequence of functions in Lemma N4. Then, for all $i=1,\cdots,S$, and $j=1,\cdots,n$,*

$$\begin{aligned}
&\sup_{(z,v)\in\mathcal{A}_T\cap\mathcal{B}_T}\left|\frac{\nabla_{\theta_j}\pi_{T,S}(z|\,v;\theta_0)\,E\left(\pi_{2T}(z,v)\right)\xi_{1T}(v)\,\pi_{1T}(v)}{\pi_{1T}^{i}(v;\theta_0)\,\pi_{2T}(z,v)}-\nabla_{\theta_j}\pi(z|\,v;\theta_0)\xi_1(v)\right| \\
&=O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q-1}\alpha_T^{-1}\delta_T^{-3}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q}\alpha_T^{-2}\delta_T^{-1}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\alpha_T^{-1}\delta_T^{-4}\right) \\
&+O_p\left(\lambda_T^{r}\alpha_T^{-1}\delta_T^{-4}\right)+O_p\left(\lambda_T^{r}\alpha_T^{-2}\delta_T^{-1}\right)+O_p\left(\lambda_T^{r}\delta_T^{-1}\alpha_T^{-1}\right).
\end{aligned}$$

Proof. Similarly as in the previous lemmas, we demonstrate the result with the trimming set $\hat{\mathcal{A}}_T \cap \hat{\mathcal{B}}_T$ (w.l.o.g.). We have,

$$
\begin{aligned}
& \sup_{(z,v)\in\hat{\mathcal{A}}_T\cap\hat{\mathcal{B}}_T} \left| \frac{\nabla_{\theta_j}\pi_{T,S}(z|\,v;\theta_0)E\left[\pi_{2T}\left(z,v\right)\right]\xi_{1T}\left(v\right)\pi_{1T}\left(v\right)}{\pi^i_{1T}\left(v;\theta_0\right)\pi_{2T}\left(z,v\right)} - \nabla_{\theta_j}\pi(z|\,v;\theta_0)\xi_1\left(v\right)\right| \\
\leq & \sup_{(z,v)\in\hat{\mathcal{A}}_T\cap\hat{\mathcal{B}}_T} \frac{E\left[\pi_{2T}\left(z,v\right)\right]}{\pi^i_{1T}\left(v;\theta_0\right)} \cdot \left| \frac{\nabla_{\theta_j}\pi_{T,S}(z|\,v;\theta_0)\pi_{1T}\left(v\right)\xi_{1T}\left(v\right)}{\pi_{2T}\left(z,v\right)} - \frac{\nabla_{\theta_j}\pi(z|\,v;\theta_0)\pi_1\left(v;\theta_0\right)\xi_1\left(v\right)}{\pi_2\left(z,v;\theta_0\right)}\right| \\
& + \sup_{(z,v)\in\hat{\mathcal{A}}_T\cap\hat{\mathcal{B}}_T} \left|\nabla_{\theta_j}\pi(z|\,v;\theta_0)\xi_1\left(v\right)\right| \left| \frac{\pi_1\left(v;\theta_0\right)E\left[\pi_{2T}\left(z,v\right)\right] - \pi_2\left(z,v;\theta_0\right)\pi^i_{1T}\left(v;\theta_0\right)}{\pi_2\left(z,v;\theta_0\right)\pi^i_{1T}\left(v;\theta_0\right)}\right| \\
\equiv & \; Q_{1T} + Q_{2T}.
\end{aligned}
$$

By Lemma N4,

$$
\begin{aligned}
Q_{1T} = & \; O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q-1}\alpha_T^{-1}\delta_T^{-3}\right) + O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)-1}\alpha_T^{-1}\delta_T^{-3}\right) + O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\alpha_T^{-1}\delta_T^{-4}\right) \\
& + O_p\left(\lambda_T^r\alpha_T^{-1}\delta_T^{-4}\right) + O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\alpha_T^{-1}\delta_T^{-1}\right) + O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q}\alpha_T^{-2}\delta_T^{-1}\right) + O_p\left(\lambda_T^r\alpha_T^{-2}\delta_T^{-1}\right).
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
Q_{2T} \leq & \sup_{(z,v)\in\hat{\mathcal{A}}_T\cap\hat{\mathcal{B}}_T} \frac{\left|\nabla_{\theta_j}\pi(z|\,v;\theta_0)\right|\xi_1\left(v\right)\pi_1\left(v;\theta_0\right)}{\pi_2\left(z,v;\theta_0\right)\pi^i_{1T}\left(v;\theta_0\right)} \left|E\left[\pi_{2T}\left(z,v\right)\right] - \pi_2\left(z,v;\theta_0\right)\right| \\
& + \sup_{(z,v)\in\hat{\mathcal{A}}_T\cap\hat{\mathcal{B}}_T} \frac{\left|\nabla_{\theta_j}\pi(z|\,v;\theta_0)\right|\xi_1\left(v\right)}{\pi^i_{1T}\left(v;\theta_0\right)} \left|\pi^i_{1T}\left(v;\theta_0\right) - \pi_1\left(v;\theta_0\right)\right| \\
= & \; O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q}\alpha_T^{-1}\delta_T^{-1}\right) + O_p\left(\lambda_T^r\alpha_T^{-1}\delta_T^{-1}\right) + O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\delta_T^{-1}\right) + O_p\left(\lambda_T^r\alpha_T^{-1}\right),
\end{aligned}
$$

by Lemma C0, and boundedness of $\left|\nabla_{\theta_j}\pi(z|\,v;\theta_0)\right|$ (Assumption 1-(b)) and $\xi_1\left(v\right)$. ■

A.3 Remarks on the bandwidth conditions in Assumptions 9 and 10

In the next remarks, we discuss the bandwidth conditions in Assumptions 9 and 10 of Al-M. For the reader's convenience, we rewrite these assumptions before producing the corresponding remarks.

Assumption 9 (Al-M, Section 3.1) *As* $T \to \infty$*,* $\delta_T \to 0$*. Moreover:*

(a) $\lambda_T \to \bar{\lambda}$, *where* $0 \leq \bar{\lambda} < \infty$, *and* $T^{\frac{1}{2}}\lambda_T^q\delta_T \to \infty$*;*

(b) $\lambda_T \to 0$, $T^{\frac{1}{2}}\lambda_T^{q+1}\delta_T^4 \to \infty$, *and* $\delta_T^{-1}\lambda_T^{\psi} \to 0$, *where* $\psi \equiv \min\left\{q^*+1, \frac{1}{5}r\right\}$.

Remark 1 (on Assumption 9, and its implication for Lemmas C1, N1, N2 and N3) (Lemmas C1, N1, N2 and N3 are needed to show Theorem 1 in Al-M)

It is easily seen that all the bandwidth conditions in Lemma C1-(a) hold under Assumption 9-(a). Below, we also show that under Assumption 9-(b), the condition in Lemma C1-(b) that

$\lambda_T^r \delta_T^{-1} \to 0$ holds as well. It is also easily seen that all the suprema in Lemmas N1-N3 go to zero in probability under the conditions that $\lambda_T \to 0$ and $T^{\frac{1}{2}} \lambda_T^{q+1} \delta_T^4 \to \infty$ in Assumption 9-(b). The only nontrivial conditions that must be shown to hold are that in Lemma N3, (i) $\lambda_T^r \delta_T^{-5} \to 0$ and (ii) $T^{1/2} \lambda_T^{q-q^*} \delta_T^5 \to \infty$. But by the second part of Assumption 9-(b), $(T^{\frac{1}{2}} \lambda_T^{q-q^*} \delta_T^5) \lambda_T^{q^*+1} \delta_T^{-1} \to \infty$. Hence, under the second part of Assumption 9-(b), we have that $T^{\frac{1}{2}} \lambda_T^{q-q^*} \delta_T^5 \to \infty$ holds if $\lambda_T^{q^*+1} \delta_T^{-1} \to 0$. So we must simultaneously have $\lambda_T^r \delta_T^{-5} \to 0$ and $\lambda_T^{q^*+1} \delta_T^{-1} \to 0$, that is $\delta_T^{-1} \lambda_T^{\min\{q^*+1, \frac{1}{5} r\}} \to 0$, which then implies that the condition in Lemma C1-(b) that $\lambda_T^r \delta_T^{-1} \to 0$ holds as well, as we initially claimed.

Assumption 10 (Al-M, Section 3.3) *As $T \to \infty$, $\alpha_T \to 0$, $\delta_T \to 0$ and $\delta_T / \alpha_T \to \kappa$, where κ is a constant. Moreover:*
(a) $\lambda_T \to \bar{\lambda}$, *where* $0 \leq \bar{\lambda} < \infty$, *and* $T^{\frac{1}{2}} \lambda_T^q \alpha_T^4 \to \infty$.
(b) $\lambda_T \to 0$, $T^{\frac{1}{2}} \lambda_T^{q+1} \alpha_T^4 \to \infty$, *and* $\alpha_T^{-1} \lambda_T^{\psi} \to 0$, *where* $\psi \equiv \min\left\{q^* + 1, \frac{1}{5} r\right\}$.

Remark 2 (on Assumption 10, and its implication for Lemmas C2, C3, N4 and N5) (Lemmas C2, C3, N4 and N5 are needed to show Theorem 2 in Al-M)

It is easily seen that the bandwidth conditions in Lemmas C2-(a) and C3-(a) hold if $\alpha_T \to 0$, $\delta_T \to 0$, $T^{\frac{1}{2}} \lambda_T^q \alpha_T^3 \to \infty$ and $T^{\frac{1}{2}} \lambda_T^q \alpha_T^2 \delta_T^2 \to \infty$. In turn, these conditions are satisfied if, for some constant κ, $\delta_T / \alpha_T \to \kappa$, and $T^{\frac{1}{2}} \lambda_T^q \alpha_T^4 \to \infty$, as required by Assumption 10-(a). Below, we also show that under Assumption 10-(b), the conditions $\alpha_T^3 \lambda_T^{-r} \to \infty$ (in Lemma C2-(b)) and $\alpha_T^2 \delta_T \lambda_T^{-r} \to \infty$ and $\alpha_T^2 \delta_T^2 \lambda_T^{-r} \to \infty$ (in Lemma C3-(b)) hold as well. It is also easily seen that all the suprema in Lemmas N4-N5 go to zero in probability under the conditions that $\lambda_T \to 0$, $\alpha_T \to 0$, $\delta_T \to 0$, and,

(i) $T^{\frac{1}{2}} \lambda_T^q \alpha_T^3 \to \infty$

(ii) $T^{\frac{1}{2}} \lambda_T^q \alpha_T^2 \delta_T \to \infty$

(iii) $T^{\frac{1}{2}} \lambda_T^{q+1} \alpha_T \delta_T^3 \to \infty$

(iv) $T^{\frac{1}{2}} \lambda_T^{q-q^*} \alpha_T \delta_T^4 \to \infty$

(v) $\lambda_T^r \alpha_T^{-3} \to 0$

(vi) $\lambda_T^r \alpha_T^{-2} \delta_T^{-2} \to 0$

(vii) $\lambda_T^r \alpha_T^{-1} \delta_T^{-4} \to 0$

If $\alpha_T \to 0$, $\delta_T \to 0$, and $\delta_T / \alpha_T \to \kappa$, as required by Assumption 10, the previous conditions can be simplified to, (i) $T^{\frac{1}{2}} \lambda_T^q \alpha_T^4 \to \infty$ (also required by Assumption 10-(b)); (ii) $T^{\frac{1}{2}} \lambda_T^{q-q^*} \alpha_T^5 \to \infty$ and (iii) $\lambda_T^r \alpha_T^{-5} \to 0$. By the same arguments produced in Remark 1, one has that (ii) and (iii) are satisfied if $\alpha_T^{-1} \lambda_T^{\min\{q^*+1, \frac{1}{5} r\}} \to 0$, as required by Assumption 10-(b). Clearly, these results also imply that $\alpha_T^3 \lambda_T^{-3} \to \infty$ (in Lemma C2-(b)) and $\alpha_T^2 \delta_T \lambda_T^{-r} \to \infty$ (in Lemma C3-(b)) (with $\delta_T / \alpha_T \to \kappa$), as we initially claimed.

B. Asymptotics for the J-SNE

B.1 Consistency

We have:

Proposition 1. *Let Assumptions 1-(a), 2, 3, 4-(a) and 11 hold. Then* $\forall\theta \in \Theta$, $L_{T,S}^{\mathrm{J}}(\theta) \xrightarrow{p} L^{\mathrm{J}}(\theta)$ *as* $T \to \infty$.

According to a well-known result (see Newey (1991, Thm. 2.1 p. 1162)), the following conditions are equivalent:

C1: $\lim_{T\to\infty} P\left(\sup_{\theta\in\Theta}\left|L_{T,S}^{\mathrm{J}}(\theta) - L^{\mathrm{J}}(\theta)\right| > \epsilon\right) = 0.$

C2: $\forall\theta \in \Theta$, $L_{T,S}^{\mathrm{J}}(\theta) \xrightarrow{p} L^{\mathrm{J}}(\theta)$, *and* $L_{T,S}^{\mathrm{J}}(\theta)$ *is stochastically equicontinuous.*

By Newey and McFadden (1994, Lemma 2.9 p. 2138), Assumption 6 guarantees that $L_{T,S}^{\mathrm{J}}(\theta)$ is stochastically equicontinuous, and so weak consistency follows from the equivalence of C1 and C2 above, Assumptions 11-12, compactness of Θ, and a classical argument (e.g., White (1994, Theorem 3.4)). So we are only left to prove Proposition 1.

Remark 3 (on the proof of Proposition 1) Before stating our proof of Proposition 1, it is useful to describe the main ideas underlying this proof. Our concern is to show that the integrand of $[L_{T,S}^{\mathrm{J}}(\theta) - L^{\mathrm{J}}(\theta)]$ is bounded by integrable functions independent of the sample size T, and that it converges in probability pointwise to zero as T goes to infinity. To establish these facts, we shall rely on an inequality which is a standard component of the consistency proof for methods of moments estimators (see, e.g., Duffie and Singleton (1993, eq. (A5), p. 949)). Let M_T and W_T be two sequences converging in probability to M_0 and W_0, respectively. (In our proof, M_T and W_T will be two estimated functions.) Then, the following inequality holds true,

$$\left|M_T^2 W_T - M_0^2 W_0\right| \leq |M_0|\,|W_T - W_0|\,|M_T| + |M_T - M_0|\left(W_T\,|M_T| + W_0\,|M_0|\right). \tag{B1}$$

We shall also use the inequality (B1) to establish consistency of the CD-SNE in Sections C and D.

Proof of Proposition 1. We produce the arguments that apply to the case in which the bandwidth sequence $\lambda_T \to \bar{\lambda} \in (0,\infty)$. Accordingly, we will make use of Lemma C0-(a). The case of a bandwidth sequence that satisfies $\lambda_T \to 0$ and $T^{1/2}\lambda_T^{q+1} \to \infty$ (which is assumed in Theorem 3 of Al-M, and used to show asymptotic normality in Section B.2 below) is dealt with similarly, given Assumption 4, by replacing Lemma C0-(a) with Lemma C0-(b). Below, we will denote $\bar{w}(x) \equiv E[w_T(x)]$.

We claim that:

$$\left|L_{T,S}^{\mathrm{J}}(\theta) - L^{\mathrm{J}}(\theta)\right| \leq \int \left(\sigma_{1T,S}(x;\theta) + \sigma_{2T,S}(x;\theta)\right) dx, \tag{B2}$$

where

$$
\begin{aligned}
\sigma_{1T,S}(x;\theta) &\equiv \left|\pi_{2T,S}(x;\theta)-\pi_{2T}(x)\right|\cdot\left|m_2(x;\theta)-m_2(x;\theta_0)\right|\cdot\left|w_T(x)-\bar{w}(x)\right| \\
\sigma_{2T,S}(x;\theta) &\equiv \left|\left[\pi_{2T,S}(x;\theta)-m_2(x;\theta)\right]-\left[\pi_{2T}(x)-m_2(x;\theta_0)\right]\right|\cdot\left[\bar{\phi}_{T,S}(x;\theta)+\bar{\phi}(x;\theta)\right] \\
\bar{\phi}_{T,S}(x;\theta) &\equiv \left|\pi_{2T,S}(x;\theta)-\pi_{2T}(x)\right|\cdot w_T(x) \\
\bar{\phi}(x;\theta) &\equiv \left|m_2(x;\theta)-m_2(x;\theta_0)\right|\cdot\bar{w}(x)
\end{aligned}
$$

provided the right hand side of (B2) is finite. Indeed, (B2) follows by an application of the inequality (B1) to the integrand of $[L^{\mathrm{J}}_{T,S}(\theta)-L^{\mathrm{J}}(\theta)]$, after setting $M_T\equiv[\pi_{2T,S}(x;\theta)-\pi_{2T}(x)]$, $M_0\equiv[m_2(x;\theta)-m_2(x;\theta_0)]$, $W_T=w_T(x)$ and $W_0=\bar{w}(x)$.

Next, we claim that for all $\theta\in\Theta$, $\int\sigma_{1T,S}(x;\theta)\xrightarrow{p}0$. We show this claim with a proof similar to that of the Glick's (1974) theorem, as given by Devroye and Györfy (1985, p. 10-11).

We have that $m_2(x;\theta)$, $w_T(x)$ and $\bar{w}(x)$ are bounded, and so

$$
\sigma_{1T,S}(x;\theta)\leq C\cdot\left(\left|m_2(x;\theta)-\pi_{2T,S}(x;\theta)\right|+\left|m_2(x;\theta)-\pi_{2T}(x)\right|\right),
$$

where C is some positive constant. Next, let $\Delta_{T,S}(x;\theta)\equiv m_2(x;\theta)-\pi_{2T,S}(x;\theta)$. We have, for all $\theta\in\Theta$,

$$
\begin{aligned}
\int\left|\Delta_{T,S}(x;\theta)\right|dx &= \int\left(\mathbb{I}_{\Delta_{T,S}(x;\theta)\geq0}-\mathbb{I}_{\Delta_{T,S}(x;\theta)<0}-\mathbb{I}_{\Delta_{T,S}(x;\theta)\geq0}+\mathbb{I}_{\Delta_{T,S}(x;\theta)\geq0}\right)\Delta_{T,S}(x;\theta)\,dx \\
&= 2\int\mathbb{I}_{\Delta_{T,S}(x;\theta)\geq0}\Delta_{T,S}(x;\theta)\,dx \\
&= 2\int\max\left\{\pi_{2T,S}(x;\theta)-m_2(x;\theta),0\right\}dx,
\end{aligned}
$$

where $\mathbb{I}.$ is the indicator function, and where the second equality follows because $\int(\mathbb{I}_{\Delta_{T,S}(x;\theta)<0}+\mathbb{I}_{\Delta_{T,S}(x;\theta)\geq0})\Delta_{T,S}(x;\theta)\,dx=\int\left[m_2(x;\theta)-\pi_{2T,S}(x;\theta)\right]dx=1-1=0$. Clearly, $\max\left\{m_2(x;\theta)-\pi_{2T,S}(x;\theta),0\right\}\leq m_2(x;\theta)$, which is integrable. By the same argument,

$$
\int\left|m_2(x;\theta)-\pi_{2T}(x)\right|dx=2\int\max\left\{m_2(x;\theta)-\pi_{2T}(x;\theta),0\right\}dx,
$$

where again, $\max\left\{m_2(x;\theta)-\pi_{2T}(x;\theta),0\right\}\leq m_2(x;\theta)$. Hence, for all $\theta\in\Theta$, $\sigma_{1T,S}(x;\theta)$ is bounded by integrable functions independent of T. As $T\to\infty$, $\sigma_{1T,S}(x;\theta)\xrightarrow{p}0$, x-pointwise. By dominated convergence, $\lim_{T\to\infty}E\left[\sigma_{1T,S}(x;\theta)\right]=E\left[\lim_{T\to\infty}\sigma_{1T,S}(x;\theta)\right]=0$ all $(x,\theta)\in\mathbb{R}^q\times\Theta$. By Fubini, $E\left[\int\sigma_{1T,S}(x;\theta)\,dx\right]=\int E\left[\sigma_{1T,S}(x;\theta)\right]dx$ all $\theta\in\Theta$. Again by dominated convergence,

$$
\lim_{T\to\infty}E\left[\int\sigma_{1T,S}(x;\theta)\,dx\right]=\lim_{T\to\infty}\int E\left[\sigma_{1T,S}(x;\theta)\right]dx=\int\lim_{T\to\infty}E\left[\sigma_{1T,S}(x;\theta)\right]dx=0,\ \forall\theta\in\Theta.
$$

By Markov's inequality:

$$
\forall\epsilon>0,\ \ P\left\{\int\sigma_{1T,S}(x;\theta)\,dx>\epsilon\right\}\leq\frac{E\left[\int\sigma_{1T,S}(x;\theta)\,dx\right]}{\epsilon},\ \ \forall\theta\in\Theta.
$$

Hence, for all $\theta\in\Theta$, $\int\sigma_{1T,S}(x;\theta)\,dx\xrightarrow{p}0$. By the same arguments, for all $\theta\in\Theta$, $\int\sigma_{2T,S}(x;\theta)\,dx\xrightarrow{p}0$. Hence, for all $\theta\in\Theta$, $\left|L^{\mathrm{J}}_{T,S}(\theta)-L^{\mathrm{J}}(\theta)\right|\xrightarrow{p}0$. ■

B.2 Asymptotic normality

By Assumption 7-(a), the order of derivation and integration in $\nabla_\theta L^{\mathrm{J}}_{T,S}(\theta)$ can be interchanged (see Newey and McFadden (1994, Lemma 3.6 p. 2152-2153)), and the first order conditions satisfied by the J-SNE are,

$$\mathbf{0}_n = \int \left[\pi_{2T,S}(x;\theta^{\mathrm{J}}_{T,S}) - \pi_{2T}(x)\right] \nabla_\theta \pi_{2T,S}\left(x;\theta^{\mathrm{J}}_{T,S}\right) w_T(x)dx.$$

Let $\theta(c) \equiv c \circ (\theta_0 - \theta^{\mathrm{J}}_{T,S}) + \theta^{\mathrm{J}}_{T,S}$, where, for any $c \in (0,1)^n$ and $\theta \in \Theta$, $c \circ \theta$ denotes the vector in Θ whose i-th element is $c^{(i)}\theta^{(i)}$. By Assumption 7-(a), there exists a c^* in $(0,1)^n$ such that:

$$\begin{aligned}
\mathbf{0}_n &= \sqrt{T}\int \left[\pi_{2T,S}(x;\theta_0) - \pi_{2T}(x)\right] \nabla_\theta \pi_{2T,S}\left(x;\theta_0\right) w_T(x)dx \\
&\quad + \left[\int \left|\nabla_\theta \pi_{2T,S}(x;\bar{\theta})\right|_2 w_T(x)dx + \left(\bar{\theta}-\theta_0\right)\cdot k_{1T,S}\left(\bar{\theta}\right) + k_{2T,S}\left(\bar{\theta}\right)\right]\cdot \sqrt{T}(\theta^{\mathrm{J}}_{T,S}-\theta_0), \qquad \text{(B3)}
\end{aligned}$$

where $\bar{\theta} \equiv \theta(c^*)$, and for some θ^*, $k_{1T,S}\left(\bar{\theta}\right)$ and $k_{2T,S}\left(\bar{\theta}\right)$ are such that,

$$\begin{aligned}
\left|k_{1T,S}\left(\bar{\theta}\right)\right| &\leq \int \left|\nabla_\theta \pi_{2T,S}(x;\theta^*)\right| \left|\nabla_{\theta\theta}\pi_{2T,S}\left(x;\bar{\theta}\right)\right| w_T(x)dx \\
\left|k_{2T,S}\left(\bar{\theta}\right)\right| &\leq \int \left|\pi_{2T,S}(x;\theta_0) - \pi_{2T}(x)\right| \left|\nabla_{\theta\theta}\pi_{2T,S}\left(x;\bar{\theta}\right)\right| w_T(x)dx
\end{aligned}$$

By Assumption 7-(a), the term $\nabla_{\theta\theta}\pi_{2T,S}\left(x;\bar{\theta}\right)$ is bounded in probability as T becomes large. Hence (i) so is $\left|k_{1T,S}\left(\bar{\theta}\right)\right|$; and (ii) by Lemma C0, $\left|k_{2T,S}\left(\bar{\theta}\right)\right| \xrightarrow{p} \mathbf{0}_{n\times n}$. Moreover,

$$\int \left|\nabla_\theta \pi_{2T,S}(x;\bar{\theta})\right|_2 w_T(x)dx = \int \left|\nabla_\theta \pi_{2T,S}(x;\theta_0)\right|_2 w_T(x)dx + R_{T,S}\left(\bar{\theta}\right),$$

where $R_{T,S}\left(\bar{\theta}\right)$ is such that,

$$\left|R_{T,S}\left(\bar{\theta}\right)\right|_{i,j} \leq \int \left|\left|\nabla_\theta \pi_{2T,S}(x;\bar{\theta})\right|_2 - \left|\nabla_\theta \pi_{2T,S}(x;\theta_0)\right|_2\right|_{i,j} w_T(x)dx.$$

Since $\int |w_T - w| \xrightarrow{p} 0$ and $\bar{\theta} \xrightarrow{p} \theta_0$, then by Lemma N1-(i), $\left|R_{T,S}\left(\bar{\theta}\right)\right|_{i,j} \xrightarrow{p} 0$ for all i,j. Hence,

$$\int \left|\nabla_\theta \pi_{2T,S}(x;\bar{\theta})\right|_2 w_T(x)dx + \left(\bar{\theta}-\theta_0\right) k_{1T,S}\left(\bar{\theta}\right) + k_{2T,S}\left(\bar{\theta}\right) \xrightarrow{p} \int \left|\nabla_\theta \pi_2\left(x;\theta_0\right)\right|_2 w(x)dx. \qquad \text{(B4)}$$

Next, consider the first term in (B3). For all $x \in \mathbb{R}^q$ and fixed T, $E\left[\pi^i_{2T}(x;\theta_0)\right] = E\left[\pi_{2T}(x)\right]$ $(i = 1,\cdots,S)$. Hence,

$$\begin{aligned}
&\sqrt{T}\int \left[\pi_{2T,S}(x;\theta_0) - \pi_{2T}(x)\right] \nabla_\theta \pi_{2T,S}(x;\theta_0) w_T(x)dx \\
&= \int \sqrt{T}\left[\pi_{2T,S}(x;\theta_0) - E(\pi_{2T,S}(x;\theta_0))\right] \nabla_\theta \pi_{2T,S}(x;\theta_0) w_T(x)dx \\
&- \int \sqrt{T}\left[\pi_{2T}(x) - E(\pi_{2T}(x))\right] \nabla_\theta \pi_{2T,S}(x;\theta_0) w_T(x)dx. \qquad \text{(B5)}
\end{aligned}$$

Let $\mathbb{G}$ be a measurable V-C subgraph class of uniformly bounded functions (see, e.g., Arcones and Yu (1994, Definition 2.2 p. 51)). By Arcones and Yu (1994, Corollary 2.1 p. 59-60), for each $G \in \mathbb{G}$, $T^{-1/2} \sum_{t=1+l}^{T} [G(x_t) - EG]$ converges in law to a Gaussian process under Assumption 2. Now $\lambda_T^{-q} K_q ((x_t - x)/\lambda_T) \in \mathbb{G}$. Let $F(x;\theta) = \int_0^x \pi_2(\xi;\theta)\, d\xi$, $F_T(x) = \int_0^x \pi_{2T}(\xi)\, d\xi$ and $F(x) = \int_0^x \pi_2(\xi;\theta_0)\, d\xi$. Under the theorem's assumptions,

$$A_T(x) \equiv \sqrt{T}\left[F_T(x) - E(F_T(x))\right] \Rightarrow \omega^0(F(x)),$$

where $\omega^0(F(x))$ is a Gaussian process with covariance kernel,

$$\min\left\{F(x), F(x')\right\}\left[1 - F(x')\right] + \sum_{k=1}^{\infty}\left[F^k(x,x') + F^k(x',x) - 2F(x)F(x')\right],$$

and $F^k(x,x') \equiv P(x_0 \leq x, x_k \leq x')$. We have,

$$\begin{aligned} J_{T,S} &\equiv \sqrt{T}\int \left[\pi_{2T}(x) - E(\pi_{2T}(x))\right] \nabla_\theta \pi_{2T,S}(x;\theta_0) w_T(x) dx \\ &= \int [w_T(x) - w(x)]\left[\nabla_\theta \pi_{2T,S}(x;\theta_0) - \nabla_\theta \pi_2(x;\theta_0)\right] dA_T(x) \\ &+ \int \left[\nabla_\theta \pi_{2T,S}(x;\theta_0) - \nabla_\theta \pi_2(x;\theta_0)\right] w(x)\, dA_T(x) \\ &+ \int [w_T(x) - w(x)]\, \nabla_\theta \pi_2(x;\theta_0)\, dA_T(x) + \int \nabla_\theta \pi_2(x;\theta_0)\, w(x)\, dA_T(x) \\ &\equiv J_{1T,S} + J_{2T,S} + J_{3T} + J_{4T}. \end{aligned}$$

By the continuous mapping theorem (e.g. Dudley (1990) (Prop. 9.3.7)),

$$J_{4T} \xrightarrow{d} J_4 \equiv \int \nabla_\theta \pi_2(x;\theta_0)\, w(x) d\omega^0(F(x)). \tag{B6}$$

By Assumption 4-(b) on the weighting functions $w_T(x)$ and $w(x)$, and boundedness of $\nabla_\theta \pi_2(x;\theta_0)$ (Assumption 1-(b)), and by Lemma N1-(i), $J_{iT,S} = [O_p(T^{-\frac{1}{2}}\lambda_T^{-q-1}) + O_p(\lambda_T^r)]$, $i = 1, 2$. By Assumption 4-(b) and boundedness of $\nabla_\theta \pi_2(x;\theta_0)$ (Assumption 1-(b)), $J_{3T} \xrightarrow{p} 0$. Therefore, by the bandwidth conditions given in Theorem 3 of Al-M, we have that:

$$J_{T,S} \xrightarrow{d} N(0, W), \qquad W \equiv var(J_4).$$

By the same computations in Aït-Sahalia (1994) (proof of Thm. 1 p. 21-22) and Aït-Sahalia (1996) (proof of eq. (12), p. 420-421),

$$\begin{aligned} W &= var\left[\nabla_\theta \pi_2(x_t;\theta_0)\, w(x_t)\right] + \sum_{k=1}^{\infty}\left\{cov\left[\nabla_\theta \pi_2(x_t;\theta_0)\, w(x_t), \nabla_\theta \pi_2(x_{t+k};\theta_0)\, w(x_{t+k})\right]\right. \\ &\left. + \ cov\left[\nabla_\theta \pi_2(x_{t+k};\theta_0)\, w(x_{t+k}), \nabla_\theta \pi_2(x_t;\theta_0)\, w(x_t)\right]\right\}, \end{aligned} \tag{B7}$$

which is finite by the mixing condition in Assumption 2 and the assumption that $E[\|\nabla_\theta \pi_2(x_t;\theta_0)\, w(x_t)\|^{\vartheta}]^{1/\vartheta} < \infty$, for some $\vartheta > 2$, by, e.g., Politis and Romano (1994) (Thm. 2.3 p. 466). Finally, let $F_T^i(x;\theta) \equiv \int_0^x \pi_{2T}^i(v;\theta)dv$, $i = 1, \cdots, S$. As for $A_T(x)$, we have that $A_T^i(x) \equiv \sqrt{T}[F_T^i(x;\theta_0) - E(F_T^i(x;\theta_0))] \Rightarrow \omega_i^0(F(x))$, where $\omega_i^0(F)$ are independent Gaussian processes. Hence,

$$\sqrt{T}\sum_{i=1}^{S}\left[F_T^i(x;\theta_0) - E(F_T^i(x;\theta_0))\right] \Rightarrow \sum_{i=1}^{S}\omega_i^0(F(x)).$$

Since $E(F_T^i(x;\theta_0)) = E(F_T^j(x;\theta_0))$ for all $i, j = 1, \cdots, S$, we have, similarly as for the J_T term,

$$\int [\sqrt{T}\,(\pi_{2T,S}(x;\theta_0) - E(\pi_{2T,S}(x;\theta_0)))]\nabla_\theta \pi_{2T,S}(x;\theta_0) w_T(x) dx \xrightarrow{d} N\left(0, \frac{1}{S}W\right),$$

where W is as in (B7). Finally, $A_T(x)$ and $A_T^i(x)$, $i = 1, \cdots, S$, are all independent. Therefore, by (B5),

$$\sqrt{T}\int\left[\pi_{2T}(x;\theta_0) - \pi_{2T}(x)\right]\nabla_\theta \pi_{2T,S}(x;\theta_0)\, w_T(x)\, dx \xrightarrow{d} N\left(0, \left(1+\frac{1}{S}\right)W\right). \qquad \text{(B8)}$$

Hence by (B3), (B4), (B8) and Slutzky's theorem,

$$\sqrt{T}\left(\theta_{T,S}^{\mathrm{J}} - \theta_0\right) \xrightarrow{d} N\left(0, \left(1+\frac{1}{S}\right)\mathcal{D}^{-1}W\mathcal{D}^{\top -1}\right),$$

where $\mathcal{D} \equiv \int |\nabla_\theta \pi_2(x;\theta_0)|_2\, w(x)\, dx$ and W is as in (B7).

C. Asymptotics for the CD-SNE

C.1 Consistency

We produce the arguments that apply to the case in which the bandwidth sequence satisfies Assumption 9-(a) of Al-M. Accordingly, we will make a repeated use of Lemma C1-(a). The case of a bandwidth sequence that satisfies Assumption 9-(b) of Al-M (which is needed to prove asymptotic normality in Section C.2) is dealt with similarly, by replacing Lemma C1-(a) with Lemma C1-(b). As in Section B.1 of the present Appendix, below we will denote $\bar{w}(z,v) \equiv E[w_T(z,v)]$.

We claim that the objective function of the CD-SNE $L_{T,S}^{\text{CD}}(\theta)$ satisfies

$$\left|L_{T,S}^{\text{CD}}(\theta) - L^{\text{CD}}(\theta)\right| \leq \iint \left(a_{1T,S}(z,v;\theta) + a_{2T,S}(z,v;\theta)\right) dzdv, \tag{C1}$$

where

$$\begin{aligned}
a_{1T,S}(z,v;\theta) &\equiv \left|\pi_{T,S}(z|\,v;\theta) - \pi_T(z|\,v)\right| \mathbb{T}_{T,S}(v;\theta) \cdot \left|m(z|\,v;\theta) - m(z|\,v;\theta_0)\right| \cdot \left|w_T(z,v) - \bar{w}(z,v)\right|; \\
a_{2T,S}(z,v;\theta) &\equiv \left|\left[\pi_{T,S}(z|\,v;\theta)\mathbb{T}_{T,S}(v;\theta) - m(z|\,v;\theta)\right] - \left[\pi_T(z|\,v)\mathbb{T}_{T,S}(v;\theta) - m(z|\,v;\theta_0)\right]\right| \\
&\quad \times \left[\phi_{T,S}(z,v;\theta) + \phi(z,v;\theta)\right]; \\
\phi_{T,S}(z,v;\theta) &\equiv \left|\pi_{T,S}(z|\,v;\theta) - \pi_T(z|\,v)\right| \mathbb{T}_{T,S}(v;\theta) \cdot w_T(z,v); \\
\phi(z,v;\theta) &\equiv \left|m(z|\,v;\theta) - m(z|\,v;\theta_0)\right| \cdot \bar{w}(z,v).
\end{aligned} \tag{C2}$$

provided the right hand side of (C1) is finite. Indeed, (C1) follows by the remarks on the proof of Proposition 1 in Section B.1 (Remark 3), and by an application of the inequality (B1) to the integrand of $[L_{T,S}^{\text{CD}}(\theta) - L^{\text{CD}}(\theta)]$, after setting $M_T \equiv [\pi_{T,S}(z|\,v;\theta) - \pi_T(z|\,v)]\,\mathbb{T}_{T,S}(v;\theta)$, $M_0 \equiv [m(z|\,v;\theta) - m(z|\,v;\theta_0)]$, $W_T = w_T(z,v)$ and $W_0 = \bar{w}(z,v)$.

Next, we show that $\iint (a_{1T,S} + a_{2T,S}) \xrightarrow{p} 0$ for all $\theta \in \Theta$. We study the two integrals separately.

- For all $(z,v,\theta) \in \mathbb{R}^{q^*} \times \mathbb{R}^{q-q^*} \times \Theta$, $a_{1T,S}(z,v;\theta) \leq \ell_T(z,v;\theta) \cdot \phi_{2T,S}(z,v;\theta)$, where

$$\begin{aligned}
\ell_T(z,v;\theta) &\equiv \left|m(z|\,v;\theta) - m(z|\,v;\theta_0)\right| \cdot \left|w_T(z,v) - \bar{w}(z,v)\right| \\
\phi_{2T,S}(z,v;\theta) &\equiv \frac{1}{S}\sum_{i=1}^{S} \left|\pi_T^i(z|\,v;\theta) - m(z|\,v;\theta)\right| \mathbb{T}_{T,S}(v;\theta) + \left|\pi_T(z|\,v) - m(z|\,v;\theta_0)\right| \mathbb{T}_{T,S}(v;\theta) \\
&\quad + \left|m(z|\,v;\theta) - m(z|\,v;\theta_0)\right| \mathbb{T}_{T,S}(v;\theta).
\end{aligned} \tag{C3}$$

  By Assumptions 1-(a), 3 and 4, we have that for each $\theta \in \Theta$, the function ℓ_T is bounded by integrable functions independent of T. Moreover, by Assumption 4, for all $\theta \in \Theta$, $\ell_T(z,v;\theta) \xrightarrow{p} 0$ (z,v)-pointwise. Finally, by Lemma C1-(a),

$$\sup_{(z,v)\in\mathbb{R}^q} \left|\pi_T^i(z|\,v;\theta) - m(z|\,v;\theta)\right| \mathbb{T}_{T,S}(v;\theta) \xrightarrow{p} 0, \quad i = 1, \cdots, S.$$

  This result clearly holds for the $S+1$ terms of $\phi_{2T,S}$ in (C3) as well. Finally, $|m(z|\,v;\theta) - m(z|\,v;\theta_0)|$ is bounded. Therefore, for all $\theta \in \Theta$,

$$\iint a_{1T,S}(z,v;\theta)\, dzdv \xrightarrow{p} 0. \tag{C4}$$

- For all $(z, v, \theta) \in \mathbb{R}^{q^*} \times \mathbb{R}^{q-q^*} \times \Theta$,

$$\begin{aligned} a_{2T,S}(z, v; \theta) \;\; \leq \;\; & \frac{1}{S} \sum_{i=1}^{S} \left| \pi_T^i(z|\, v; \theta) \mathbb{T}_{T,S}(v; \theta) - m(z|\, v; \theta) \right| \phi_{3T,S}(z, v; \theta) \\ & + \left| \pi_T(z|\, v) \mathbb{T}_{T,S}(v; \theta) - m(z|\, v; \theta_0) \right| \phi_{3T,S}(z, v; \theta), \end{aligned} \tag{C5}$$

where $\phi_{3T,S}(z, v; \theta) \equiv \phi(z, v; \theta) + \phi_{T,S}(z, v; \theta) \leq \phi(z, v; \theta) + \phi_{2T,S}(z, v; \theta)\, w_T(z, v)$. For each $i = 1, \cdots, S$, and $(z, v, \theta) \in \mathbb{R}^{q^*} \times \mathbb{R}^{q-q^*} \times \Theta$,

$$\begin{aligned} & \left| \pi_T^i(z|\, v; \theta) \mathbb{T}_{T,S}(v; \theta) - m(z|\, v; \theta) \right| \phi_{3T}(z, v; \theta) \\ & \leq m(z|\, v; \theta) \left[1 - \mathbb{T}_{T,S}(v; \theta) \right] \cdot \left[\phi(z, v; \theta) + \phi_{2T,S}(z, v; \theta)\, w_T(z, v) \right] \\ & + \left| \pi_T^i(z|\, v; \theta) - m(z|\, v; \theta) \right| \mathbb{T}_{T,S}(v; \theta) \cdot \left[\phi(z, v; \theta) + \phi_{2T,S}(z, v; \theta)\, w_T(z, v) \right] \\ & \equiv a_{21T,S}(z, v; \theta) + a_{22T,S}(z, v; \theta), \end{aligned}$$

where the inequality holds by the triangle inequality. Since w_T, ϕ and m are bounded, and w_T and ϕ are also integrable, $\iint a_{22T,S}(z, v; \theta) \xrightarrow{p} 0$ for all $\theta \in \Theta$ by Lemma C1-(a). As for the $a_{21T,S}$ term, $|1 - \mathbb{T}_{T,S}(v; \theta)| \leq 1$. Moreover, $[1 - \mathbb{T}_{T,S}(v; \theta)] \xrightarrow{p} 0$ pointwise. Hence, by the previous results on $\phi_{2T,S}$ and Lemma C1-(a), $\iint a_{21T,S}(z, v; \theta) \xrightarrow{p} 0$ for all $\theta \in \Theta$. By reiterating the previous arguments, one shows that the same result holds for the second term in (C5) and, hence, for all $\theta \in \Theta$,

$$\iint a_{2T,S}(z, v; \theta)\, dz dv \xrightarrow{p} 0. \tag{C6}$$

Hence, the consistency proof is complete, by eqs. (C1), (C4) and (C6).

C.2 Asymptotic normality

Remarks on the use of Lemmas N1-N3.

(a) Lemma N1 is needed to show that the $\mathcal{J}_{T,S}$ term in the asymptotic expansion in (C30) below (see (C31a)) converges in probability to $\mathcal{J}$, where $\mathcal{J}$ is given in eq. (C34c). Lemmas N2 and N3 are needed to show that the terms $\mathcal{I}_{1T,S}^{i}$ and $\mathcal{I}_{2T,S}^{i}$ in the asymptotic expansion in (C30) (see (C31b)-(C31c)) converge in distribution to the Gaussian terms $\mathcal{I}_1^i$ and $\mathcal{I}_2^i$ given in eqs. (C34a)-(C34b)

(b) The bandwidth conditions in Assumption 9-(b) of Al-M guarantee that the suprema in Lemmas N1-N3 go to zero in probability, as explained in the Remarks 1 of Section A.3 of the present Appendix.

The following remark is useful.

Remark 4. We have,

$$\begin{aligned}
\nabla_\theta \mathbb{T}_{T,S}(v;\theta_0) &= G_{\delta_T}(\pi_{1T}(v))\,\nabla_\theta \prod_{i=1}^{S} G_{\delta_T}\left(\pi_{1T}^{i}(v;\theta_0)\right)\\
&= G_{\delta_T}(\pi_{1T}(v))\sum_{i=1}^{S}\left[\nabla_\theta G_{\delta_T}\left(\pi_{1T}^{i}(v;\theta_0)\right)\right]\prod_{j\neq i} G_{\delta_T}\left(\pi_{1T}^{j}(v;\theta_0)\right)\\
&= G_{\delta_T}(\pi_{1T}(v))\sum_{i=1}^{S} g_{\delta_T}\left(\pi_{1T}^{i}(v;\theta_0)\right)\left[\nabla_\theta \pi_{1T}^{i}(v;\theta_0)\right]\prod_{j\neq i} G_{\delta_T}\left(\pi_{1T}^{j}(v;\theta_0)\right)\\
&= \sum_{i=1}^{S} g_{\delta_T}\left(\pi_{1T}^{i}(v;\theta_0)\right)\left[\nabla_\theta \pi_{1T}^{i}(v;\theta_0)\right] G_{\delta_T}^{(-i)}(v;\theta_0),
\end{aligned}$$

where

$$G_{\delta_T}^{(-i)}(v;\theta_0) \equiv \prod_{j=0\ ;\ j\neq i}^{S} G_{\delta_T}\left(\pi_{1T}^{j}(v;\theta_0)\right),$$

and g_δ is the function introduced in Assumption 8.

Remark 5. For all $\ell = 1, \cdots, S$, we have,

$$\nabla_{\theta_i}\pi_T^{\ell}(z|\,v;\theta) = \frac{\nabla_{\theta_i}\pi_{2T}^{\ell}(z,v;\theta)}{\pi_{1T}^{\ell}(v;\theta)} - \frac{\pi_{2T}^{\ell}(z,v;\theta)}{\pi_{1T}^{\ell}(v;\theta)^2}\nabla_{\theta_i}\pi_{1T}^{\ell}(v;\theta),$$

$$\begin{aligned}
\nabla_{\theta_i\theta_j}\pi_T^{\ell}(z|\,v;\theta) &= \frac{\nabla_{\theta_j\theta_i}\pi_{2T}^{\ell}(z,v;\theta)}{\pi_{1T}^{\ell}(v;\theta)}\\
&\quad - \frac{\nabla_{\theta_i}\pi_{2T}^{\ell}(z,v;\theta)\,\nabla_{\theta_j}\pi_{1T}^{\ell}(v;\theta) + \nabla_{\theta_j}\pi_{2T}^{\ell}(z,v;\theta)\,\nabla_{\theta_i}\pi_{1T}^{\ell}(v;\theta) + \pi_{2T}^{\ell}(z,v;\theta)\,\nabla_{\theta_j\theta_i}\pi_{1T}^{\ell}(v;\theta)}{\pi_{1T}^{\ell}(v;\theta)^2}\\
&\quad + 2\frac{\pi_{2T}^{\ell}(z,v;\theta)\,\nabla_{\theta_i}\pi_{1T}^{\ell}(v;\theta)\,\nabla_{\theta_j}\pi_{1T}^{\ell}(v;\theta)}{\pi_{1T}^{\ell}(v;\theta)^3},
\end{aligned}$$

at all points of continuity.

We now demonstrate our asymptotic normality claims. By Remark 5 and Assumption 7-(a), we may interchange the order of derivation and integration in $\nabla_\theta L_{T,S}^{\mathrm{CD}}(\theta)$ (similarly as for the J-SNE in Section B.2). Thus, the CD-SNE satisfies the following first order conditions:

$$\begin{aligned}\mathbf{0}_n &= \frac{1}{S}\sum_{i=1}^{S}\iint\left[\frac{\pi_{2T}^i(z,v;\theta_{T,S})}{\pi_{1T}^i(v;\theta_{T,S})}-\frac{\pi_{2T}(z,v)}{\pi_{1T}(v)}\right]\nabla_\theta\pi_{T,S}(z|v;\theta_{T,S})\,w_T(z,v)\mathbb{T}_{T,S}^2(v;\theta_{T,S})\,dzdv\\ &\quad+\iint\left[\pi_{T,S}(z|v;\theta_{T,S})-\pi_T(z|v)\right]^2 w_T(z,v)\mathbb{T}_{T,S}(v;\theta_{T,S})\nabla_\theta\mathbb{T}_{T,S}(v;\theta_{T,S})\,dzdv.\end{aligned}$$

We have, for some convex combination $\bar{\theta}$ of θ_0 and $\theta_{T,S}$,

$$\begin{aligned}\mathbf{0}_n &= \frac{1}{S}\sum_{i=1}^{S}\sqrt{T}\iint\left[\frac{\pi_{2T}^i(z,v;\theta_0)}{\pi_{1T}^i(v;\theta_0)}-\frac{\pi_{2T}(z,v)}{\pi_{1T}(v)}\right]\nabla_\theta\pi_{T,S}(z|v;\theta_0)\,w_T(z,v)\mathbb{T}_{T,S}^2(v;\theta_0)\,dzdv\\ &\quad+B_{T,S}+C_{T,S}\cdot\sqrt{T}\left(\theta_{T,S}-\theta_0\right),\end{aligned}\tag{C7}$$

where

$$B_{T,S}\equiv\sqrt{T}\iint\left[\pi_{T,S}(z|v;\theta_0)-\pi_T(z|v)\right]^2 w_T(z,v)\,\mathbb{T}_{T,S}(v;\theta_0)\,\nabla_\theta\mathbb{T}_{T,S}(v;\theta_0)\,dzdv.\tag{C8}$$

and,

$$\begin{aligned}C_{T,S}&\equiv\iint\nabla_\theta\left\{\left[\pi_{T,S}(z|v;\bar{\theta})-\pi_T(z|v)\right]\nabla_\theta\pi_{T,S}(z|v;\bar{\theta})\,\mathbb{T}_{T,S}^2(v;\bar{\theta})\right\}w_T(z,v)\,dzdv\\ &\quad+\iint\nabla_\theta\left\{\left[\pi_{T,S}(z|v;\bar{\theta})-\pi_T(z|v)\right]^2\mathbb{T}_{T,S}(v;\bar{\theta})\,\nabla_\theta\mathbb{T}_{T,S}(v;\bar{\theta})\right\}w_T(z,v)\,dzdv.\end{aligned}\tag{C9}$$

We now study the two terms $B_{T,S}$ and $C_{T,S}$, and then elaborate on the first order conditions in (C7).

Analysis of the $B_{T,S}$ term

Claim 1. *We have,*

$$B_{T,S}\xrightarrow{p}\mathbf{0}_n.\tag{C10}$$

To prove Claim 1, note, first, that the term $B_{T,S}$ in (C8) is:

$$\begin{aligned}B_{T,S}&=\sqrt{T}\iint\frac{1}{S}\sum_{i=1}^{S}\left[\frac{\pi_{2T}^i(z,v;\theta_0)}{\pi_{1T}^i(v;\theta_0)}-\frac{\pi_{2T}(z,v)}{\pi_{1T}(v)}\right]\left[\pi_{T,S}(z|v;\theta_0)-\pi_T(z|v)\right]\\ &\qquad\times\left[w_T(z,v)-w(z,v)\right]\mathbb{T}_{T,S}(v;\theta_0)\,\nabla_\theta\mathbb{T}_{T,S}(v;\theta_0)\,dzdv\\ &\quad+\sqrt{T}\iint\frac{1}{S}\sum_{i=1}^{S}\left[\frac{\pi_{2T}^i(z,v;\theta_0)}{\pi_{1T}^i(v;\theta_0)}-\frac{\pi_{2T}(z,v)}{\pi_{1T}(v)}\right]\left[\pi_{T,S}(z|v;\theta_0)-\pi_T(z|v)\right]\\ &\qquad\times w(z,v)\,\mathbb{T}_{T,S}(v;\theta_0)\,\nabla_\theta\mathbb{T}_{T,S}(v;\theta_0)\,dzdv\\ &\equiv B_{1T,S}+B_{2T,S}.\end{aligned}\tag{C11}$$

We show that $B_{1T,S} \xrightarrow{p} \mathbf{0}_n$. Clearly,

$$\begin{aligned}B_{1T,S} &= \sqrt{T}\iint \frac{1}{S}\sum_{i=1}^{S}\left[\frac{\pi_{2T}^{i}\left(z,v;\theta_0\right)}{\pi_{1T}^{i}\left(v;\theta_0\right)}-\frac{\pi_{2T}\left(z,v\right)}{\pi_{1T}\left(v\right)}\right]\left[\pi_{T,S}\left(z|\,v;\theta_0\right)-\pi\left(z|\,v;\theta_0\right)\right]\\ &\qquad\times\left[w_T\left(z,v\right)-w\left(z,v\right)\right]\mathbb{T}_{T,S}\left(v;\theta_0\right)\nabla_\theta\mathbb{T}_{T,S}\left(v;\theta_0\right)dzdv\\ &\quad-\sqrt{T}\iint \frac{1}{S}\sum_{i=1}^{S}\left[\frac{\pi_{2T}^{i}\left(z,v;\theta_0\right)}{\pi_{1T}^{i}\left(v;\theta_0\right)}-\frac{\pi_{2T}\left(z,v\right)}{\pi_{1T}\left(v\right)}\right]\left[\pi_{T}\left(z|\,v\right)-\pi\left(z|\,v;\theta_0\right)\right]\\ &\qquad\times\left[w_T\left(z,v\right)-w\left(z,v\right)\right]\mathbb{T}_{T,S}\left(v;\theta_0\right)\nabla_\theta\mathbb{T}_{T,S}\left(v;\theta_0\right)dzdv \end{aligned}\tag{C12}$$

Moreover, $E\left(\pi_{2T}\left(z,v\right)\right)=E\left(\pi_{2T}^{i}\left(z,v;\theta_0\right)\right)$. Therefore, at all points of continuity,

$$\begin{aligned}&\frac{\pi_{2T}^{i}\left(z,v;\theta_0\right)}{\pi_{1T}^{i}\left(v;\theta_0\right)}-\frac{\pi_{2T}\left(z,v\right)}{\pi_{1T}\left(v\right)}\\ &=\frac{\pi_{2T}^{i}\left(z,v;\theta_0\right)-E\left(\pi_{2T}^{i}\left(z,v;\theta_0\right)\right)}{\pi_{1T}^{i}\left(v;\theta_0\right)}-\frac{\pi_{2T}(z,v)-E\left(\pi_{2T}\left(z,v\right)\right)}{\pi_{1T}\left(v\right)}+\frac{E\left(\pi_{2T}\left(z,v\right)\right)\left[\pi_{1T}(v)-\pi_{1T}^{i}(v;\theta_0)\right]}{\pi_{1T}^{i}\left(v;\theta_0\right)\cdot\pi_{1T}\left(v\right)}.\end{aligned}\tag{C13}$$

By replacing (C13) into (C12) leaves,

$$\begin{aligned}B_{1T,S} &= \frac{1}{S}\sum_{i=1}^{S}\iint u_{T,S}^{i}\left(z,v\right)\left[w_T\left(z,v\right)-w\left(z,v\right)\right]dA_T^{i}\left(z,v\right)\\ &\quad-\iint u_{T,S}^{0}\left(z,v\right)\left[w_T\left(z,v\right)-w\left(z,v\right)\right]dA_T\left(z,v\right)\\ &\quad+\frac{1}{S}\sum_{i=1}^{S}\iint v_{T,S}^{i}\left(z,v\right)\left[w_T\left(z,v\right)-w\left(z,v\right)\right]\left[dA_T\left(v\right)-dA_T^{i}\left(v\right)\right]\\ &\equiv \breve{B}_{1T,S}^{(1)}+\breve{B}_{1T,S}^{(2)}+\breve{B}_{1T,S}^{(3)},\end{aligned}\tag{C14}$$

where, for $i=0,1,\cdots,S$,

$$\begin{aligned}u_{T,S}^{i}\left(z,v\right) &= \frac{1}{\pi_{1T}^{i}\left(v;\theta_0\right)}\left\{\left[\pi_{T,S}\left(z|\,v;\theta_0\right)-\pi\left(z|\,v;\theta_0\right)\right]-\left[\pi_T\left(z|\,v\right)-\pi\left(z|\,v;\theta_0\right)\right]\right\}\\ &\quad\times\mathbb{T}_{T,S}\left(v;\theta_0\right)\nabla_\theta\mathbb{T}_{T,S}\left(v;\theta_0\right)\end{aligned}\tag{C15}$$

$\left(\pi_{2T}^{0}\left(v;\theta_0\right)=\pi_{2T}\left(v\right)\right)$ and, for $i=1,\cdots,S$,

$$\begin{aligned}v_{T,S}^{i}\left(z,v\right) &= \frac{E\left(\pi_{2T}\left(z,v\right)\right)}{\pi_{1T}^{i}\left(v;\theta_0\right)\pi_{1T}\left(v\right)}\left\{\left[\pi_{T,S}\left(z|\,v;\theta_0\right)-\pi\left(z|\,v;\theta_0\right)\right]-\left[\pi_T\left(z|\,v\right)-\pi\left(z|\,v;\theta_0\right)\right]\right\}\\ &\quad\times\mathbb{T}_{T,S}\left(v;\theta_0\right)\nabla_\theta\mathbb{T}_{T,S}\left(v;\theta_0\right)\end{aligned}\tag{C16}$$

and, finally, similarly as in Section B.2,

$$\begin{aligned}&dA_T^{0}\left(z,v\right)\equiv\sqrt{T}\left[\pi_{2T}\left(z,v\right)-E\left(\pi_{2T}\left(z,v\right)\right)\right]dzdv\\ &dA_T^{i}\left(z,v\right)\equiv\sqrt{T}\left[\pi_{2T}^{i}\left(z,v;\theta_0\right)-E\left(\pi_{2T}^{i}\left(z,v;\theta_0\right)\right)\right]dzdv\quad(i=1,\cdots,S)\\ &dA_T^{0}\left(v\right)\equiv\sqrt{T}\left[\pi_{1T}\left(v\right)-E\left(\pi_{1T}\left(v\right)\right)\right]dv\\ &dA_T^{i}\left(v\right)\equiv\sqrt{T}\left[\pi_{1T}^{i}\left(v;\theta_0\right)-E\left(\pi_{1T}^{i}\left(v;\theta_0\right)\right)\right]dv\qquad(i=1,\cdots,S)\end{aligned}\tag{C17}$$

In particular, and as in Section B.2, $A_T^i(z,v) \equiv \int_{-\infty}^{z}\int_{-\infty}^{v} dA_T^i(s',s)$ and $A_T^i(v) \equiv \int_{-\infty}^{v} dA_T^i(s)$ converge weakly to Gaussian processes.

We now study the terms $u_{T,S}^i$ in eqs. (C14) and (C15). By Remark 4,

$$
\begin{aligned}
&\left|u_{T,S}^i(z,v)\right| \\
&\equiv \frac{1}{\pi_{1T}^i(v;\theta_0)}\left\{\left|\pi_{T,S}(z|v;\theta_0)-\pi(z|v;\theta_0)\right|+\left|\pi_T(z|v)-\pi(z|v;\theta_0)\right|\right\}\mathbb{T}_{T,S}(v;\theta_0)\left|\nabla_\theta\mathbb{T}_{T,S}(v;\theta_0)\right| \\
&\leq \frac{1}{\pi_{1T}^i(v;\theta_0)}\left\{\left|\pi_{T,S}(z|v;\theta_0)-\pi(z|v;\theta_0)\right|+\left|\pi_T(z|v)-\pi(z|v;\theta_0)\right|\right\}\mathbb{T}_{T,S}(v;\theta_0) \\
&\times\sum_{j=1}^{S}\left|\nabla_\theta\pi_{1T}^j(v;\theta_0)-\nabla_\theta\pi_1(v;\theta_0)\right| g_{\delta_T}\left(\pi_{1T}^j(v;\theta_0)\right)G_{\delta_T}^{(-j)}(v;\theta_0) \\
&+\frac{1}{\pi_{1T}^i(v;\theta_0)}\left\{\left|\pi_{T,S}(z|v;\theta_0)-\pi(z|v;\theta_0)\right|+\left|\pi_T(z|v)-\pi(z|v;\theta_0)\right|\right\}\mathbb{T}_{T,S}(v;\theta_0) \\
&\times\sum_{j=1}^{S}\left|\nabla_\theta\pi_1(v;\theta_0)\right| g_{\delta_T}\left(\pi_{1T}^j(v;\theta_0)\right)G_{\delta_T}^{(-j)}(v;\theta_0) \\
&\equiv b_{1T,S}(z,v)+b_{2T,S}(z,v).
\end{aligned}
\tag{C18}
$$

By Assumption 8, $g_{\delta_T}=\delta_T^{-1}\times\breve{k}_T$, where $\breve{k}_T$ is bounded in probability. Therefore, by Lemma C1 (see eq. (A3b)) and Lemma N1,

$$
\begin{aligned}
&\sup_{(z,v)\in\mathbb{R}^q} b_{1T,S}(z,v) \\
&=\delta_T^{-1}\times\left[O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q}\delta_T^{-1}\right)+O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\delta_T^{-2}\right)+O_p\left(\lambda_T^r\delta_T^{-2}\right)\right]\times\left[O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)-1}\right)+O_p\left(\lambda_T^r\right)\right] \\
&\overset{p}{\to}\mathbf{0}_n,
\end{aligned}
\tag{C19}
$$

where the convergence follows by Assumption 9 in Al-M, and the Remark 1 in Section A.3 of the present Appendix. By similar arguments, and boundedness of $\nabla_\theta\pi_1(v;\theta_0)$ (Assumption 1-(b)), it follows that $\sup_{(z,v)\in\mathbb{R}^q} b_{2T,S}(z,v)\overset{p}{\to}\mathbf{0}_n$. Therefore, by eq. (C18),

$$
\sup_{(z,v)\in\mathbb{R}^q}\left|u_{T,S}^i(z,v)\right|\overset{p}{\to}\mathbf{0}_n,\quad i=0,1,\cdots,S. \tag{C20}
$$

Eq. (C20), combined with Assumption 4 on the weighting function, imply that $\left(\breve{B}_{1T,S}^{(1)}+\breve{B}_{1T,S}^{(2)}\right)\overset{p}{\to}\mathbf{0}_n$ in (C14). Next, we show that in (C14), $\breve{B}_{1T,S}^{(3)}\overset{p}{\to}\mathbf{0}_n$ as well. We have,

$$
\begin{aligned}
&\left|v_{T,S}^i(z,v)\right| \\
&\leq\frac{E\left(\pi_{2T}(z,v)\right)}{\pi_{1T}^i(v;\theta_0)\,\pi_{1T}(v)}\left\{\left|\pi_{T,S}(z|v;\theta_0)-\pi(z|v;\theta_0)\right|+\left|\pi_T(z|v)-\pi(z|v;\theta_0)\right|\right\} \\
&\times\mathbb{T}_{T,S}(v;\theta_0)\left|\nabla_\theta\mathbb{T}_{T,S}(v;\theta_0)\right|.
\end{aligned}
$$

By Remark 4,

$$
\begin{aligned}
&\frac{1}{\pi_{1T}^{i}\left(v;\theta_{0}\right)\cdot\pi_{1T}\left(v\right)}\left\{\left|\pi_{T,S}\left(z\right|v;\theta_{0}\right)-\pi\left(z\right|v;\theta_{0}\right)\right|+\left|\pi_{T}\left(z\right|v\right)-\pi\left(z\right|v;\theta_{0}\right)\right|\right\} \\
&\qquad\times\mathbb{T}_{T,S}\left(v;\theta_{0}\right)\left|\nabla_{\theta}\mathbb{T}_{T,S}\left(v;\theta_{0}\right)\right| \\
&\leq\frac{1}{\pi_{1T}^{i}\left(v;\theta_{0}\right)\cdot\pi_{1T}\left(v\right)}\left\{\left|\pi_{T,S}\left(z\right|v;\theta_{0}\right)-\pi\left(z\right|v;\theta_{0}\right)\right|+\left|\pi_{T}\left(z\right|v\right)-\pi\left(z\right|v;\theta_{0}\right)\right|\right\} \\
&\qquad\times\mathbb{T}_{T,S}\left(v;\theta_{0}\right)\sum_{j=1}^{S}\left|\nabla_{\theta}\pi_{1T}^{j}\left(v;\theta_{0}\right)-\nabla_{\theta}\pi_{1}\left(v;\theta_{0}\right)\right|g_{\delta_{T}}\left(\pi_{1T}^{j}\left(v;\theta_{0}\right)\right)G_{\delta_{T}}^{(-j)}\left(v;\theta_{0}\right) \\
&+\frac{1}{\pi_{1T}^{i}\left(v;\theta_{0}\right)\pi_{1T}\left(v\right)}\left\{\left|\pi_{T,S}\left(z\right|v;\theta_{0}\right)-\pi\left(z\right|v;\theta_{0}\right)\right|+\left|\pi_{T}\left(z\right|v\right)-\pi\left(z\right|v;\theta_{0}\right)\right|\right\} \\
&\qquad\times\mathbb{T}_{T,S}\left(v;\theta_{0}\right)\sum_{j=1}^{S}\left|\nabla_{\theta}\pi_{1}\left(v;\theta_{0}\right)\right|g_{\delta_{T}}\left(\pi_{1T}^{j}\left(v;\theta_{0}\right)\right)G_{\delta_{T}}^{(-j)}\left(v;\theta_{0}\right) \\
&\equiv b_{1T,S}^{*}\left(z,v\right)+b_{2T,S}^{*}\left(z,v\right). \qquad\qquad (C21)
\end{aligned}
$$

By arguments nearly identical to those we used to show (C20), we have that:

$$
\begin{aligned}
&\sup_{(z,v)\in\mathbb{R}^{q}}b_{1T,S}^{*}\left(z,v\right) \\
&=\delta_{T}^{-2}\times\left[O_{p}\left(T^{-\frac{1}{2}}\lambda_{T}^{-q}\delta_{T}^{-1}\right)+O_{p}\left(T^{-\frac{1}{2}}\lambda_{T}^{-(q-q^{*})}\delta_{T}^{-2}\right)+O_{p}\left(\lambda_{T}^{r}\delta_{T}^{-2}\right)\right]\times\left[O_{p}\left(T^{-\frac{1}{2}}\lambda_{T}^{-(q-q^{*})-1}\right)+O_{p}\left(\lambda_{T}^{r}\right)\right] \\
&\xrightarrow{p}\mathbf{0}_{n}, \qquad\qquad (C22)
\end{aligned}
$$

where the convergence follows by Assumption 9 in Al-M, and the Remark 1 in Section A.3 of the present Appendix. By similar arguments, and boundedness of $\nabla_{\theta}\pi_{1}\left(v;\theta_{0}\right)$ (Assumption 1-(b)), we also have that $\sup_{(z,v)\in\mathbb{R}^{q}}b_{2T,S}^{*}\left(z,v\right)\xrightarrow{p}\mathbf{0}_{n}$. It follows that $\sup_{(z,v)\in\mathbb{R}^{q}}\left|\upsilon_{T,S}^{i}\left(z,v\right)\right|\xrightarrow{p}\mathbf{0}_{n}$ and, by Assumption 4 on the weighting function, that $\breve{B}_{1T,S}^{(3)}\xrightarrow{p}\mathbf{0}_{n}$ in (C14). We have thus established that $B_{1T,S}\xrightarrow{p}\mathbf{0}_{n}$ in (C14).

By eq. (C11), we then have that $B_{T,S}\xrightarrow{p}\mathbf{0}_{n}$ whenever $B_{2T,S}\xrightarrow{p}\mathbf{0}_{n}$. We now show this is the case. Indeed, the $B_{2T,S}$ term can be re-written as:

$$
\begin{aligned}
B_{2T,S}&=\frac{1}{S}\sum_{i=1}^{S}\iint u_{T,S}^{i}\left(z,v\right)w\left(z,v\right)dA_{T}^{i}\left(z,v\right)-\iint u_{T,S}^{0}\left(z,v\right)w\left(z,v\right)dA_{T}\left(z,v\right) \\
&\quad+\frac{1}{S}\sum_{i=1}^{S}\iint\upsilon_{T,S}^{i}\left(z,v\right)w\left(z,v\right)\left[dA_{T}\left(v\right)-dA_{T}^{i}\left(v\right)\right]. \qquad\qquad (C23)
\end{aligned}
$$

where the functions $u_{T,S}^{i}\left(z,v\right)$ and $\upsilon_{T,S}^{i}\left(z,v\right)$ are as in eqs. (C15) and (C16). Therefore, $B_{2T,S}\xrightarrow{p}\mathbf{0}_{n}$ follows by boundedness of the limiting weighting function $w\left(z,v\right)$ (Assumption 4) and the same arguments used to show that $B_{1T,S}\xrightarrow{p}\mathbf{0}_{n}$. This proves that the convergence in (C10) of Claim 1 holds true.

Analysis of the $C_{T,S}$ term

Claim 2. *We have,*

$$C_{T,S} = \iint \left|\nabla_\theta \pi_{T,S}(z|\, v;\theta_0)\mathbb{T}_{T,S}(v;\theta_0)\right|_2 w_T(z,v)dzdv + o_p(1). \tag{C24}$$

To prove Claim 2, note that the term $C_{T,S}$ in (C9) is:

$$C_{T,S} = C_{1T,S} + C_{2T,S}, \tag{C25}$$

where

$$\begin{aligned} C_{1T,S} &\equiv \iint \nabla_\theta \left\{ \left[\pi_{T,S}\left(z|\, v;\bar\theta\right) - \pi_T\left(z|\, v\right)\right] \nabla_\theta \pi_{T,S}\left(z|\, v;\bar\theta\right) \mathbb{T}^2_{T,S}\left(v;\bar\theta\right)\right\} w_T(z,v)\, dzdv; \\ C_{2T,S} &\equiv \iint \nabla_\theta \left\{ \left[\pi_{T,S}\left(z|\, v;\bar\theta\right) - \pi_T\left(z|\, v\right)\right]^2 \mathbb{T}_{T,S}\left(v;\bar\theta\right) \nabla_\theta \mathbb{T}_{T,S}\left(v;\bar\theta\right)\right\} w_T(z,v)\, dzdv. \end{aligned}$$

We study these two integrals separately.

- By performing the inner differentiation,

$$\begin{aligned} & C_{1T,S} \\ &= \iint \left\{\left|\nabla_\theta \pi_{T,S}\left(z|\, v;\bar\theta\right)\right|_2 + \left[\pi_{T,S}\left(z|\, v;\bar\theta\right) - \pi_T\left(z|\, v\right)\right] \nabla_{\theta\theta}\pi_{T,S}\left(z|\, v;\bar\theta\right)\right\} \mathbb{T}^2_{T,S}\left(v;\bar\theta\right) w_T(z,v)\, dzdv \\ &+ 2\iint \left[\pi_{T,S}\left(z|\, v;\bar\theta\right) - \pi_T\left(z|\, v\right)\right] \nabla_\theta \pi_{T,S}\left(z|\, v;\bar\theta\right) \mathbb{T}_{T,S}\left(v;\bar\theta\right) \nabla_\theta \mathbb{T}_{T,S}\left(v;\bar\theta\right) w_T(z,v)\, dzdv \\ &\equiv C_{11T,S} + C_{12T,S}. \end{aligned} \tag{C26}$$

We have,

$$\begin{aligned} C_{11T,S} = & \iint \left|\nabla_\theta \pi_{T,S}\left(z|\, v;\theta_0\right) \mathbb{T}_{T,S}\left(v;\theta_0\right)\right|_2 w_T(z,v)\, dzdv \\ &+ \iint \left[\pi\left(z|\, v;\theta_0\right) - \pi_T\left(z|\, v\right)\right] \nabla_{\theta\theta}\pi_{T,S}\left(z|\, v;\bar\theta\right) \mathbb{T}^2_{T,S}\left(v;\theta_0\right) w_T(z,v)\, dzdv \\ &+ \iint \left[\pi\left(z|\, v;\bar\theta\right) - \pi\left(z|\, v;\theta_0\right)\right] \nabla_{\theta\theta}\pi_{T,S}\left(z|\, v;\bar\theta\right) \mathbb{T}^2_{T,S}\left(v;\theta_0\right) w_T(z,v)\, dzdv \\ &+ \iint \left[\pi_{T,S}\left(z|\, v;\bar\theta\right) - \pi\left(z|\, v;\bar\theta\right)\right] \nabla_{\theta\theta}\pi_{T,S}\left(z|\, v;\bar\theta\right) \mathbb{T}^2_{T,S}\left(v;\theta_0\right) w_T(z,v)\, dzdv + o_p(1). \end{aligned}$$

By Remark 5, Assumptions 1-(a), 3 and 7, by Θ compact, and by Lemma C1, we have that the last three terms converge in probability to zero, under the bandwidth conditions of Assumptions 9-(b). Hence,

$$C_{11T,S} = \iint \left|\nabla_\theta \pi_{T,S}\left(z|\, v;\theta_0\right) \mathbb{T}_{T,S}\left(v;\theta_0\right)\right|_2 w_T(z,v)\, dzdv + o_p(1). \tag{C27}$$

As regards the $C_{12T,S}$ term in (C26), we have that $C_{12T,S} \xrightarrow{p} \mathbf{0}_{n\times n}$ by the same arguments, and the additional arguments we made to show that $B_{1T,S} \xrightarrow{p} \mathbf{0}_n$ in eq. (C14) above. Hence, $C_{1T,S} = C_{11T,S} + o_p(1)$, where $C_{11T,S}$ is given in eq. (C27).

- By differentiating,

$$
\begin{aligned}
&C_{2T,S}\\
&= 2\iint \left[\pi_{T,S}\left(z|\,v;\bar{\theta}\right) - \pi_T\left(z|\,v\right)\right] \nabla_\theta \pi_{T,S}\left(z|\,v;\bar{\theta}\right) \mathbb{T}_{T,S}\left(v;\bar{\theta}\right) \nabla_\theta \mathbb{T}_{T,S}\left(v;\bar{\theta}\right) w_T\left(z,v\right) dzdv\\
&+ \iint \left[\pi_{T,S}\left(z|\,v;\bar{\theta}\right) - \pi_T\left(z|\,v\right)\right]^2 \left[\left|\nabla_\theta \mathbb{T}_{T,S}\left(v;\bar{\theta}\right)\right|_2 + \mathbb{T}_{T,S}\left(v;\bar{\theta}\right) \nabla_{\theta\theta} \mathbb{T}_{T,S}\left(v;\bar{\theta}\right)\right] w_T\left(z,v\right) dzdv.
\end{aligned}
\tag{C28}
$$

We have, $C_{2T,S} \xrightarrow{p} \mathbf{0}_{n\times n}$ by the same arguments produced to deal with the $C_{1T,S}$ term, by Remark 4, by $\pi_{1T}^i \leq \delta_T \Rightarrow g = 0$ (by Assumption 8), by noticing again that by Remark 4, $\nabla_{\theta\theta}\mathbb{T}_{T,S}\left(v;\bar{\theta}\right) = \delta_T^{-2}\hat{k}_T$ (where $\hat{k}_T$ is a term bounded in probability) and, finally, by Assumption 9 in Al-M, and the Remark 1 in Section A.3 of the present Appendix.

Therefore, eq. (C24) in Claim 2 follows by eqs. (C25), (C26), (C27) and the previous results that $C_{12T,S} \xrightarrow{p} \mathbf{0}_{n\times n}$ and $C_{2T,S} \xrightarrow{p} \mathbf{0}_{n\times n}$.

<u>First order conditions in eq. (C7)</u>

By the previous results on $B_{T,S}$ (Claim 1, eq. (C10)) and $C_{T,S}$ (Claim 2, eq. (C24)), the first order conditions in eq. (C7) are,

$$
\begin{aligned}
\mathbf{0}_n &= \frac{1}{S}\sum_{i=1}^{S}\sqrt{T}\iint \left[\frac{\pi_{2T}^i\left(z,v;\theta_0\right)}{\pi_{1T}^i\left(v;\theta_0\right)} - \frac{\pi_{2T}\left(z,v\right)}{\pi_{1T}\left(v\right)}\right] \nabla_\theta \pi_{T,S}(z|\,v;\theta_0) w_T(z,v) \mathbb{T}_{T,S}^2\left(v;\theta_0\right) dzdv + o_p(1)\\
&+ \left[\iint \left|\nabla_\theta \pi_{T,S}(z|\,v;\theta_0)\mathbb{T}_{T,S}\left(v;\theta_0\right)\right|_2 w_T(z,v) dzdv + o_p(1)\right]\cdot\sqrt{T}\left(\theta_{T,S}-\theta_0\right).
\end{aligned}
\tag{C29}
$$

We now elaborate on eq. (C29). By replacing (C13) into eq. (C29) we obtain,

$$
\begin{aligned}
&o_p(1)\\
&= \frac{1}{S}\sum_{i=1}^{S}\sqrt{T}\iint \left[\frac{\pi_{2T}^i\left(z,v;\theta_0\right) - E\left(\pi_{2T}^i\left(z,v;\theta_0\right)\right)}{\pi_{1T}^i\left(v;\theta_0\right)}\right] \nabla_\theta \pi_{T,S}(z|\,v;\theta_0) w_T(z,v) \mathbb{T}_{T,S}^2\left(v;\theta_0\right) dzdv\\
&- \frac{1}{S}\sum_{i=1}^{S}\sqrt{T}\iint \left[\frac{\pi_{2T}(z,v) - E\left(\pi_{2T}\left(z,v\right)\right)}{\pi_{1T}\left(v\right)}\right] \nabla_\theta \pi_{T,S}(z|\,v;\theta_0) w_T(z,v) \mathbb{T}_{T,S}^2\left(v;\theta_0\right) dzdv\\
&+ \frac{1}{S}\sum_{i=1}^{S}\sqrt{T}\iint \left[\frac{E\left(\pi_{2T}\left(z,v\right)\right)\left[\pi_{1T}(v) - \pi_{1T}^i(v;\theta_0)\right]}{\pi_{1T}^i\left(v;\theta_0\right)\cdot\pi_{1T}\left(v\right)}\right] \nabla_\theta \pi_{T,S}(z|\,v;\theta_0) w_T(z,v) \mathbb{T}_{T,S}^2\left(v;\theta_0\right) dzdv\\
&+ \left[\iint \left|\nabla_\theta \pi_{T,S}(z|\,v;\theta_0)\mathbb{T}_{T,S}\left(v;\theta_0\right)\right|_2 w_T(z,v) dzdv + o_p(1)\right]\cdot\sqrt{T}\left(\theta_{T,S}-\theta_0\right)\\
&\equiv \frac{1}{S}\sum_{i=1}^{S}\left[\left(\mathcal{I}_{1T,S}^i - \mathcal{I}_{1T,S}^0\right) + \left(\mathcal{I}_{2T,S}^i - \mathcal{I}_{2T,S}^0\right)\right] + \left[\mathcal{J}_{T,S} + o_p(1)\right]\cdot\sqrt{T}\left(\theta_{T,S}-\theta_0\right),
\end{aligned}
\tag{C30}
$$

where

$$\mathcal{J}_{T,S} \equiv \iint \left| \nabla_\theta \pi_{T,S}\left(z|\, v;\theta_0\right) \mathbb{T}_{T,S}\left(v;\theta_0\right)\right|_2 w_T\left(z,v\right) dzdv \tag{C31a}$$

$$\mathcal{I}^i_{1T,S} \equiv \iint \frac{\nabla_\theta \pi_{T,S}(z|\, v;\theta_0) w_T(z,v)}{\pi^i_{1T}\left(v;\theta_0\right)} \mathbb{T}^2_{T,S}\left(v;\theta_0\right) dA^i_T\left(z,v\right) \tag{C31b}$$

$$\mathcal{I}^i_{2T,S} \equiv \iint \frac{\nabla_\theta \pi_{T,S}(z|\, v;\theta_0) E\left(\pi_{2T}\left(z,v\right)\right) w_T(z,v)}{\pi^i_{1T}\left(v;\theta_0\right)\pi_{1T}\left(v\right)} \mathbb{T}^2_{T,S}\left(v;\theta_0\right) dzdA^i_T\left(v\right) \tag{C31c}$$

and $dA^i_T\left(z,v\right)$, and $dA^i_T\left(v\right)$ are as in eq. (C17) above. Also, note that,

$$\mathcal{I}^i_{2T,S} = \int \gamma^i_{T,S}\left(v\right) \mathbb{T}^2_{T,S}\left(v;\theta_0\right) dA^i_T\left(v\right)$$

where, for all $i = 1, \cdots, S$,

$$\gamma^i_{T,S}\left(v\right) \equiv \int \frac{\nabla_\theta \pi_{T,S}(z|\, v;\theta_0) E\left(\pi_{2T}\left(z,v\right)\right) w_T(z,v)}{\pi^i_{1T}\left(v;\theta_0\right)\pi_{1T}\left(v\right)} dz. \tag{C32}$$

Next let $\omega^0_i \circ F\left(z,v;\theta_0\right)$, $i = 0, 1, \cdots, S$, denote independent Gaussian processes. Let, also, $\omega^0 \circ F\left(v;\theta_0\right)$ and $\omega^0_i \circ F\left(v;\theta_0\right)$, $i = 1, \cdots, S$, denote independent Gaussian processes. Finally, let

$$\gamma\left(v\right) \equiv \int \frac{\nabla_\theta \pi(z|\, v;\theta_0) \pi_2\left(z,v;\theta_0\right) w(z,v)}{\pi_1\left(v;\theta_0\right)^2} dz. \tag{C33}$$

We now demonstrate that for $i = 0, 1, \cdots, S$,

$$\mathcal{I}^i_{1T,S} \xrightarrow{d} \mathcal{I}^i_1 \equiv \iint \frac{\nabla_\theta \pi\left(z|\, v;\theta_0\right) w\left(z,v\right)}{\pi_1\left(v;\theta_0\right)} d\omega^0_i\left(F\left(z,v;\theta_0\right)\right); \tag{C34a}$$

that for $i = 0, 1, \cdots, S$,

$$\mathcal{I}^i_{2T,S} \xrightarrow{d} \mathcal{I}^i_2 \equiv \int \gamma\left(v\right) d\omega^0_i\left(F\left(v;\theta_0\right)\right); \tag{C34b}$$

and finally that,

$$\mathcal{J}_{T,S} \xrightarrow{p} \mathcal{J} \equiv \iint \left|\nabla_\theta \pi\left(z|\, v;\theta_0\right)\right|_2 w(z,v) dzdv. \tag{C34c}$$

We show eq. (C34a) for the $\mathcal{I}^0_{1T,S}$ term only, the proof for the other $\mathcal{I}^i_{1T,S}$ terms ($i = 1, \cdots, S$) being identical. We have,

$$\mathcal{I}^0_{1T,S} = \iint \frac{\nabla_\theta \pi\left(z|\, v;\theta_0\right) w\left(z,v\right)}{\pi_1\left(v;\theta_0\right)} dA^i_T\left(z,v\right) + \hat{\mathcal{I}}^0_{1T,S} \tag{C35}$$

where,

$$\begin{aligned}
\hat{\mathcal{I}}^0_{1T,S} \equiv & \iint \left[\frac{\nabla_\theta \pi_{T,S}(z|\, v;\theta_0) w_T(z,v)}{\pi_{1T}\left(v\right)} - \frac{\nabla_\theta \pi(z|\, v;\theta_0) w(z,v)}{\pi_1\left(v;\theta_0\right)} \right] \mathbb{T}^2_{T,S}\left(v;\theta_0\right) dA^0_T\left(z,v\right) \\
& - \iint \frac{\nabla_\theta \pi(z|\, v;\theta_0) w(z,v)}{\pi_1\left(v;\theta_0\right)} \left[1 - \mathbb{T}^2_{T,S}\left(v;\theta_0\right)\right] dA^0_T\left(z,v\right).
\end{aligned}$$

By Lemma N2, and by $\left|1-\mathbb{T}_{T,S}^{2}\left(v;\theta_{0}\right)\right|<1$, $\left|1-\mathbb{T}_{T,S}^{2}\left(v;\theta_{0}\right)\right| \xrightarrow{p} 0$, we have that $\hat{\mathcal{I}}_{1T,S}^{0} \xrightarrow{p} 0$. Eq. (C34a) then follows by eq. (C35) and the continuous mapping theorem, as in eq. (B6) in Section B.2 of the present Appendix.

We now turn to demonstrate the convergence in eq. (C34b). We have,

$$\begin{aligned}\mathcal{I}_{2T,S}^{i} &= \int \gamma\left(v\right) dA_{T}^{i}\left(v\right) + \int \left[\gamma_{T,S}^{i}\left(v\right)-\gamma\left(v\right)\right] \mathbb{T}_{T,S}^{2}\left(v;\theta_{0}\right) dA_{T}^{i}\left(v\right) - \int \left[1-\mathbb{T}_{T,S}^{2}\left(v;\theta_{0}\right)\right] \gamma\left(v\right) dA_{T}^{i}\left(v\right) \\ &\equiv \int \gamma\left(v\right) dA_{T}^{i}\left(v\right) + \mathcal{I}_{2aT,S}^{i} + \mathcal{I}_{2bT,S}^{i}. \end{aligned} \tag{C36}$$

As regards the $\mathcal{I}_{2aT,S}^{i}$ term,

$$\left|\gamma_{T,S}^{i}\left(v\right)-\gamma\left(v\right)\right| \leq \int \Gamma_{T,S}^{i}\left(z,v\right) dz, \tag{C37}$$

where, provided the right hand side of (C37) is finite, we have that for all $(z,v) \in \mathbb{R}^{q^{*}} \times \mathcal{B}_{T}$,

$$\begin{aligned}\Gamma_{T,S}^{i}\left(z,v\right) &\equiv \left|\frac{\nabla_{\theta}\pi_{T,S}\left(z|\,v;\theta_{0}\right) E\left(\pi_{2T}\left(z,v\right)\right) w_{T}\left(z,v\right)}{\pi_{1T}^{i}\left(v;\theta_{0}\right)\pi_{1T}\left(v\right)} - \frac{\nabla_{\theta}\pi\left(z|\,v;\theta_{0}\right)\pi_{2}\left(z,v;\theta_{0}\right) w\left(z,v\right)}{\pi_{1}\left(v;\theta_{0}\right)^{2}}\right| \\ &\leq E\left(\pi_{2T}\left(z,v\right)\right)\cdot \sup_{(z,v)\in\mathbb{R}^{q^{*}}\times\mathcal{B}_{T}} \left|\frac{\nabla_{\theta}\pi_{T,S}\left(z|\,v;\theta_{0}\right)}{\pi_{1T}^{i}\left(v;\theta_{0}\right)\pi_{1T}\left(v\right)} - \frac{\nabla_{\theta}\pi\left(z|\,v;\theta_{0}\right)}{\pi_{1}\left(v;\theta_{0}\right)^{2}}\right| \left|w_{T}\left(z,v\right)-w\left(z,v\right)\right| \\ &+ w\left(z,v\right)\cdot \sup_{(z,v)\in\mathbb{R}^{q^{*}}\times\mathcal{B}_{T}} \left|\frac{\nabla_{\theta}\pi_{T,S}\left(z|\,v;\theta_{0}\right) E\left(\pi_{2T}\left(z,v\right)\right)}{\pi_{1T}^{i}\left(v;\theta_{0}\right)\pi_{1T}\left(v\right)} - \frac{\nabla_{\theta}\pi\left(z|\,v;\theta_{0}\right)\pi_{2}\left(z,v;\theta_{0}\right)}{\pi_{1}\left(v;\theta_{0}\right)^{2}}\right| \\ &+ E\left(\pi_{2T}\left(z,v\right)\right)\cdot \sup_{(z,v)\in\mathbb{R}^{q^{*}}\times\mathcal{B}_{T}} \left|\frac{\nabla_{\theta}\pi\left(z|\,v;\theta_{0}\right)}{\pi_{1}\left(v;\theta_{0}\right)^{2}}\right| \left|w_{T}\left(z,v\right)-w\left(z,v\right)\right| \\ &\equiv \Gamma_{1T,S}^{i}\left(z,v\right) + \Gamma_{2T,S}^{i}\left(z,v\right) + \Gamma_{3T}\left(z,v\right), \end{aligned} \tag{C38}$$

and where $\mathcal{B}_{T}$ is the same trimming set introduced in Lemma C1 of Section A. Since $\int E\left(\pi_{2T}\left(z,v\right)\right) dz$ is bounded and constant, and since $w_{T}\left(z,v\right)$ and $w\left(z,v\right)$ are bounded and integrable, we have, by Lemma N3, that,

$$\sup_{v\in\mathbb{R}^{q-q^{*}}} \int \left[\Gamma_{1T,S}^{i}\left(z,v\right)+\Gamma_{2T,S}^{i}\left(z,v\right)\right] \mathbb{T}_{T,S}^{2}\left(v;\theta_{0}\right) dz \xrightarrow{p} 0. \tag{C39}$$

Moreover,

$$\sup_{(z,v)\in\mathbb{R}^{q^{*}}\times\mathcal{B}_{T}} \left|\frac{\nabla_{\theta}\pi\left(z|\,v;\theta_{0}\right)}{\pi_{1}\left(v;\theta_{0}\right)^{2}}\right| \left|w_{T}\left(z,v\right)-w\left(z,v\right)\right| = O_{p}\left(T^{-\frac{1}{2}}\lambda_{T}^{q}\delta_{T}^{-2}\right) + O_{p}\left(\lambda_{T}^{r}\delta_{T}^{-2}\right) \xrightarrow{p} 0,$$

where the equality follows by Assumption 4-(b) on the weighting function and the convergence follows by the bandwidth conditions in Assumption 9-(b). Thus, $\sup_{v\in\mathbb{R}^{q-q^{*}}} \int \Gamma_{3T}\left(z,v\right)\mathbb{T}_{T,S}^{2}\left(v;\theta_{0}\right) dz \xrightarrow{p} 0$, and, by (C37), (C38) and (C39), $\sup_{v\in\mathcal{B}_{T}} \left|\gamma_{T,S}^{i}\left(v\right)-\gamma\left(v\right)\right| \xrightarrow{p} 0$. Then, by the decomposition in (C36),

$$\mathcal{I}_{2aT,S}^{i} \xrightarrow{p} 0.$$

By the properties of $\mathbb{T}^2_{T,S}(v;\theta_0)$ used to show that $\hat{\mathcal{I}}^0_{1T,S} \xrightarrow{p} 0$, we also have $\mathcal{I}^i_{2bT,S} \xrightarrow{p} 0$, which completes the proof that for all $i = 0, 1, \cdots, S$,

$$\mathcal{I}^i_{2T,S} = \int \gamma(v)\, dA^i_T(v) + o_p(1).$$

The convergence of in (C34b) follows, again, by the continuous mapping theorem.

Finally, eq. (C34c) follows by $\iint |w_T - w| \xrightarrow{p} 0$ and Lemma N1-(ii) and arguments nearly identical to those we used to demonstrate (C34a) and (C34b).

The normality claim in Theorem 2 immediately follows: as in Section B.2, the terms $\mathcal{I}^i_1$, $i = 0, 1, \cdots, S$ in eqs. (C34a)-(C34b) are all independent and asymptotically centered Gaussian. Therefore, $\sqrt{T}(\theta_{T,S} - \theta_0)$ is asymptotically centered normal with variance,

$$V_S \equiv \mathcal{J}^{-1} \cdot var\left(\frac{1}{S}\sum_{i=1}^{S}\left[\left(\mathcal{I}^i_1 - \mathcal{I}^0_1\right) + \left(\mathcal{I}^i_2 - \mathcal{I}^0_2\right)\right]\right) \cdot \mathcal{J}^{\top -1},$$

where $\mathcal{J}$ is defined in eq. (C34c), and

$$\begin{aligned}
& var\left(\frac{1}{S}\sum_{i=1}^{S}\left[\left(\mathcal{I}^i_1 - \mathcal{I}^0_1\right) + \left(\mathcal{I}^i_2 - \mathcal{I}^0_2\right)\right]\right) \\
&= \left(1 + \frac{1}{S}\right)\left[var\left(\mathcal{I}^0_1\right) + var\left(\mathcal{I}^0_2\right)\right] + 2cov\left(\frac{1}{S}\sum_{i=1}^{S}\mathcal{I}^i_1 - \mathcal{I}^0_1, \frac{1}{S}\sum_{i=1}^{S}\mathcal{I}^i_2 - \mathcal{I}^0_2\right) \\
&= \left(1 + \frac{1}{S}\right)\left[var\left(\mathcal{I}^0_1\right) + var\left(\mathcal{I}^0_2\right)\right] + 2\left(1 + \frac{1}{S}\right) cov\left(\mathcal{I}^0_1, \mathcal{I}^0_2\right) \\
&= \left(1 + \frac{1}{S}\right) var\left(\mathcal{I}^0_1 + \mathcal{I}^0_2\right),
\end{aligned}$$

and, by the same computations leading to eq. (B7) in Section B.2,

$$\begin{aligned}
& var\left(\mathcal{I}^0_1 + \mathcal{I}^0_2\right) \\
&= var\left[\Upsilon(z_t, v_t)\right] + \sum_{k=1}^{\infty}\left\{cov\left[\Upsilon(z_t, v_t), \Upsilon(z_{t+k}, v_{t+k})\right] + cov\left[\Upsilon(z_{t+k}, v_{t+k}), \Upsilon(z_t, v_t)\right]\right\},
\end{aligned} \tag{C40}$$

where

$$\Upsilon(z, v) = \frac{\nabla_\theta \pi(z|\, v;\theta_0)\, w(z, v)}{\pi_1(v;\theta_0)} + \gamma(v),$$

and $\gamma(v)$ is as in eq. (C33). Finally, these variance terms are finite by the mixing condition in Assumption 2 and the assumption that $E[\|\Upsilon(z_t, v_t)\|^{\vartheta}]^{1/\vartheta} < \infty$, for some $\vartheta > 2$, by, e.g., Politis and Romano (1994) (Thm. 2.3 p. 466).

C.3 The estimator in Al-M, Section 2.3

In Al-M (Section 2.3), we defined the following estimator,

$$\begin{aligned}\hat{\theta}_{T,S} &= \arg\min_{\theta\in\Theta} \mathcal{L}_{T,S}(\theta) \\ &\equiv \arg\min_{\theta\in\Theta} \sum_{k=1}^{l} \int_{\mathbb{R}^{2q^*}} \left[\pi_{T,S}\left(y^o \middle| y^o_{-k};\theta\right) - \pi_T\left(y^o \middle| y^o_{-k}\right)\right]^2 w_T\left(y^o, y^o_{-k}\right) \mathbb{T}^2_{T,S}\left(y^o_{-k};\theta\right) dy^o dy^o_{-k},\end{aligned} \tag{C41}$$

where $\pi_T\left(y^o \middle| y^o_{-k}\right)$ is the estimate of the conditional density of two observations that are k lags apart in eq. (9) of Al-M, $\pi_{T,S}\left(y^o \middle| y^o_{-k};\theta\right)$ is its simulated counterpart, and $\mathbb{T}_{T,S}\left(y^o_{-k};\theta\right)$ is a trimming function.

The regularity conditions we need to establish consistency and asymptotic normality of $\hat{\theta}_{T,S}$ mirror those in Al-M: In addition to Al-M (Assumptions 1 and 2), we require that the kernels K_{2q^*} and K_{q^*} satisfy the same conditions as K_q and K_{q-q^*} in Al-M (Assumption 3), where K_d is a d-dimensional kernel (as defined by Al-M (Section 2.2)); the trimming function w_T satisfies Al-M (Assumption 4); the criterion $\mathcal{L}_{T,S}(\theta)$ in (C41) satisfies the same conditions satisfied by $L^{\mathrm{CD}}_{T,S}(\theta)$ in Al-M (Assumptions 5 and 6); the gradients, $\nabla_\theta K_{2q^*}(\cdot)$ and $\nabla_\theta K_{q^*}(\cdot)$, satisfy the same conditions satisfied by $\nabla_\theta K_q(\cdot)$ and $\nabla_\theta K_{q-q^*}(\cdot)$ in Al-M (Assumption 7); finally, with $2q^*$ replacing q, the trimming function $\mathbb{T}_{T,S}\left(y^o_{-k};\theta\right)$ satisfies the same conditions satisfied by $\mathbb{T}_{T,S}(v;\theta)$ in Al-M (Assumptions 8 and 9).

It is straight forward to see that consistency of $\hat{\theta}_{T,S}$ obtains under the same conditions in Al-M (as modified above), by arguments nearly identical to those in Section C.1 of the present Appendix. As regards asymptotic normality, it is also simple to find that under the previous conditions, the first order conditions satisfied by $\hat{\theta}_{T,S}$ are, asymptotically:

$$\begin{aligned}0 &= \sum_{k=1}^{l} \int_{\mathbb{R}^{2q^*}} \sqrt{T}\left[\pi_{T,S}\left(y^o \middle| y^o_{-k};\theta_0\right) - \pi_T\left(y^o \middle| y^o_{-k}\right)\right] \nabla_\theta \pi_{T,S}\left(y^o \middle| y^o_{-k};\theta_0\right) w_T\left(y^o, y^o_{-k}\right) \mathbb{T}^2_{T,S}\left(y^o_{-k};\theta_0\right) dy^o dy^o_{-k} \\ &+ \left[\sum_{k=1}^{l} \int_{\mathbb{R}^{2q^*}} \left|\nabla_\theta \pi_{T,S}(y^o | y^o_{-k};\theta_0) \mathbb{T}_{T,S}\left(y^o_{-k};\theta_0\right)\right|_2 w_T\left(y^o, y^o_{-k}\right) dy^o dy^o_{-k}\right] \sqrt{T}(\hat{\theta}_{T,S} - \theta_0) + o_p(1).\end{aligned}$$

Next, let $\pi_{2T}\left(y^o, y^o_{-k}\right)$ (resp. $\pi_{1T}\left(y^o_{-k}\right)$) be the estimate of the joint density of two observations that are k lags apart in eq. (9) of Al-M (resp. the marginal density of the observables), and $\pi^i_{2T}\left(y^o, y^o_{-k};\theta\right)$ (resp. $\pi^i_{1T}\left(y^o_{-k}\right)$) be the simulated counterpart, obtained at the i-th simulation, for $i = 1, \cdots, S$. By the usual arguments in Sections B.2 and C.2 of the present Appendix, we have that:

$$-\hat{\mathcal{J}}_{T,S}\sqrt{T}\left(\theta_{T,S} - \theta_0\right) = \sum_{k=1}^{l} \left(\frac{1}{S}\sum_{i=1}^{S}\left(\breve{\mathcal{I}}^i_{1T,k} + \breve{\mathcal{I}}^i_{2T,k}\right) - \left(\breve{\mathcal{I}}^0_{1T,k} + \breve{\mathcal{I}}^0_{2T,k}\right)\right) + o_p(1),$$

where

$$\hat{\mathcal{J}}_{T,S} \equiv \sum_{k=1}^{l} \int_{\mathbb{R}^{2q^*}} \left| \nabla_\theta \pi_{T,S}(y^o | \, y^o_{-k}; \theta_0) \mathbb{T}_{T,S} \left(y^o_{-k}; \theta_0 \right) \right|_2 w_T \left(y^o, y^o_{-k} \right) dy^o dy^o_{-k};$$

$$\breve{\mathcal{I}}^i_{1T,k} \equiv \iint \eta \left(y^o, y^o_{-k} \right) dA^i_T \left(y^o, y^o_{-k} \right) \quad \text{and} \quad \breve{\mathcal{I}}^i_{2T,k} \equiv \int \gamma \left(y^o_{-k} \right) dA^i_T \left(y^o_{-k} \right), \quad i = 0, 1, \cdots, S;$$

$$\eta \left(y^o, y^o_{-k} \right) = \frac{\nabla_\theta \pi \left(y^o | \, y^o_{-k}; \theta_0 \right) w \left(y^o, y^o_{-k} \right)}{\pi_1 \left(y^o_{-k}; \theta_0 \right)}; \quad \gamma \left(y^o_{-k} \right) = \int \frac{\nabla_\theta \pi(y^o | \, y^o_{-k}; \theta_0) \pi_2 \left(y^o, y^o_{-k}; \theta_0 \right) w(y^o, y^o_{-k})}{\pi_1 \left(y^o_{-k}; \theta_0 \right)^2} dy^o;$$

and, for, $i = 0, 1, \cdots, S$, $dA^i_T \left(y^o, y^o_{-k} \right) \equiv \sqrt{T} \left[\pi^i_{2T} \left(y^o, y^o_{-k}; \theta_0 \right) - E \left(\pi^i_{2T} \left(y^o, y^o_{-k}; \theta_0 \right) \right) \right] dy^o dy^o_{-k}$ and $dA^i_T \left(y^o_{-k} \right) \equiv \sqrt{T} \left[\pi^i_{1T} \left(y^o_{-k}; \theta_0 \right) - E \left(\pi^i_{1T} \left(y^o_{-k}; \theta_0 \right) \right) \right] dy^o_{-k}$, with $\pi^0_{2T} \left(y^o, y^o_{-k}; \theta_0 \right)$ denoting the joint density estimate on sample data, $\pi_{2T} \left(y^o, y^o_{-k} \right)$; and $\pi^0_{1T} \left(y^o_{-k}; \theta_0 \right)$ denoting the marginal density estimate on sample data, $\pi_{1T} \left(y^o_{-k} \right)$. Under the previous conditions, and arguments nearly identical to those in Sections B.2 and C.2 of the present Appendix,

$$\hat{\mathcal{J}}_{T,S} \xrightarrow{p} \hat{\mathcal{J}} \equiv \sum_{k=1}^{l} \int_{\mathbb{R}^{2q^*}} \left| \nabla_\theta \pi(y^o | \, y^o_{-k}; \theta_0) \right|_2 w \left(y^o, y^o_{-k} \right) dy^o dy^o_{-k};$$

and for $i = 0, 1, \cdots, S$,

$$\begin{aligned} \breve{\mathcal{I}}^i_{1T,k} \xrightarrow{p} \breve{\mathcal{I}}^i_{1,k} &\equiv \iint \eta \left(y^o, y^o_{-k} \right) d\omega^0_i \left(F \left(y^o, y^o_{-k}; \theta_0 \right) \right) \\ \breve{\mathcal{I}}^i_{2T,k} \xrightarrow{p} \breve{\mathcal{I}}^i_{2,k} &\equiv \int \gamma \left(y^o_{-k} \right) d\omega^0_i \left(F \left(y^o_{-k}; \theta_0 \right) \right) \end{aligned}$$

where $\omega^0_i \circ F \left(y^o, y^o_{-k}; \theta_0 \right)$ and $\omega^0_i \circ F \left(y^o_{-k}; \theta_0 \right)$ denote independent Gaussian processes. Therefore, $\sqrt{T}(\hat{\theta}_{T,S} - \theta_0)$ is asymptotically centered normal with variance

$$\hat{V}_S = \hat{\mathcal{J}}^{-1} var \left[\sum_{k=1}^{l} \left(\frac{1}{S} \sum_{i=1}^{S} \left(\breve{\mathcal{I}}^i_{1,k} + \breve{\mathcal{I}}^i_{2,k} \right) - \left(\breve{\mathcal{I}}^0_{1,k} + \breve{\mathcal{I}}^0_{2,k} \right) \right) \right] \hat{\mathcal{J}}^{\top -1},$$

where

$$\begin{aligned} & var \left[\sum_{k=1}^{l} \left(\frac{1}{S} \sum_{i=1}^{S} \left(\breve{\mathcal{I}}^i_{1,k} + \breve{\mathcal{I}}^i_{2,k} \right) - \left(\breve{\mathcal{I}}^0_{1,k} + \breve{\mathcal{I}}^0_{2,k} \right) \right) \right] \\ & = var \left[\sum_{k=1}^{l} \left(\frac{1}{S} \sum_{i=1}^{S} \left(\breve{\mathcal{I}}^i_{1,k} + \breve{\mathcal{I}}^i_{2,k} \right) - \left(\breve{\mathcal{I}}^0_{1,k} + \breve{\mathcal{I}}^0_{2,k} \right) \right) \right] \\ & = var \left[\frac{1}{S} \sum_{i=1}^{S} \left(\sum_{k=1}^{l} \left(\breve{\mathcal{I}}^i_{1,k} - \breve{\mathcal{I}}^0_{1,k} \right) + \sum_{k=1}^{l} \left(\breve{\mathcal{I}}^i_{2,k} - \breve{\mathcal{I}}^0_{2,k} \right) \right) \right] \\ & = \left(1 + \frac{1}{S} \right) \left[var \left(\sum_{k=1}^{l} \breve{\mathcal{I}}^0_{1,k} \right) + var \left(\sum_{k=1}^{l} \breve{\mathcal{I}}^0_{2,k} \right) \right] + 2 cov \left[\sum_{k=1}^{l} \left(\frac{1}{S} \sum_{i=1}^{S} \breve{\mathcal{I}}^i_{1,k} - \breve{\mathcal{I}}^0_{1,k} \right), \sum_{k=1}^{l} \left(\frac{1}{S} \sum_{i=1}^{S} \breve{\mathcal{I}}^i_{2,k} - \breve{\mathcal{I}}^0_{2,k} \right) \right] \\ & = \left(1 + \frac{1}{S} \right) \left[var \left(\sum_{k=1}^{l} \breve{\mathcal{I}}^0_{1,k} \right) + var \left(\sum_{k=1}^{l} \breve{\mathcal{I}}^0_{2,k} \right) \right] + 2 \left(1 + \frac{1}{S} \right) cov \left(\sum_{k=1}^{l} \breve{\mathcal{I}}^0_{1,k}, \sum_{k=1}^{l} \breve{\mathcal{I}}^0_{2,k} \right) \\ & = \left(1 + \frac{1}{S} \right) var \left(\sum_{k=1}^{l} \left(\breve{\mathcal{I}}^0_{1,k} + \breve{\mathcal{I}}^0_{2,k} \right) \right), \end{aligned}$$

and where by the same computations leading to eqs. (B7) and (C40) in Sections B.2 and C.2 of the present Appendix,

$$var\left(\sum_{k=1}^{l}\left(\breve{\mathcal{I}}_{1,k}^{0}+\breve{\mathcal{I}}_{2,k}^{0}\right)\right)$$
$$=var\left(\hat{\Upsilon}\left(y_{t}^{o},\cdots,y_{t-l}^{o}\right)\right)$$
$$+\sum_{k=1}^{\infty}\left[cov\left(\hat{\Upsilon}\left(y_{t}^{o},\cdots,y_{t-l}^{o}\right),\hat{\Upsilon}\left(y_{t+k}^{o},\cdots,y_{t+k-l}^{o}\right)\right)+cov\left(\hat{\Upsilon}\left(y_{t+k}^{o},\cdots,y_{t+k-l}^{o}\right),\hat{\Upsilon}\left(y_{t}^{o},\cdots,y_{t-l}^{o}\right)\right)\right],$$

and

$$\hat{\Upsilon}\left(y_{t}^{o},\cdots,y_{t-l}^{o}\right)=\sum_{k=1}^{l}\left[\eta\left(y_{t}^{o},y_{t-k}^{o}\right)+\gamma\left(y_{t-k}^{o}\right)\right].$$

D. Asymptotics for the efficient CD-SNE

D.1 Consistency

We produce the arguments that apply to the case in which the bandwidth sequence satisfies Assumption 10-(a) of Al-M. Accordingly, we will make a repeated use of Lemmas C1-(a), C2-(a) and C3-(a). The case of a bandwidth sequence that satisfies Assumption 10-(b) of Al-M (which is used to prove asymptotic normality in Section D.2 below) is dealt with similarly, by replacing Lemma Cℓ-(a) with Lemma Cℓ-(b), $\ell = 1, 2, 3$.

By the remarks on the proof of Proposition 1 in Section B.1 (Remark 3), and eq. (B2), we only have to show that for all $\theta \in \Theta$,

$$\iint a_{iT,S}(z, v; \theta)\, dzdv \xrightarrow{p} 0, \quad i = 1, 2, \tag{D1}$$

where the terms $a_{iT,S}$ are defined as in Section C.1 (eqs. (C2)), but with weighting function $w_T(z, v) = [\pi_{1T}(v)/\pi_T(z|v)]\,\mathbb{T}_{2T}(z, v)$ and $w(z, v) = m_1(v; \theta_0)/m(z|v; \theta_0)$. We proceed as in Section C.1, and study these two integrals separately.

- For all $(z, v, \theta) \in \mathbb{R}^{q^*} \times \mathbb{R}^{q-q^*} \times \Theta$,

$$\begin{aligned}
a_{1T,S}(z, v; \theta) &\leq |\pi_{T,S}(z|v; \theta) - \pi_T(z|v)|\, \mathbb{T}_{T,S}(v; \theta)\, m_2(z, v; \theta_0)\, |m(z|v; \theta) - m(z|v; \theta_0)| \\
&\quad \times \frac{1}{m_2(z, v; \theta_0)} \left| \frac{\pi_{1T}(v)}{\pi_T(z|v)} - \frac{m_1(v; \theta_0)}{m(z|v; \theta_0)} \right| \mathbb{T}_{2T}(z, v) \\
&\quad + |\pi_{T,S}(z|v; \theta) - \pi_T(z|v)|\, \mathbb{T}_{T,S}(v; \theta)\, |m(z|v; \theta) - m(z|v; \theta_0)| \frac{m_1(v; \theta_0)}{m(z|v; \theta_0)} \\
&\quad \times [1 - \mathbb{T}_{2T}(z, v)] \\
&\leq \ell_{1T,S}(z, v; \theta) \cdot \ell_{2T}(z, v; \theta) \cdot m_2(z, v; \theta_0)\, |m(z|v; \theta) - m(z|v; \theta_0)| \\
&\quad + \ell_{2T}(z, v; \theta) \cdot [m(z|v; \theta) - m(z|v; \theta_0)]^2\, m_2(z, v; \theta_0)\, \mathbb{T}_{T,S}(v; \theta) \\
&\quad + \ell_{3T,S}(z, v; \theta)\, |m(z|v; \theta) - m(z|v; \theta_0)|\, m_2(z, v; \theta_0)\, m_1(v; \theta_0)\, [1 - \mathbb{T}_{2T}(z, v)] \\
&\quad + \frac{m_1(v; \theta_0)}{m(z|v; \theta_0)} [m(z|v; \theta) - m(z|v; \theta_0)]^2\, \mathbb{T}_{T,S}(v; \theta)\, [1 - \mathbb{T}_{2T}(z, v)] \\
&\equiv a_{11T,S}(z, v; \theta) + a_{12T,S}(z, v; \theta) + a_{13T,S}(z, v; \theta) + a_{14T,S}(z, v; \theta),
\end{aligned}$$

where

$$\begin{aligned}
\ell_{1T,S}(z, v; \theta) &\equiv \left[\frac{1}{S} \sum_{i=1}^{S} \left| \pi_T^i(z|v; \theta) - m(z|v; \theta) \right| + |\pi_T(z|v) - m(z|v; \theta_0)| \right] \mathbb{T}_{T,S}(v; \theta) \\
\ell_{2T}(z, v; \theta) &\equiv \frac{1}{m_2(z, v; \theta_0)} \left| \frac{\pi_{1T}(v)}{\pi_T(z|v)} - \frac{m_1(v; \theta_0)}{m(z|v; \theta_0)} \right| \mathbb{T}_{2T}(z, v) \\
\ell_{3T,S}(z, v; \theta) &\equiv \frac{\ell_{1T,S}(z, v; \theta)}{m_2(z, v; \theta_0)\, m(z|v; \theta_0)}
\end{aligned}$$

We have that for all $\theta \in \Theta$: (i) $\iint a_{11T,S} \xrightarrow{p} 0$, by Lemma C1-(a) in Appendix A.1, by Lemma C2, and by boundedness and integrability of the function $m_2(z, v; \theta)\, |m(z|v; \theta) - m(z|v; \theta_0)|$;

(ii) $\iint a_{12T,S} \xrightarrow{p} 0$, by Lemma C2-(a), and by boundedness and integrability of the function $[m(z|v;\theta) - m(z|v;\theta_0)]^2 m_2(z,v;\theta)$; (iii) $\iint a_{13T,S} \xrightarrow{p} 0$, by Lemmas C1-(a), C2-(a) and C3-(a). As regards the $\iint a_{14T,S}$ term, note that the function,

$$m(z|v;\theta_0)^{-1}[m(z|v;\theta) - m(z|v;\theta_0)]^2 m_1(v;\theta_0)$$

is the integrand of the asymptotic criterion $L^{\mathrm{CD}}(\theta)$ in eq. (15) of Al-M, with $w(z,v) = m_1(v;\theta_0)^2/m_2(z,v;\theta_0)$, which is bounded and integrable by the assumption that $L^{\mathrm{CD}}(\theta)$ is continuous and bounded on Θ (Assumption 5). Moreover, for all $(z,v,\theta) \in \mathbb{R}^{q^*} \times \mathbb{R}^{q-q^*} \times \Theta$, we have that $|\mathbb{T}_{T,S}(v;\theta)[1 - \mathbb{T}_{2T}(z,v)]| \leq 1$, and $[1 - \mathbb{T}_{2T}(z,v)] \xrightarrow{p} 0$ pointwise. Hence $\iint a_{14T,S} \xrightarrow{p} 0$ for all $\theta \in \Theta$ and, hence, for all $\theta \in \Theta$,

$$\iint a_{1T,S}(z,v;\theta)\, dzdv \xrightarrow{p} 0. \tag{D2}$$

- For all $(z,v,\theta) \in \mathbb{R}^{q^*} \times \mathbb{R}^{q-q^*} \times \Theta$,

$$\begin{aligned} a_{2T,S}(z,v;\theta) &\leq \left[\phi_{T,S}(z,v;\theta) + \phi(z,v;\theta)\right] \\ &\quad \times [|\pi_{T,S}(z|v;\theta) - m(z|v;\theta)| + |\pi_T(z|v) - m(z|v;\theta_0)|]\, \mathbb{T}_{T,S}(v;\theta) \\ &\quad + \left[\phi_{T,S}(z,v;\theta) + \phi(z,v;\theta)\right][m(z|v;\theta) - m(z|v;\theta_0)][1 - \mathbb{T}_{T,S}(v;\theta)] \\ &\equiv a_{21T,S}(z,v;\theta) + a_{22T,S}(z,v;\theta). \end{aligned} \tag{D3}$$

For all $(z,v,\theta) \in \mathbb{R}^{q^*} \times \mathbb{R}^{q-q^*} \times \Theta$,

$$\begin{aligned} & a_{21T,S}(z,v;\theta) \\ & \leq |\pi_{T,S}(z|v;\theta) - m(z|v;\theta)|\, \mathbb{T}_{T,S}(v;\theta) \frac{1}{m_2(z,v;\theta_0)} \left| \frac{\pi_{1T}(v)}{\pi_T(z|v)} - \frac{m_1(v;\theta_0)}{m(z|v;\theta_0)} \right| \mathbb{T}_{2T}(z,v) \\ & \times m_2(z,v;\theta_0)[|\pi_{T,S}(z|v;\theta) - m(z|v;\theta)| + |\pi_T(z|v) - m(z|v;\theta_0)|]\, \mathbb{T}_{T,S}(v;\theta) \\ & + |\pi_T(z|v) - m(z|v;\theta_0)|\, \mathbb{T}_{T,S}(v;\theta) \frac{1}{m_2(z,v;\theta_0)} \left| \frac{\pi_{1T}(v)}{\pi_T(z|v)} - \frac{m_1(v;\theta_0)}{m(z|v;\theta_0)} \right| \mathbb{T}_{2T}(z,v) \\ & \times m_2(z,v;\theta_0)[|\pi_{T,S}(z|v;\theta) - m(z|v;\theta)| + |\pi_T(z|v) - m(z|v;\theta_0)|]\, \mathbb{T}_{T,S}(v;\theta) \\ & + |m(z|v;\theta) - m(z|v;\theta_0)|\, \mathbb{T}_{T,S}(v;\theta) \frac{1}{m_2(z,v;\theta_0)} \left| \frac{\pi_{1T}(v)}{\pi_T(z|v)} - \frac{m_1(v;\theta_0)}{m(z|v;\theta_0)} \right| \mathbb{T}_{2T}(z,v) \\ & \times m_2(z,v;\theta_0)[|\pi_{T,S}(z|v;\theta) - m(z|v;\theta)| + |\pi_T(z|v) - m(z|v;\theta_0)|]\, \mathbb{T}_{T,S}(v;\theta) \\ & + |m(z|v;\theta) - m(z|v;\theta_0)|\, m_1(v;\theta_0)\, m_2(z,v;\theta_0) \\ & \times \frac{1}{m(z|v;\theta_0)\, m_2(z,v;\theta_0)} [|\pi_{T,S}(z|v;\theta) - m(z|v;\theta)| + |\pi_T(z|v) - m(z|v;\theta_0)|]\, \mathbb{T}_{T,S}(v;\theta) \\ & + |\pi_{T,S}(z|v;\theta) - \pi_T(z|v)|\, \mathbb{T}_{T,S}(v;\theta) \frac{m_1(v;\theta_0)}{m(z|v;\theta_0)} \mathbb{T}_{2T}(z,v) \\ & \times [|\pi_{T,S}(z|v;\theta) - m(z|v;\theta)| + |\pi_T(z|v) - m(z|v;\theta_0)|]\, \mathbb{T}_{T,S}(v;\theta) \\ & \equiv a_{211T,S}(z,v;\theta) + a_{212T,S}(z,v;\theta) + a_{213T,S}(z,v;\theta) + a_{214T,S}(z,v;\theta) + a_{215T,S}(z,v;\theta). \end{aligned}$$

We have that for all $\theta \in \Theta$: (i) $\iint \sum_{j=1,2,3} a_{21jT,S} \xrightarrow{p} 0$ by Lemma C1-(a) in Appendix A.1, by Lemma C2-(a), and by boundedness and integrability of $m(z|v;\theta)$ and $m_2(z,v;\theta)$; (ii) $\iint a_{214T,S} \xrightarrow{p} 0$ by Lemma C3-(a) and boundedness and integrability of $m(z|v;\theta)$ and $m_2(z,v;\theta)$; (iii) $\iint a_{215T,S} \xrightarrow{p} 0$ by Lemmas C1-(a) and C3-(a). Therefore, for all $\theta \in \Theta$,

$$\iint a_{21T,S}(z,v;\theta)\, dzdv \xrightarrow{p} 0. \tag{D4}$$

Next, for all $(z,v,\theta) \in \mathbb{R}^{q^*} \times \mathbb{R}^{q-q^*} \times \Theta$,

$$a_{22T,S}(z,v;\theta) \leq \left\{ a^*_{22T,S}(z,v;\theta)\, m_2(z,v;\theta_0) + [m(z|v;\theta) - m(z|v;\theta_0)]^2 \frac{m_1(v;\theta_0)}{m(z|v;\theta_0)} \right\} [1 - \mathbb{T}_{T,S}(v;\theta)],$$

where

$$\begin{aligned}
& a^*_{22T,S}(z,v;\theta) \\
& \equiv [|\pi_{T,S}(z|v;\theta) - m(z|v;\theta)| + |\pi_T(z|v) - m(z|v;\theta_0)| + |m(z|v;\theta) - m(z|v;\theta_0)|]\, \mathbb{T}_{T,S}(v;\theta) \\
& \times \frac{1}{m_2(z,v;\theta_0)} \left| \frac{\pi_{1T}(v)}{\pi_T(z|v)} - \frac{m_1(v;\theta_0)}{m(z|v;\theta_0)} \right| \mathbb{T}_{2T}(z,v) \cdot |m(z|v;\theta) - m(z|v;\theta_0)| \\
& + \frac{1}{m_2(z,v;\theta_0)\, m(z|v;\theta_0)} |\pi_{T,S}(z|v;\theta) - \pi_T(z|v)|\, \mathbb{T}_{T,S}(v;\theta)\, m_1(v;\theta_0) \\
& \times |m(z|v;\theta) - m(z|v;\theta_0)|\, \mathbb{T}_{2T}(z,v).
\end{aligned}$$

As in Appendix C.1, we have that for all $\theta \in \Theta$, $1 - \mathbb{T}_{T,S}(v;\theta) \xrightarrow{p} 0$ pointwise. Since for all $(v,\theta) \in \mathbb{R}^{q-q^*} \times \Theta$, $1 - \mathbb{T}_{T,S}(v;\theta) \leq 1$, and the functions $m_2(z,v;\theta_0)$ and $m(z|v;\theta_0)^{-1}[m(z|v;\theta) - m(z|v;\theta_0)]^2 m_1(v;\theta_0)$ are bounded and integrable (by the assumption that the asymptotic criterion $L^{\mathrm{CD}}(\theta)$ is continuous and bounded on Θ), then, for all $\theta \in \Theta$,

$$\iint a_{22T,S}(z,v;\theta)\, dzdv \xrightarrow{p} 0. \tag{D5}$$

Hence, by (D3), (D4) and (D5),

$$\iint a_{2T,S}(z,v;\theta)\, dzdv \xrightarrow{p} 0. \tag{D6}$$

Hence, eq. (D1) holds by eqs. (D2) and (D6).

D.2 Asymptotic normality

Remarks on the use of Lemmas N4-N5.

(a) Lemma N4 is needed to show that the term $\mathcal{J}_{T,S}$ defined in eq. (D10) below converges in probability to the term $\mathcal{J}$ defined in eq. (D11) below. Lemma N5 is needed to show that the terms $\mathcal{I}_{1T,S}^{i}$ and $\mathcal{I}_{2T,S}^{i}$ defined in eqs. (D12b) and (D12c) below converge in distribution to the Gaussian terms provided in eqs. (D13) and (D14) below.

(b) The bandwidth conditions in Assumption 10-(b) of Al-M ensure that the suprema in Lemmas N4-N5 go to zero in probability, as explained in the Remarks 2 of Section A.3 of the present Appendix.

We now proceed to prove our asymptotic normality claims. We shall provide a result that is slightly more general than needed — namely, a result that holds for more general weighting functions than that considered in Theorem 2 of Al-M. Let

$$\mathbb{W}_T \equiv \left\{ w_T(z,v) : w_T(z,v) = \xi_{1T}(v) \frac{\pi_{1T}(v)^2}{\pi_{2T}(z,v)} \mathbb{T}_{2T}(z,v) \right\},$$

where the function $\xi_{1T}(v)$ satisfies the conditions in Lemma N4.

We study the asymptotic behavior of the CD-SNE for weighting functions $w_T \in \mathbb{W}_T$. First, by Remark 5 and Assumption 7-(a), we may still interchange the order of differentiation and integration in the criterion. The first order conditions of the CD-SNE are still as in eq. (C7) for $w_T \in \mathbb{W}_T$. So we only need to check that for every $w_T \in \mathbb{W}_T$, the terms $B_{T,S}$ and $C_{T,S}$ in eq. (C7) behave as in Claim 1 (eq. (C10)) in Claim 2 (eq. (C24)).

Claim 1

To check that Claim 1 holds for every $w_T \in \mathbb{W}_T$ and the additional Assumption 10 in Al-M, we need to verify that $B_{1T,S} + B_{2T,S} \xrightarrow{p} \mathbf{0}_n$ in the decomposition of $B_{T,S}$ given in (C11).

- As regards the $B_{1T,S}$ term in (C14), we only have to verify that for $w_T \in \mathbb{W}_T$, the terms $u_{T,S}^{i}(z,v)\left[w_T(z,v) - w(z,v)\right]$ $(i = 0, 1, \cdots, S)$ and the terms $v_{T,S}^{i}(z,v)\left[w_T(z,v) - w(z,v)\right]$ $(i = 1, \cdots, S)$ converge uniformly to zero, for $w(z,v) = \pi_1(v;\theta_0)^2 / \pi_2(z,v;\theta_0)$. For every $w_T \in \mathbb{W}_T$, we have,

$$\begin{aligned}
& w_T(z,v) - w(z,v) \\
&= \left[\xi_{1T}(v) \frac{\pi_{1T}(v)^2}{\pi_{2T}(z,v)} - \xi_1(v) \frac{\pi_1(v;\theta_0)^2}{\pi_2(z,v;\theta_0)} \right] \mathbb{T}_{2T}(z,v) - \xi_1(v) \frac{\pi_1(v;\theta_0)^2}{\pi_2(z,v;\theta_0)} \left[1 - \mathbb{T}_{2T}(z,v)\right] \\
&= \left[\xi_{1T}(v) - \xi_1(v)\right] \frac{\pi_{1T}(v)}{\pi_T(z|v)} \mathbb{T}_{2T}(z,v) + \xi_1(v) \left(\frac{\pi_{1T}(v)}{\pi_T(z|v)} - \frac{\pi_1(v;\theta_0)}{\pi(z|v;\theta_0)} \right) \mathbb{T}_{2T}(z,v) \\
&\quad - \xi_1(v) \frac{\pi_1(v;\theta_0)^2}{\pi_2(z,v;\theta_0)} \left[1 - \mathbb{T}_{2T}(z,v)\right] \\
&\equiv \kappa_{1T}(z,v) + \kappa_{2T}(z,v) + \kappa_{3T}(z,v). \qquad \text{(D7)}
\end{aligned}$$

By the assumption on the functions $\xi_{1T}(v)$ and $\xi_1(v)$, we have that $\sup_{(z,v)\in\mathbb{R}^q}|\kappa_{1T}(z,v)| = O_p\left(T^{-1/2}\lambda_T^{-(q-q^*)}\alpha_T^{-1}\right) + O_p\left(\lambda_T^r\alpha_T^{-1}\right) \xrightarrow{p} \mathbf{0}_n$, where the convergence result follows by the bandwidth conditions given in Assumption 10. By Lemma C2, $\sup_{(z,v)\in\mathbb{R}^q}|\kappa_{2T}(z,v)| \xrightarrow{p} \mathbf{0}_n$. Hence, we are left to show that the terms $u^i_{T,S}(z,v)\,\kappa_{3T}(z,v)$ and the terms $\upsilon^i_{T,S}(z,v)\,\kappa_{3T}(z,v)$ converge uniformly to zero. By the computations made to show that $\sup_{(z,v)\in\mathbb{R}^q}\left|u^i_{T,S}(z,v)\right| \xrightarrow{p} 0$ and $\sup_{(z,v)\in\mathbb{R}^q}\left|\upsilon^i_{T,S}(z,v)\right| \xrightarrow{p} 0$ we only need to show that $\sup_{(z,v)\in\mathbb{R}^q}|b_{1T,S}(z,v)\,\kappa_{3T}(z,v)| \xrightarrow{p} \mathbf{0}_n$ and $\sup_{(z,v)\in\mathbb{R}^q}\left|b^*_{1T,S}(z,v)\,\kappa_{3T}(z,v)\right| \xrightarrow{p} \mathbf{0}_n$, where $b_{1T,S}(z,v)$ is as in eq. (C18) and $b^*_{1T,S}(z,v)$ is as in eq. (C21). We have, by eq. (C19), and the definition of $\kappa_{3T}(z,v)$ in (D7),

$$\begin{aligned}
&\sup_{(z,v)\in\mathbb{R}^q}|b_{1T,S}(z,v)\,\kappa_{3T}(z,v)| \\
&= \alpha_T^{-1}\delta_T^{-1} \times \left[O_p\left(T^{-\frac{1}{2}}\lambda_T^{-q}\delta_T^{-1}\right) + O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)}\delta_T^{-2}\right) + O_p\left(\lambda_T^r\delta_T^{-2}\right)\right] \\
&\quad \times \left[O_p\left(T^{-\frac{1}{2}}\lambda_T^{-(q-q^*)-1}\right) + O_p\left(\lambda_T^r\right)\right] \\
&\xrightarrow{p} \mathbf{0}_n,
\end{aligned} \tag{D8}$$

where the convergence follows by the bandwidth conditions in Assumptions 10 (see the Remarks 2 in Section A.3 of the present Appendix). Moreover, it follows by eq. (C22) that $\sup_{(z,v)\in\mathbb{R}^q}\left|b^*_{1T,S}(z,v)\,\kappa_{3T}(z,v)\right| \xrightarrow{p} \mathbf{0}_n$, at the same rate as (D8). Hence, for every $w_T \in \mathbb{W}_T$, we have that $B_{1T,S} \xrightarrow{p} \mathbf{0}_n$ in the decomposition of $B_{T,S}$ given in (C11).

- Consider, next, the $B_{2T,S}$ term. By eq. (C23), we have that $B_{2T,S} \xrightarrow{p} \mathbf{0}_n$ if (i) $\sup_{(z,v)\in\mathbb{R}^q}\left|u^i_{T,S}(z,v)\,w(z,v)\right| \xrightarrow{p} \mathbf{0}_n$ $(i = 0, 1, \cdots, S)$, and (ii) $\sup_{(z,v)\in\mathbb{R}^q}\left|\upsilon^i_{T,S}(z,v)\,w(z,v)\right| \xrightarrow{p} \mathbf{0}_n$ $(i = 1, \cdots, S)$. But for every $w_T \in \mathbb{W}_T$, we have that $\kappa_{3T}(z,v) = [1 - \mathbb{T}_{2T}(z,v)]\,w(z,v)$. Therefore, (i) and (ii) follow by the previous arguments we used to show that for every $w_T \in \mathbb{W}_T$, $B_{1T,S} \xrightarrow{p} \mathbf{0}_n$. Hence, $B_{2T,S} \xrightarrow{p} \mathbf{0}_n$ and so $B_{1T,S} + B_{2T,S} \xrightarrow{p} \mathbf{0}_n$.

Claim 2

To check Claim 2 with $w_T \in \mathbb{W}_T$ and the additional Assumption 10 in Al-M, we need to verify that in the decomposition of $C_{T,S} = C_{1T,S} + C_{2T,S}$ given in eq. (C25),

$$C_{1T,S} = \mathcal{J}_{T,S} + o_p(1)\,; \qquad C_{2T,S} \xrightarrow{p} \mathbf{0}_{n\times n} \tag{D9}$$

and,

$$\mathcal{J}_{T,S} = \iint \left|\frac{\nabla_\theta \pi_{T,S}(z|\,v;\theta_0)\,\pi_{1T}(v)}{\pi_{2T}(z,v)}\mathbb{T}_{T,S}(v;\theta_0)\sqrt{\mathbb{T}_{2T}(z,v)}\right|_2 \xi_{1T}(v)\,\pi_{2T}(z,v)\,dzdv. \tag{D10}$$

- Eq. (D9) follows (i) by applying Lemmas C2 and C3 to the term $C_{11T,S}$ in (C26) evaluated at $w_T \in \mathbb{W}_T$; and (ii) because under the bandwidth conditions in Assumption 10-(b), the term $C_{12T,S}$ in (C26) is such that $C_{12T,S} \xrightarrow{p} \mathbf{0}_{n\times n}$ by the same arguments we used to show that $B_{1T} + B_{2T} \xrightarrow{p} \mathbf{0}_n$.

- The proof that $C_{2T,S} \xrightarrow{p} \mathbf{0}_{n\times n}$ for every $w_T \in \mathbb{W}_T$ follows by the decomposition in (C28), by eq. (D7), by Lemma C2, and by the same steps used in Section C.2 to show that $C_{2T,S} \xrightarrow{p} \mathbf{0}_{n\times n}$ for $w_T(z,v)$ satisfying Assumption 4 in Al-M.

Hence, given (D9), we have that $C_{T,S} = \mathcal{J}_{T,S} + o_p(1)$, where the $\mathcal{J}_{T,S}$ term in eq. (D10) is

$$\begin{aligned}\mathcal{J}_{T,S} &= \iint \left|\nabla_\theta \pi_{T,S}(z|\,v;\theta_0)\mathbb{T}_{T,S}(v;\theta_0)\right|_2 \frac{\pi_{1T}(v)}{\pi_T(z|\,v)}\xi_{1T}(v)\,\mathbb{T}_{2T}(z,v)\,dzdv \\ &= \iint \left|\frac{\nabla_\theta \pi_{T,S}(z|\,v;\theta_0)}{\pi_T(z|\,v)}\mathbb{T}_{T,S}(v;\theta_0)\sqrt{\mathbb{T}_{2T}(z,v)}\right|_2 \xi_{1T}(v)\,\pi_{2T}(z,v)\,dzdv,\end{aligned}$$

and by Lemma N4,

$$\mathcal{J}_{T,S} \xrightarrow{p} \mathcal{J} \equiv \iint \left|\frac{\nabla_\theta \pi(z|\,v;\theta_0)}{\pi(z|\,v;\theta_0)}\right|_2 \xi_1(v)\,\pi_2(z,v;\theta_0)\,dzdv. \tag{D11}$$

The asymptotic expansion

By the previous results that Claim 1 and Claim 2 in Section C.2 hold for every $w_T \in \mathbb{W}_T$, we may rewrite the first order conditions of the CD-SNE in eq. (C7) as follows,

$$o_p(1) = \frac{1}{S}\sum_{i=1}^{S}\left[\left(\mathcal{I}^i_{1T,S} - \mathcal{I}^0_{1T,S}\right) + \left(\mathcal{I}^i_{2T,S} - \mathcal{I}^0_{2T,S}\right)\right] + \left[\mathcal{J}_{T,S} + o_p(1)\right]\cdot\sqrt{T}\left(\theta_{T,S} - \theta_0\right), \tag{D12a}$$

where now, for every $w_T \in \mathbb{W}_T$, $\mathcal{J}_{T,S}$ behaves as in (D11), and the terms $\mathcal{I}^i_{1T,S}$ and $\mathcal{I}^i_{2T,S}$ are given by,

$$\mathcal{I}^i_{1T,S} = \iint \frac{\nabla_\theta \pi_{T,S}(z|\,v;\theta_0)\,\pi_{1T}(v)^2}{\pi^i_{1T}(v;\theta_0)\,\pi_{2T}(z,v)}\xi_{1T}(v)\mathbb{T}_{2T}(z,v)\,\mathbb{T}^2_{T,S}(v;\theta_0)\,dA^i_T(z,v) \tag{D12b}$$

$$\mathcal{I}^i_{2T,S} = \int \gamma^i_{T,S}(v)\,\mathbb{T}^2_{T,S}(v;\theta_0)\,dA^i_T(v) \tag{D12c}$$

where (D12c) follows by evaluating the function $\gamma^i_{T,S}(v)$ in (C32) at $w_T \in \mathbb{W}_T$,

$$\gamma^i_{T,S}(v) = \int \frac{\nabla_\theta \pi_{T,S}(z|\,v;\theta_0)\,E(\pi_{2T}(z,v))\,\xi_{1T}(v)\,\pi_{1T}(v)}{\pi^i_{1T}(v;\theta_0)\,\pi_{2T}(z,v)}\mathbb{T}_{2T}(z,v)\,dz.$$

By Lemma N5 applied to the $\mathcal{I}^i_{1T,S}$ terms, Assumption 10, and additional arguments nearly identical to those made in Section C.2, we have that for $i = 0, 1, \cdots, S$,

$$\mathcal{I}^i_{1T,S} \xrightarrow{d} \mathcal{I}^i_1 \equiv \iint \frac{\nabla_\theta \pi(z|\,v;\theta_0)}{\pi(z|\,v;\theta_0)}\xi_1(v)\,d\omega^0_i(F(z,v)). \tag{D13}$$

As regards the terms $\mathcal{I}^i_{2T,S}$ in (D12c), we claim that for $i = 0, 1, \cdots, S$,

$$\mathcal{I}^i_{2T,S} \xrightarrow{p} \mathbf{0}_n. \tag{D14}$$

By (C36), the proof of (D14) is nearly identical to the proof of (C34b), once we show that

$$\sup_{v\in\mathcal{B}_T}\left|\gamma_{T,S}^{i}(v)\right| \xrightarrow{p} \mathbf{0}_n, \tag{D15}$$

where $\mathcal{B}_T$ is the same trimming set introduced in the proof of Lemma C1 in Section A.

We have

$$\left|\gamma_{T,S}^{i}(v)\right| \leq \int \Lambda_{T,S}^{i}(z,v)\,dz, \tag{D16}$$

where, for all $(z,v)\in\mathcal{A}_T\cap\mathcal{B}_T$,

$$\begin{aligned}
\Lambda_{T,S}^{i}(z,v) &\equiv \left|\frac{\nabla_\theta \pi_{T,S}(z|\,v;\theta_0)\,\xi_{1T}(v)\,\pi_{1T}(v)\,E(\pi_{2T}(z,v))}{\pi_{1T}^{i}(v;\theta_0)\,\pi_{2T}(z,v)}\mathbb{T}_{2T}(z,v)\right| \\
&\leq E(\pi_{2T}(z,v))\cdot \sup_{(z,v)\in\mathcal{A}_T\cap\mathcal{B}_T}\left|\frac{\nabla_\theta \pi_{T,S}(z|\,v;\theta_0)\,\xi_{1T}(v)\,\pi_{1T}(v)}{\pi_{1T}^{i}(v;\theta_0)\,\pi_{2T}(z,v)} - \frac{\nabla_\theta \pi(z|\,v;\theta_0)\,\xi_1(v)}{\pi_2(z,v;\theta_0)}\right| \\
&+ \xi_1(v)\,\nabla_\theta\pi(z|\,v;\theta_0)\cdot \sup_{(z,v)\in\mathcal{A}_T\cap\mathcal{B}_T}\left|\frac{E(\pi_{2T}(z,v)) - \pi_2(z,v;\theta_0)}{\pi_2(z,v;\theta_0)}\right| \\
&+ \xi_1(v)\,\nabla_\theta\pi(z|\,v;\theta_0)\cdot \sup_{(z,v)\in\mathcal{A}_T\cap\mathcal{B}_T}\mathbb{T}_{2T}(z,v) \\
&\equiv \Lambda_{1T,S}^{i}(z,v) + \Lambda_{2T}(z,v) + \Lambda_{3T}(z,v),
\end{aligned}$$

and where $\mathcal{A}_T$ is the same trimming set introduced in Lemma C2.

Since $\int E(\pi_{2T}(z,v))\,dz$ is bounded and constant, we have, by Lemma N5, that

$$\sup_{v\in\mathbb{R}^{q^*}}\int \Lambda_{1T,S}^{i}(z,v)\,\mathbb{T}_{T,S}^{2}(v;\theta_0)\,dz \xrightarrow{p} \mathbf{0}_n.$$

Moreover,

$$\sup_{(z,v)\in\mathcal{A}_T\cap\mathcal{B}_T}\left|\frac{E(\pi_{2T}(z,v)) - \pi_2(z,v;\theta_0)}{\pi_2(z,v;\theta_0)}\right| = O_p\left(\lambda_T^{r}\alpha_T^{-1}\right) \xrightarrow{p} \mathbf{0}_n,$$

by the bandwidth conditions in Assumption 10-(b). This result, combined with $0 \equiv \xi_1(v)\int \nabla_\theta\pi(z|\,v;\theta_0)\,dz$, implies that:

$$\sup_{v\in\mathbb{R}^{q^*}}\int \Lambda_{2T}(z,v)\,\mathbb{T}_{T,S}^{2}(v;\theta_0)\,dz \xrightarrow{p} 0.$$

Finally, $\sup_{(z,v)\in\mathbb{R}^{q^*}\times\mathcal{B}_T}\mathbb{T}_{2T}(z,v) \leq 1$. Hence, $\int \Lambda_{3T,S}(z,v)\,dz \leq \xi_1(v)\int \nabla_\theta\pi(z|\,v;\theta_0)\,dz \equiv 0$. These results for $\Lambda_{1T,S}^{i}(z,v)$, $\Lambda_{2T}(z,v)$ and $\Lambda_{3T}(z,v)$, combined with the inequality in (D16), imply that the uniform convergence in (D15) holds true.

Hence, the asymptotic expansion in (D12a)-(D12b)-(D12c) is:

$$o_p(1) = \frac{1}{S}\sum_{i=1}^{S}\left(\mathcal{I}_{1T,S}^{i} - \mathcal{I}_{1T,S}^{0}\right) + \left[\mathcal{J}_{T,S} + o_p(1)\right]\cdot\sqrt{T}\left(\theta_{T,S} - \theta_0\right), \tag{D17}$$

where the terms $\mathcal{I}_{1T,S}^{i}$ satisfy (D13).

Asymptotic normality

The next result follows by (D11), (D13), (D14), (D17), the Slutzky's theorem, the assumption that $E[\|\nabla_\theta \log \pi(z_t|\, v_t; \theta_0)\|^\vartheta]^{1/\vartheta} < \infty$, for some $\vartheta > 2$ (and boundedness of $\xi_1(v)$), the mixing condition in Assumption 2, and the same computations leading to eq. (B7) in Section B.2 and eq. (C40) in Section C.2:

Proposition 2. *Under the Assumptions of Theorem 2 in Al-M, the CD-SNE with weighting functions $w_T \in \mathbb{W}_T$ is consistent and asymptotically normal with variance/covariance matrix*

$$\left(1 + \frac{1}{S}\right)\left(var(\Phi_t) + \sum_{k=1}^{\infty}\left[cov(\Phi_t, \Phi_{t+k}) + cov(\Phi_{t+k}, \Phi_t)\right]\right),$$

where $\Phi_t \equiv \Phi(z_t, v_t)$ and,

$$\Phi(z, v) \equiv \left[\iint \left|\frac{\nabla_\theta \pi(s'|\, s; \theta_0)}{\pi(s'|\, s; \theta_0)}\right|_2 \xi_1(s)\, \pi_2(s', s; \theta_0)\, ds' ds\right]^{-1} \frac{\nabla_\theta \pi(z|\, v; \theta_0)}{\pi(z|\, v; \theta_0)} \xi_1(v). \tag{D18}$$

Theorem 2 in Al-M is a special case of Proposition 2. Precisely, set $\xi_1(\cdot) = \xi_{1T}(\cdot) \equiv 1$ and $(z, v) = (y_2, y_1)$. The function Φ in (D18) is then,

$$\Phi(y_2, y_1) = \left[\iint \left|\frac{\nabla_\theta \pi(s'|\, s; \theta_0)}{\pi(s'|\, s; \theta_0)}\right|_2 \pi_2(s', s; \theta_0)\, ds' ds\right]^{-1} \frac{\nabla_\theta \pi(y_2|\, y_1; \theta_0)}{\pi(y_2|\, y_1; \theta_0)}.$$

The efficiency claim follows by the standard score martingale difference argument (see, e.g., Wooldridge (1994, Lemma 5.2 p. 2677)).

E. Diffusion models

E.1 Background

In this section, we study the large sample behavior of our estimators for stochastic differential equations.

Let $\Theta \subset \mathbb{R}^n$ be a compact parameter set, and for a given parameter vector θ_0 in the interior of Θ, consider the following data generating process $y = \{y(\tau)\}_{\tau \geq 0}$:

$$dy(\tau) = b\left(y(\tau), \theta_0\right) d\tau + a\left(y(\tau), \theta_0\right) dW(\tau),\ \tau \geq 0, \tag{E1}$$

where W is a standard d-dimensional Brownian motion; b and a are vector and matrix valued functions in $\mathbb{R}^d$ and $\mathbb{R}^{d \times d}$, respectively; a is full rank; and y takes values in $\mathbb{R}^d$. Similarly as in the discrete-time setup in Al-M, we partition $y(\tau)$ as $y(\tau) \equiv [y^o(\tau) \;\; y^u(\tau)]$, where $y^o(\tau) \in \mathbb{R}^{q^*}$ is the vector of the observable variables. We assume that the data are sampled at regular intervals. Thus we still let $q \equiv q^*(1+l)$ and $x_t \equiv \left[y_t^o, \cdots, y_{t-l}^o\right]$ $(t = 1+l, \cdots, T)$, where $\{y_t^o\}_{t=1}^T$ is the sequence of the observations and T is the sample size.

We consider the following regularity condition:

Maintained assumptions. *The system (E1) has a strong solution and it is strictly stationary. Furthermore, Assumptions 1 and 2 (with mixing coefficients $\bar{\beta}_k$ and exponent $\bar{\mu} > 1$, say) hold in the context of model (E1).*

Chen, Hansen and Carrasco (1999) provide primitive conditions guaranteeing that Assumption 2 in Al-M holds in the case of scalar diffusions. A scalar diffusion is β-mixing with exponential decay if their "pull measure", defined as $\frac{b}{a} - \frac{1}{2}\frac{\partial a}{\partial y}$, is negative (positive) at the right (left) boundary (the authors also provide conditions ensuring β-mixing with polynomial decay in the case of zero pull measure at one of the boundaries (see their Remark 5)). As regards multidimensional diffusions, β-mixing with exponential decay can be checked through results developed by Meyn and Tweedie (1993) for exponential ergodicity, as in Carrasco, Hansen and Chen (1999). Finally, Carrasco, Hansen and Chen (1999) provide more specific results pertaining to partially observed diffusions.

To generate simulated paths of the observable variables in (E1), various discretization schemes can be used (see, e.g., Kloeden and Platen (1999)). In this section, we consider the simple Euler-Maruyama discrete time approximation to (E1):

$${}_h y_{h(k+1)} - {}_h y_{hk} = b\left({}_h y_{hk}, \theta\right) h + a\left({}_h y_{hk}, \theta\right) \sqrt{h}\epsilon_{k+1}, \qquad k = 0, 1, \cdots, \tag{E2}$$

where h is the discretization step and $\{\epsilon_k\}_{k=1,\cdots}$ is a sequence of independent and identically distributed $\mathbb{R}^d$-valued random variables. Let $x_h^i(\theta) = \{x_{t,h}^i(\theta)\}_{t=1+l}^T$ denote the "pseudo"-skeleton of the i-th simulation path $(i = 1, \cdots, S)$ at the parameter value θ. (We use the wording "pseudo"-skeleton because h is nonzero.) That is, $x_{t,h}^i(\theta)$ is the i-th simulation of the t-th observation when the parameter vector is θ and the discretization step is h. We employ the following additional pieces of notation. We let, $\pi_{2T,h}^i(x;\theta) \equiv \left(T_T \lambda_T^q\right)^{-1} \sum_{t=1+l}^T K_q((x_{t,h}^i(\theta) - x)/\lambda_T)$, where $K_q(\cdot)$ is as in Assumption 3 of of Al-M. Accordingly, we set $\pi_{2T,S,h}(x;\theta) \equiv S^{-1} \sum_{i=1}^S \pi_{2T,h}^i(x;\theta)$;

and adopt a similar notation for the marginal density estimate of $\pi_1(v;\theta)$. Finally, we set $\pi_{T,S,h}(z,v;\theta) \equiv S^{-1}\sum_{i=1}^{S}\pi_{2T,h}^{i}(z,v;\theta)/\pi_{1T,h}^{i}(v;\theta)$. Let $y(\theta_0)$ denote the solution to (E1).

We assume that the high frequency simulator satisfies the following conditions:

Assumption E.1. *For all $\theta \in \Theta$,* **(a)** *the high frequency simulator (E2) converges weakly to the solution of (E1) i.e., for each i, $y_h^i(\theta) \Rightarrow y(\theta)$ as $h \downarrow 0$.* **(b)** *The diffusion and drift functions a and b are Lipschitz continuous in y; their components are four times continuously differentiable in y; and a, b and their partial derivatives up to the fourth order have polynomial growth in y.* **(c)** *Finally, h is a function of the sample size T, and is such that $h \equiv h_T \downarrow 0$ and:* **(c.1)** $\sqrt{T}h_T \to 0$*; or* **(c.2)** $\sqrt{T}h_T\delta_T^{-1} \to 0$*, where δ_T is the trimming parameter introduce in Assumption 8 of Al-M.*

The maintained Assumption that (E1) is stationary implies that the "observed skeleton" of the diffusion inherits the same features of the continuous-time process. Since the simulation step h can not be zero in practice, we extend Assumption 2 in Al-M to cover the "pseudo"-skeleton behavior:

Assumption E.2. *For all $\theta \in \Theta$, $\exists h^0 > 0$ depending on θ : for all $h \in \left(0, h^0\right)$, $y_h^i(\theta)$ is β-mixing with mixing coefficients $\beta_k(h) : \lim_{k\to\infty}\sup_{h\in(0,h^0)} k^{\mu_h}\beta_k(h) \to 0$ for some sequence $\{\mu_h\}_h > 1$; and $\lim_{h\downarrow 0}\mu_h = \bar{\mu}$, $\lim_{h\downarrow 0}\beta_k(h) = \bar{\beta}_k$, where $\bar{\mu}$ and $\bar{\beta}_k$ are as in the maintained assumptions.*

Primitive conditions ensuring that Assumption E.1-(a) holds are well-known and can be found, for instance, in Kloeden and Platen (1999). Primitive conditions guaranteeing that Assumption E.2 holds are also well-known (see, e.g., Tjøstheim (1990) for conditions ensuring that (E2) is exponentially ergodic for fixed h). Assumptions E.1-(b,c) make our estimators asymptotically free of biases arising from the *imperfect* simulation of model (E1) (model (E1) is imperfectly simulated so long as $h > 0$). Precisely, these biases arise through terms taking the form $\sqrt{T}[E(K_q(x_{t,h}^i(\theta_0)) - E(K_q(x_t))]$. By results summarized in Kloeden and Platen (1999, chapter 14), however, $\sqrt{T}[E(K_q(x_{t,h}^i(\theta_0)) - E(K_q(x_t))] = O(h \cdot \sqrt{T})$ whenever Assumptions E.1-(a,b) hold and K_q is as differentiable as a and b are in Assumption E.1-(b). The role of Assumption E.1-(c) is to asymptotically eliminate these bias terms. Naturally, more precise high frequency simulators would allow h to shrink to zero at an even lower rate. Finally, Assumption E.1-(b) can considerably be weakened. For example, one may simply require that a, b be Hölder continuous, as in Kloeden and Platen (1999, Theorem 14.1.5 p. 460). These extensions are not considered here to keep the presentation as simple as possible.

Let $L_{T,S,h}^{\text{CD}}$ and $L_{T,S,h}^{\text{J}}$ be the criterions of the CD-SNE (Definition 1 in Al-M) and the J-SNE (Definition 2 in Al-M), and consider a sequence h_T of discretization stepsizes converging to zero. We need the following regularity conditions:

Assumption E.3 **(a)** *Either* **(a.1)** *$L_{T,S,h}^{\text{J}}$ satisfies Assumption 11; or* **(a.2)** *$L_{T,S,h}^{\text{CD}}$ satisfies Assumption 5-6.* **(b)** *With $\left|\nabla_\theta K_q((x_{t,h}^i(\theta) - x)/\lambda_T)\right|$ replacing $\left|\nabla_\theta K_q((x_t^i(\theta) - x)/\lambda_T)\right|$, Assumption 7 holds.*

Assumptions E.1-E.3 are the additional assumptions we need to prove that our estimators work as in the discrete-time setup in Al-M. We have:

Theorem E.1. *Let Assumptions E.1-(a,b) and E.2 hold. Then, under the additional Assumptions E.1-(c.1) and E.3-(a.1,b), the J-SNE is as in Section B (i.e. as in Theorem 3 in Al-M). Under the additional Assumptions E.1-(c.2) and E.3-(a.2,b), the CD-SNE is as in Section C (i.e. as in Theorem 1 in Al-M) and in Section D (i.e. as in Theorem 2 in Al-M).*

The proof of this theorem is in the next two subsections.

E.2 Consistency

We only provide the proof of consistency for the J-SNE. The proofs of consistency for the CD-SNE follow by a mere change in notation.

Similarly as in Section B.1, we only need to show that for all $\theta \in \Theta$, $L^{\mathrm{J}}_{T,S,h}(\theta) \xrightarrow{p} L^{\mathrm{J}}(\theta)$ as $T \to \infty$ and $h \downarrow 0$. Now by arguments similar to those of the proof of Proposition 1,

$$\left|L^{\mathrm{J}}_{T,S,h}(\theta) - L^{\mathrm{J}}(\theta)\right| \leq \sum_{j=1}^{2} \int \sigma_{jT,S,h}(x;\theta)\, dx$$

where

$$\begin{aligned}
\sigma_{1T,S,h}(x;\theta) &\equiv |\pi_{2T,S,h}(x;\theta) - \pi_{2T}(x)| \cdot |m_2(x;\theta) - m_2(x;\theta_0)| \cdot |w_T(x) - w(x)| \\
\sigma_{2T,S,h}(x;\theta) &\equiv |[\pi_{2T,S,h}(x;\theta) - m_2(x;\theta)] - [\pi_{2T}(x) - m_2(x;\theta_0)]| \cdot \left[\bar{\phi}_{T,S,h}(x;\theta) + \bar{\phi}(x;\theta)\right] \\
\bar{\phi}_{T,S,h}(x;\theta) &\equiv |\pi_{2T,S,h}(x;\theta) - \pi_{2T}(x)| \cdot w_T(x) \quad ; \quad \bar{\phi}(x;\theta) \equiv |m_2(x;\theta) - m_2(x;\theta_0)| \cdot w(x)
\end{aligned}$$

By Assumption E1-(a), $x^i_h(\theta) \Rightarrow x(\theta)$ as $h \downarrow 0$ (all $\theta \in \Theta$), $i = 1, \cdots, S$, where $x(\theta) = \{x_t(\theta)\}_{t=1+l}^T$, and $x_t(\theta)$ is the hypothetical discretely sampled subvector x_t from (E1), for a generic parameter value θ. By continuity of $\pi^i_{2T,h}(x;\theta)$ with respect to the simulated points $\{x^i_{t,h}(\theta)\}_{t=1+l}^T$ and independence of the simulations, $\pi^i_{2T,h}(x;\theta) \Rightarrow \pi_{2T}(x;\theta) \equiv \sum_{t=1+l}^T K_q\left((x_t(\theta) - x)/\lambda_T\right) / \left(T\lambda_T^q\right)$ as $h \downarrow 0$ (all $\theta \in \Theta$), for all $i = 1, \cdots, S$. Therefore, for all $x \in \mathbb{R}^q$, $\sigma_{jT,S,h}(x;\theta) \Rightarrow \sigma_{jT}(x;\theta)$ as $h \downarrow 0$ (all $\theta \in \Theta$), $j = 1, 2$. Consistency follows by the results proven in Section B.1 that for all $\theta \in \Theta$, $\int \sigma_{jT,S}(x;\theta)\, dx \xrightarrow{p} 0$, $j = 1, 2$.

E.3 Asymptotic normality

We now produce proofs of asymptotic normality for the J-SNE and the CD-SNE. (The proof for the CD-SNE with weighting function $w_T(z,v) = [\pi_{1T}(v)/\pi_T(z|v)]\,\mathbb{T}_{2T}(z,v)$ follows by a mere change in notation, and the Markov property of a diffusion (see, e.g., Arnold (1992), Thm. 9.2.3 p. 146)).

SNE

By the same arguments in Section B.2, the first order conditions lead to the following counterpart to eq. (B5):

$$\sqrt{T}\int\left[\pi_{2T,S,h}(x;\theta_0)-\pi_{2T}(x)\right]\nabla_\theta\pi_{2T,S,h}(x;\theta_0)\,w_T(x)dx$$

$$=\int\sqrt{T}\left[\pi_{2T,S,h}(x;\theta_0)-E(\pi_{2T,S,h}(x;\theta_0))\right]\nabla_\theta\pi_{2T,S,h}(x;\theta_0)w_T(x)dx$$

$$-\int\sqrt{T}\left[\pi_{2T}(x)-E(\pi_{2T}^i(x))\right]\nabla_\theta\pi_{2T,S,h}(x;\theta_0)w_T(x)dx$$

$$+\int\sqrt{T}\left[E(\pi_{2T,S,h}(x;\theta_0))-E(\pi_{2T}(x))\right]\nabla_\theta\pi_{2T,S,h}(x;\theta_0)w_T(x)dx$$

$$\equiv G_{1Th,S}+G_{2Th,S}+G_{3Th,S}.$$

(The presence of the additional term $G_{3Th,S}$ arises due to the imperfectness of simulations.) Under Assumption E1-(b) and Assumption 3 (the kernel is four times continuously differentiable), $G_{3Th,S}=O_p(\sqrt{T}h)$ by Kloeden and Platen (1999, Thm. 14.5.1 p. 473). By Assumption E1-(c), $G_{3Th,S}\xrightarrow{p}\mathbf{0}_n$. The terms $G_{1Th,S}$ and $G_{2Th,S}$ behave exactly as the two terms in the r.h.s. of eq. (B5) in Section B.2.

CD-SNE

The first order conditions are still formally as in Section C.2, and lead to the following expansion,

$$\mathbf{0}_n=\frac{1}{S}\sum_{i=1}^{S}\sqrt{T}\iint\left[\frac{\pi_{2T,h}^i(z,v;\theta_0)}{\pi_{1T}^i(v;\theta_0)}-\frac{\pi_{2T}(z,v)}{\pi_{1T}(v)}\right]\nabla_\theta\pi_{T,S,h}(z|\,v;\theta_0)\,w_T(z,v)\mathbb{T}_{T,S}^2(v;\theta_0)\,dzdv$$
$$+B_{T,S,h}+C_{T,S,h}\cdot\sqrt{T}\left(\theta_{T,S,h}-\theta_0\right),$$

for some convex combination $\bar{\theta}$ of θ_0 and $\theta_{T,S,h}$. Here, $\theta_{T,S,h}$ is the CD-SNE, and,

$$B_{T,S,h}\equiv\sqrt{T}\iint\left[\pi_{T,S,h}(z|\,v;\theta_0)-\pi_T(z|\,v)\right]^2 w_T(z,v)\,\mathbb{T}_{T,S}(v;\theta_0)\,\nabla_\theta\mathbb{T}_{T,S}(v;\theta_0)\,dzdv \quad \text{(E3)}$$

$$C_{T,S,h}\equiv\iint\nabla_\theta\left\{\left[\pi_{T,S,h}\left(z|\,v;\bar{\theta}\right)-\pi_T(z|\,v)\right]\nabla_\theta\pi_{T,S,h}\left(z|\,v;\bar{\theta}\right)\mathbb{T}_{T,S}^2\left(v;\bar{\theta}\right)\right\}w_T(z,v)\,dzdv \quad \text{(E4)}$$
$$+\iint\nabla_\theta\left\{\left[\pi_{T,S,h}\left(z|\,v;\bar{\theta}\right)-\pi_T(z|\,v)\right]^2\mathbb{T}_{T,S}\left(v;\bar{\theta}\right)\nabla_\theta\mathbb{T}_{T,S}\left(v;\bar{\theta}\right)\right\}w_T(z,v)\,dzdv$$

As regards the $B_{T,S,h}$ term in (E3) we have,

$$B_{T,S,h} = \sqrt{T}\iint \frac{1}{S}\sum_{i=1}^{S}\left[\frac{\pi_{2T,h}^{i}(z,v;\theta_0)}{\pi_{1T}^{i}(v;\theta_0)} - \frac{\pi_{2T}(z,v)}{\pi_{1T}(v)}\right]\left[\pi_{T,S,h}(z|v;\theta_0) - \pi_T(z|v)\right]$$
$$\times\left[w_T(z,v) - w(z,v)\right]\mathbb{T}_{T,S}(v;\theta_0)\nabla_\theta \mathbb{T}_{T,S}(v;\theta_0)\,dzdv$$
$$+\sqrt{T}\iint \frac{1}{S}\sum_{i=1}^{S}\left[\frac{\pi_{2T,h}^{i}(z,v;\theta_0)}{\pi_{1T}^{i}(v;\theta_0)} - \frac{\pi_{2T}(z,v)}{\pi_{1T}(v)}\right]\left[\pi_{T,S,h}(z|v;\theta_0) - \pi_T(z|v)\right]$$
$$\times w(z,v)\,\mathbb{T}_{T,S}(v;\theta_0)\nabla_\theta \mathbb{T}_{T,S}(v;\theta_0)\,dzdv. \tag{E5}$$

Moreover, at all points of continuity,

$$\frac{\pi_{2T,h}^{i}(z,v;\theta_0)}{\pi_{1T}^{i}(v;\theta_0)} - \frac{\pi_{2T}(z,v)}{\pi_{1T}(v)}$$
$$= \frac{\pi_{2T,h}^{i}(z,v;\theta_0) - E(\pi_{2T,h}^{i}(z,v;\theta_0))}{\pi_{1T}^{i}(v;\theta_0)} - \frac{\pi_{2T}(z,v) - E(\pi_{2T}(z,v))}{\pi_{1T}(v)} + \frac{E(\pi_{2T}(z,v))\left[\pi_{1T}(v) - \pi_{1T}^{i}(v;\theta_0)\right]}{\pi_{1T}^{i}(v;\theta_0)\cdot\pi_{1T}(v)}$$
$$+\frac{E(\pi_{2T,h}^{i}(z,v;\theta_0)) - E(\pi_{2T}(z,v))}{\pi_{1T}^{i}(v;\theta_0)}$$
$$\equiv \breve{A}_{1T} + \breve{A}_{2T} + \breve{A}_{3T} + \breve{A}_{4T}. \tag{E6}$$

(This expression is the counterpart to eq. (C13), and differs formally from (C13) because of the additional term $\breve{A}_{4T}$, which arise due to the imperfectness of simulations.) By replacing this expression into (E5),

$$B_{T,S,h} = B_{T,S,h}^{*} + B_{T,S,h}^{**},$$

where

$$B_{T,S,h}^{*} = \sqrt{T}\iint \frac{1}{S}\sum_{i=1}^{S}\left[\frac{E(\pi_{2T,h}^{i}(z,v;\theta_0)) - E(\pi_{2T}(z,v))}{\pi_{1T}^{i}(v;\theta_0)}\right]\left[\pi_{T,S,h}(z|v;\theta_0) - \pi_T(z|v)\right]$$
$$\times\left[w_T(z,v) - w(z,v)\right]\mathbb{T}_{T,S}(v;\theta_0)\nabla_\theta \mathbb{T}_{T,S}(v;\theta_0)\,dzdv$$
$$+\sqrt{T}\iint \frac{1}{S}\sum_{i=1}^{S}\left[\frac{E(\pi_{2T,h}^{i}(z,v;\theta_0)) - E(\pi_{2T}(z,v))}{\pi_{1T}^{i}(v;\theta_0)}\right]\left[\pi_{T,S,h}(z|v;\theta_0) - \pi_T(z|v)\right]$$
$$\times w(z,v)\,\mathbb{T}_{T,S}(v;\theta_0)\nabla_\theta \mathbb{T}_{T,S}(v;\theta_0)\,dzdv$$
$$\equiv B_{1T,S,h}^{*} + B_{2T,S,h}^{*},$$

and $B_{T,S,h}^{**}$ is a term such that $B_{T,S,h}^{**} \xrightarrow{p} \mathbf{0}_n$ (by the same arguments we used to show that $B_{jT,S} \xrightarrow{p} \mathbf{0}_n$ $(j = 1, 2)$ in Section C.2.) We claim that $B_{T,S,h}^{*} \xrightarrow{p} \mathbf{0}_n$ as well. Indeed,

$$\left|B_{1T,S,h}^{*}\right| \leq \sqrt{T}\iint \frac{1}{S}\sum_{i=1}^{S}\frac{\left|E(\pi_{2T,h}^{i}(z,v;\theta_0)) - E(\pi_{2T}(z,v))\right|\left|\pi_{T,S,h}(z|v;\theta_0) - \pi_T(z|v)\right|}{\pi_{1T}^{i}(v;\theta_0)\,\pi_2(z,v;\theta_0)}$$
$$\times\left[w_T(z,v) - w(z,v)\right]\mathbb{T}_{T,S}(v;\theta_0)\nabla_\theta \mathbb{T}_{T,S}(v;\theta_0)\,\pi_2(z,v;\theta_0)\,dzdv$$
$$= O_p\left(\sqrt{T}\cdot h\right)\times\delta_T^{-1}\alpha_T^{-1}\times o_p(1),$$

where the $o_p(1)$ term arises through an application of Lemma N1, and the bandwidth conditions in Assumption 9-(b). The $\left|B^*_{1T,S,h}\right|$ term behaves in the same manner. Hence, by the bandwidth conditions in Assumption 9-(b), $B^*_{jT,S,h} \xrightarrow{p} \mathbf{0}_n$ $(j=1,2)$ and, then, $B_{T,S,h} \xrightarrow{p} \mathbf{0}_n$ in (E3). Finally, the $C_{T,S,h}$ term in (E4) behaves precisely as the $C_{T,S}$ term in Section C.2 (see eq. (C24)). Therefore, the first order conditions of the CD-SNE are as in eq. (C29):

$$\mathbf{0}_n = \frac{1}{S}\sum_{i=1}^{S}\sqrt{T}\iint\left[\frac{\pi^i_{2T,h}(z,v;\theta_0)}{\pi^i_{1T}(v;\theta_0)} - \frac{\pi_{2T}(z,v)}{\pi_{1T}(v)}\right]\nabla_\theta \pi_{T,S,h}(z|v;\theta_0) w_T(z,v)\mathbb{T}^2_{T,S}(v;\theta_0)\, dzdv + o_p(1)$$
$$+\left[\iint |\nabla_\theta \pi_{T,S,h}(z|v;\theta_0)\mathbb{T}_{T,S}(v;\theta_0)|_2\, w_T(z,v)dzdv + o_p(1)\right]\cdot\sqrt{T}\left(\theta_{T,S,h}-\theta_0\right). \qquad \text{(E7)}$$

To complete the proof, we only need to deal with the first term in (E7). (The third term in (E7) behaves as the third term in (C29).) By plugging (E6) into the first term of (E7), we obtain,

$$\frac{1}{S}\sum_{i=1}^{S}\sqrt{T}\iint\left[\frac{\pi^i_{2T,h}(z,v;\theta_0)}{\pi^i_{1T}(v;\theta_0)} - \frac{\pi_{2T}(z,v)}{\pi_{1T}(v)}\right]\nabla_\theta \pi_{T,S,h}(z|v;\theta_0) w_T(z,v)\mathbb{T}^2_{T,S}(v;\theta_0)\, dzdv$$
$$= \frac{1}{S}\sum_{i=1}^{S}\left(\mathcal{I}^i_{1T,S,h} - \mathcal{I}^0_{1T,S,h}\right) + \left(\mathcal{I}^i_{2T,S,h} - \mathcal{I}^0_{2T,S,h}\right) + D^*_{T,S,h}$$

where the terms $\mathcal{I}^i_{1T,S,h}$ and $\mathcal{I}^i_{2T,S,h}$ $(i=0,1,\cdots,S)$ behave precisely as the corresponding terms $\mathcal{I}^i_{1T,S}$ and $\mathcal{I}^i_{2T,S}$ in Section C.2, and,

$$D^*_{T,S,h} \equiv \frac{1}{S}\sum_{i=1}^{S}\sqrt{T}\iint\left[\frac{E(\pi^i_{2T,h}(z,v;\theta_0)) - E(\pi_{2T}(z,v))}{\pi^i_{1T}(v;\theta_0)}\right]\nabla_\theta \pi_{T,S,h}(z|v;\theta_0) w_T(z,v)\mathbb{T}^2_{T,S}(v;\theta_0)\, dzdv.$$

By Assumption E.1-(c.2), and Lemma N1-(ii), we then have that $D^*_{T,S,h} \xrightarrow{p} \mathbf{0}_n$.

F. Identifiability and bandwidth choice, modulus of continuity issues and Neyman Chi Square measures of distance

F.1 Identifiability and bandwidth choice

For completeness, we reproduce the two examples in Appendix A.2 of Al-M. These examples relate to instances of kernels and data generating processes with both bounded and unbounded support for which the identifiability condition in Al-M (Assumption 5) holds.

Example 1. Let y_t in Al-M (eq. (8)) be independent and identically distributed as a Gaussian with unit variance and mean parameter $\theta_0 = 0$. Let the kernel be uniform, as in the example of Al-M (Section 3.1, eq. (16)), i.e. $K(y) = \frac{1}{2}\mathbb{I}_{|y|\leq 1}$, where $\mathbb{I}$ is the indicator function. In this case, the asymptotic criterion of the CD-SNE is given by:

$$\mathbb{L}\left(\theta,\bar{\lambda}\right) \equiv \int_{\mathbb{R}} \left[\pi_1^*\left(y;\theta,\bar{\lambda}\right) - \pi_1^*\left(y;\theta_0,\bar{\lambda}\right)\right]^2 w\left(y\right) dy, \tag{F1}$$

where,

$$\pi_1^*\left(y;\theta,\bar{\lambda}\right) - \pi_1^*\left(y;\theta_0,\bar{\lambda}\right) = \frac{1}{2\bar{\lambda}}\frac{1}{\sqrt{2\pi}}\int_{y-\bar{\lambda}}^{y+\bar{\lambda}}\left(e^{-\frac{1}{2}(\xi-\theta)^2} - e^{-\frac{1}{2}(\xi-\theta_0)^2}\right)d\xi, \quad \theta_0 = 0,$$

and where we take $w(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}$. Identification occurs if $\mathbb{L}\left(\theta,\bar{\lambda}\right) = 0 \Rightarrow \theta = \theta_0$, or if

$$\bar{\lambda} : \sup_{y\in\mathbb{R}}\left|\pi_1^*\left(y;\theta,\bar{\lambda}\right) - \pi_1^*\left(y;\theta_0,\bar{\lambda}\right)\right| = 0. \tag{F2}$$

Let the limiting bandwidth value $\bar{\lambda} = \frac{1}{2}$. Figure 1 below illustrates that $\mathbb{L}\left(\theta,\frac{1}{2}\right) = 0$ only with $\theta = \theta_0$. In other terms, $\theta_0 = 0$ is the only parameter value for θ that makes $\pi_1^*\left(y;\theta,\frac{1}{2}\right) - \pi_1^*\left(y;\theta_0,\frac{1}{2}\right) = 0$ for each y, which is what is formally required by (F2).

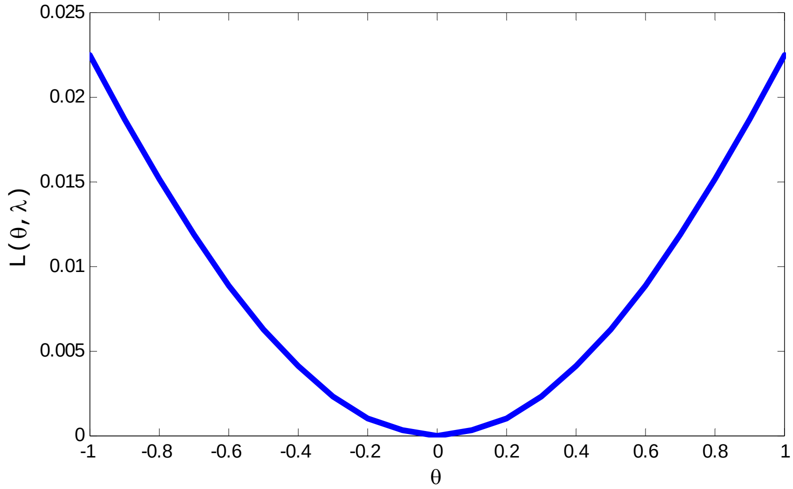

Figure 1 - Identifiability with Normal distributions and uniform kernels. This picture depicts the SNE asymptotic criterion $\mathbb{L}\left(\theta,\bar{\lambda}\right)$ in eq. (F1), evaluated at $\bar{\lambda}=\frac{1}{2}$.

Next, we develop one example in which data have *bounded* support, and identification occurs even when the limiting bandwidth value $\bar{\lambda}$ is nonzero.

Example 2. Let us assume that y_t is independent and identically distributed, generated by a Beta distribution with parameters θ,β, where β_0 is known and equal to 2. Therefore, the support is $Y=(0,1)$, and the marginal density for y_t is,

$$\pi_1\left(y;\theta\right)=\frac{\Gamma\left(\theta+2\right)}{\Gamma\left(\theta\right)}y^{\theta-1}\left(1-y\right).$$

Let $\theta_0=2$. For all $\theta\in(1,\infty)$,

$$\Delta\pi_1^*\left(y;\theta,\bar{\lambda}\right)\equiv\pi_1^*\left(y;\theta,\bar{\lambda}\right)-\pi_1^*\left(y;\theta_0,\bar{\lambda}\right)=\frac{1}{\bar{\lambda}}\int_0^1 K\left(\frac{y-\xi}{\bar{\lambda}}\right)\left[\frac{\Gamma\left(\theta+2\right)}{\Gamma\left(\theta\right)}\xi^{\theta-1}-6\xi\right]\left(1-\xi\right)d\xi.$$

Consider the kernel $K\left(x\right)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$, and let $\bar{\lambda}=1$. Figure 2 plots both $\pi_1^*\left(y;\theta,\bar{\lambda}\right)$ and $\Delta\pi_1^*\left(y;\theta,\bar{\lambda}\right)$. Note, now, that for all y, $\Delta\pi_1^*\left(y;\theta,1\right)=0\Rightarrow\theta=\theta_0=2$. Thus, in this example, the model and the limiting bandwidth $\bar{\lambda}$ satisfy an indentification condition which is even stronger than required by(F2), i.e. $\sup_{y\in(0,1)}\left|\Delta\pi_1^*\left(y;\theta,1\right)\right|=0\Rightarrow\theta=\theta_0$.

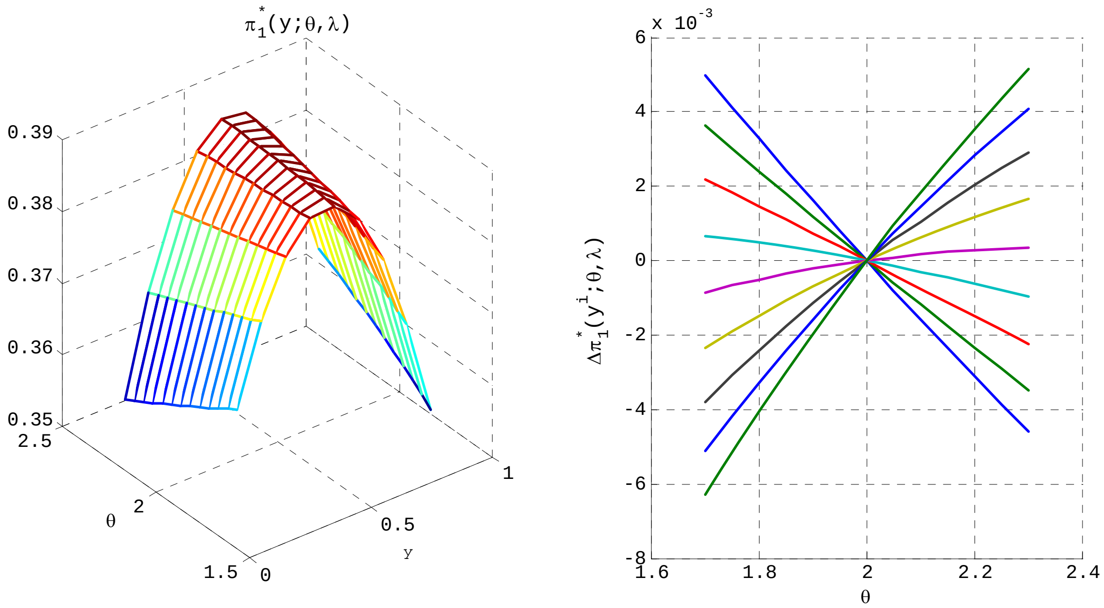

Figure 2 - Identifiability with Beta distributions. The left hand side panel depicts the function $\pi_1^*\left(y;\theta,\bar{\lambda}\right)=\bar{\lambda}^{-1}\int K\left(\frac{y-\xi}{\bar{\lambda}}\right)\pi_1\left(\xi;\theta\right)d\xi$ evaluated at the points $y^i=0.10,\ 0.20,\ \cdots,\ 0.90$ and $\theta^i=1.70,\ 1.75,\ \cdots,\ 2.30$. The right hand side panel depicts the lines $\theta\mapsto\Delta\pi_1^*\left(y;\theta,\bar{\lambda}\right)=\pi_1^*\left(y;\theta,\bar{\lambda}\right)-\Delta\pi_1^*\left(y;\theta_0,\bar{\lambda}\right)$ evaluated at the same points y^i,θ^i. In both cases, $\pi_1\left(y;\theta\right)=\frac{\Gamma(\theta+2)}{\Gamma(\theta)}y^{\theta-1}\left(1-y\right)$, $\theta_0=2$, K is the Gaussian kernel and $\bar{\lambda}=1$.

F.2 Modulus of continuity issues

On Assumption 12. An example of conditions under which the modulus of continuity condition in Assumption 12 holds is a global modulus of continuity condition on $\pi_{2T,S}(x;\cdot)$ similar to the standard one used by Duffie and Singleton (1993, p. 938) in a related problem:

$$\forall x\in\mathbb{R}^q,\quad \forall\left(\varphi,\theta\right)\in\Theta\times\Theta,\quad \left|\pi_{2T,S}\left(x;\varphi\right)-\pi_{2T,S}\left(x;\theta\right)\right|\ \leq k_{T,S}\left(x\right)\cdot\left\|\varphi-\theta\right\|^c,\quad c>0, \qquad \text{(F3)}$$

where $k_{T,S}\left(x\right)$ is a sequence of functions such that

$$\beta_{pT,S}\equiv\int k_{T,S}\left(x\right)^p w_T\left(x\right)dx<\infty,\ \text{all } T \text{ and } p=1,2.$$

In turn, condition (F3) holds for $c=1$ whenever $\nabla_\theta\pi_{2T,S}(x;\theta)$ is continuous and bounded (see, also, related results by Andrews (1992, p. 248-249)). Indeed, suppose that $\pi_{2T,S}$ has bounded derivative w.r.t θ. By the mean value theorem,

$$\forall x\in\mathbb{R}^q,\quad \forall\left(\varphi,\theta\right)\in\Theta\times\Theta,\quad \pi_{2T,S}\left(x;\varphi\right)=\pi_{2T,S}\left(x;\theta\right)+\nabla_\theta\pi_{2T,S}\left(x;\bar{\theta}\right)\cdot\left(\varphi-\theta\right),$$

for some convex combination $\bar{\theta}$ of φ and θ. Hence,

$$\forall x \in \mathbb{R}^q, \quad \forall (\varphi,\theta) \in \Theta \times \Theta, \quad \left|\pi_{2T,S}(x;\varphi) - \pi_{2T,S}(x;\theta)\right| \leq M \sum_{i=1}^{n} |\varphi - \theta|_i \leq M\sqrt{n} \left\|\varphi - \theta\right\|,$$

where $M \equiv \sup_{\theta\in\Theta} \left\{ \left|\nabla_\theta \pi_{2T,S}(x;\theta)\right|_i, \, i = 1, \cdots, n, \, x \in \mathbb{R}^q \right\}$, and the second inequality follows by the Cauchy-Schwartz inequality, and $M < \infty$ because Θ is compact. We now prove the claim that (F3) implies the modulus of continuity condition in Assumption 12.

Modulus of continuity claim. (Ineq. (F3) implies the modulus of continuity condition in Assumption 12) For all $(\varphi,\theta) \in \Theta \times \Theta$,

$$\begin{aligned} L_{T,S}^{\mathrm{J}}(\varphi) - L_{T,S}^{\mathrm{J}}(\theta) &= \int \left[\pi_{2T,S}(x;\varphi) - \pi_{2T,S}(x;\theta)\right]^2 w_T(x)dx \\ &\quad +2\int \left[\pi_{2T,S}(x;\varphi) - \pi_{2T,S}(x;\theta)\right]\left[\pi_{2T,S}(x;\theta) - \pi_{2T}(x)\right] w_T(x)dx. \end{aligned}$$

Let $B \equiv \max_{(\varphi,\theta)\in\Theta\times\Theta} \left\|\varphi - \theta\right\|^c$. By Θ compact, $B < \infty$. By condition (F3),

$$\begin{aligned} \left|L_{T,S}^{\mathrm{J}}(\varphi) - L_{T,S}^{\mathrm{J}}(\theta)\right| &\leq \left\|\varphi - \theta\right\|^{2c} \cdot \beta_{2T,S} + 2\left\|\varphi - \theta\right\|^c \cdot \int k_{T,S}(x) \left|\pi_{2T,S}(x;\theta) - \pi_{2T}(x)\right| w_T(x)\, dx \\ &\leq \left\|\varphi - \theta\right\|^{2c} \cdot \beta_{2T,S} + \left\|\varphi - \theta\right\|^c \cdot \xi_{T,S} \cdot \beta_{1T,S} \\ &\leq \left\|\varphi - \theta\right\|^c \cdot \left(B \cdot \beta_{2T,S} + \xi_{T,S} \cdot \beta_{1T,S}\right), \end{aligned}$$

where $\xi_{T,S} \equiv 2 \cdot \sup_{x\in\mathbb{R}^q,\, \theta\in\Theta} \left|\pi_{2T,S}(x;\theta) - \pi_{2T}^i(x)\right| < \infty$. Since $\beta_{1T,S}$, $\beta_{2T,S}$ and $\xi_{T,S}$ are bounded in probability as T becomes large, so is $B \cdot \beta_{2T,S} + \xi_{T,S} \cdot \beta_{1T,S}$. Set then $\kappa_{T,S} \equiv B \cdot \beta_{2T,S} + \xi_T \cdot \beta_{1T,S}$ to conclude. ■

Modulus of continuity issues. We present one primitive condition ensuring that the modulus of continuity condition in Assumption 6 holds true in the context of the asymptotics for the CD-SNE in Sections C and D, namely that for all $(z,v)\in\mathbb{R}^{q^*}\times\mathbb{R}^{q-q^*}$ and for all $(\varphi,\theta)\in\Theta\times\Theta$,

$$\begin{array}{ll}\text{(i)} & \left|\pi_{T,S}\left(z|\,v;\varphi\right)\mathbb{T}_{T,S}\left(v;\varphi\right)-\pi_{T,S}\left(z|\,v;\theta\right)\mathbb{T}_{T,S}\left(v;\theta\right)\right|\leq k_{1T,S}\left(z,v\right)\cdot\left\|\varphi-\theta\right\|^{c}\\ \text{(ii)} & \left|\mathbb{T}_{T,S}\left(v;\varphi\right)-\mathbb{T}_{T,S}\left(v;\theta\right)\right|\leq k_{2T,S}\left(z,v\right)\cdot\left\|\varphi-\theta\right\|^{c}\end{array}$$

where the functions $k_{iT,S}\left(z,v\right)$ satisfy,

$$\begin{aligned}M_{1T,S} &\equiv \iint k_{1T,S}\left(z,v\right)\left|\pi_{T,S}\left(z|\,v;\varphi\right)\mathbb{T}_{T,S}\left(v;\varphi\right)-\pi_{T,S}\left(z|\,v;\theta\right)\mathbb{T}_{T,S}\left(v;\theta\right)\right|w_{T}\left(z,v\right)dzdv\\ M_{2T,S} &\equiv \iint k_{2T,S}\left(z,v\right)\pi_{T}\left(z|\,v\right)^{2}\left[\mathbb{T}_{T,S}\left(v;\varphi\right)+\mathbb{T}_{T,S}\left(v;\theta\right)\right]w_{T}\left(z,v\right)dzdv\\ M_{3T,S} &\equiv \iint k_{2T,S}\left(z,v\right)\pi_{T,S}\left(z|\,v;\theta\right)\mathbb{T}_{T,S}\left(v;\theta\right)\pi_{T}\left(z|\,v\right)w_{T}\left(z,v\right)dzdv\\ M_{4T,S} &\equiv 2\iint k_{1T,S}\left(z,v\right)\pi_{T,S}(z|\,v;\theta)\mathbb{T}_{T,S}\left(v;\theta\right)w_{T}\left(z,v\right)dzdv\\ M_{5T,S} &\equiv 2\iint k_{1T,S}\left(z,v\right)\pi_{T}\left(z|\,v\right)\mathbb{T}_{T,S}\left(v;\theta\right)w_{T}\left(z,v\right)dzdv\end{aligned}$$

where $\sum_{j=1}^{5}M_{jT,S}\equiv M_{T,S}<\infty$. Indeed, we have that for all $(\varphi,\theta)\in\Theta\times\Theta$,

$$\begin{aligned}&L_{T,S}^{\mathrm{CD}}\left(\varphi\right)-L_{T,S}^{\mathrm{CD}}\left(\theta\right)\\ &=\iint\left[\pi_{T,S}\left(z|\,v;\varphi\right)\mathbb{T}_{T,S}\left(v;\varphi\right)-\pi_{T,S}\left(z|\,v;\theta\right)\mathbb{T}_{T,S}\left(v;\theta\right)\right]^{2}w_{T}\left(z,v\right)dzdv\\ &+\iint\pi_{T}\left(z|\,v\right)^{2}\left[\mathbb{T}_{T,S}\left(v;\varphi\right)+\mathbb{T}_{T,S}\left(v;\theta\right)\right]\left[\mathbb{T}_{T,S}\left(v;\varphi\right)-\mathbb{T}_{T,S}\left(v;\theta\right)\right]w_{T}\left(z,v\right)dzdv\\ &+\iint\pi_{T,S}\left(z|\,v;\theta\right)\mathbb{T}_{T,S}\left(v;\theta\right)\left[\mathbb{T}_{T,S}\left(v;\varphi\right)-\mathbb{T}_{T,S}\left(v;\theta\right)\right]\pi_{T}\left(z|\,v\right)w_{T}\left(z,v\right)dzdv\\ &+2\iint\left[\pi_{T,S}\left(z|\,v;\varphi\right)\mathbb{T}_{T,S}\left(v;\varphi\right)-\pi_{T,S}\left(z|\,v;\theta\right)\mathbb{T}_{T,S}\left(v;\theta\right)\right]\pi_{T,S}(z|\,v;\theta)\mathbb{T}_{T,S}\left(v;\theta\right)w_{T}\left(z,v\right)dzdv\\ &-2\iint\left[\pi_{T,S}\left(z|\,v;\varphi\right)\mathbb{T}_{T,S}\left(v;\varphi\right)-\pi_{T,S}\left(z|\,v;\theta\right)\mathbb{T}_{T,S}\left(v;\theta\right)\right]\mathbb{T}_{T,S}\left(v;\theta\right)\pi_{T}\left(z|\,v\right)w_{T}\left(z,v\right)dzdv\\ &\leq M_{T,S}\cdot\left\|\varphi-\theta\right\|^{c}.\end{aligned}$$

F.3 On Neyman Chi Square measures of distance

We provide a simple example of parameter restrictions ensuring the existence of Neyman Chi-Square distances in a dynamic context. We consider a stationary Gaussian AR(1) model,

$$y_{t}=b_{0}y_{t-1}+\epsilon_{t},\quad \epsilon_{t}\sim\mathrm{NID}\left(0,\sigma_{0}^{2}\right),\quad b_{0}\in\left(-1,1\right),\ \sigma_{0}>0,\quad t=-\infty,\cdots \tag{F4}$$

Our Neyman Chi-Square measure of distance is,

$$\mathcal{N}\left(\theta\right)=\iint\left[\frac{\pi\left(z|\,v;\theta\right)-\pi\left(z|\,v;\theta_{0}\right)}{\pi\left(z|\,v;\theta_{0}\right)}\right]^{2}\pi_{2}\left(z,v;\theta_{0}\right)dzdv,\qquad z\equiv y_{t},v\equiv y_{t-1},\theta=\left(b,\sigma\right).$$

We wish to find parameter restrictions such that the function $f(z,v;\theta)$ defined as,

$$f(z,v;\theta) \equiv \frac{\pi(z|v;\theta) - \pi(z|v;\theta_0)}{\pi(z|v;\theta_0)} \sqrt{\pi_2(z,v;\theta_0)},$$

is bounded and integrable.

In the context of model (F4),

$$\begin{aligned} \pi(z|v;\theta) &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\frac{(z-bv)^2}{\sigma^2}\right] \\ \pi_2(z,v;\theta) &= \frac{\sqrt{1-b^2}}{2\pi\sigma^2} \exp\left(-\frac{z^2-2bvz+v^2}{2\sigma^2}\right) \end{aligned}$$

and,

$$f(z,v;\theta) \leq \frac{\left(1-b_0^2\right)^{\frac{1}{4}}}{\sqrt{2\pi}\sigma_0} \left[f_1(z,v;\theta) + f_2(z,v;\theta)\right],$$

where

$$\begin{aligned} f_1(z,v;\theta) &\equiv \frac{\sigma_0}{\sigma} \exp\left(-\frac{c_1 z^2 + c_2 vz + c_3 v^2}{4\sigma^2\sigma_0^2}\right) \\ c_1 &\equiv 2\sigma_0^2 - \sigma^2 \\ c_2 &\equiv 2\sigma^2 b_0 - 4\sigma_0^2 b \\ c_3 &\equiv 2\sigma_0^2 b^2 + \left(1-2b_0^2\right)\sigma^2 \\ f_2(z,v;\theta) &\equiv \exp\left(-\frac{z^2-2b_0vz+v^2}{4\sigma_0^2}\right). \end{aligned}$$

Clearly, f_2 is bounded and integrable, and so is f_1 whenever $c_1 > 0$ and $c_2 > 0$, i.e. whenever,

$$\begin{cases} \sigma < \sqrt{2}\sigma_0 \\ 0 < 2\sigma_0^2 b^2 + \left(1-2b_0^2\right)\sigma^2 \end{cases} \tag{F5}$$

If $b_0^2 \leq \frac{1}{2}$, the system of inequalities (F5) holds whenever $\sigma < \sqrt{2}\sigma_0$ and $b \in (-1,1)$. If $b_0^2 \in \left(\frac{1}{2},1\right)$ and $b_0 > 0$, the corresponding admissible region is the shaded area in Figure 3 (the case $b_0 < 0$ is similar). Note that in this case, θ_0 can only lie below the straight line,

$$\sigma = \frac{\sqrt{2}\sigma_0}{\sqrt{-\left(1-2b_0^2\right)}} b.$$

For suppose not. Then $\sigma_0 > \left[-\left(1-2b_0^2\right)\right]^{-\frac{1}{2}} \sqrt{2}\sigma_0 b_0 \Leftrightarrow -1 + 2b_0^2 > 2b_0^2$, which is impossible.

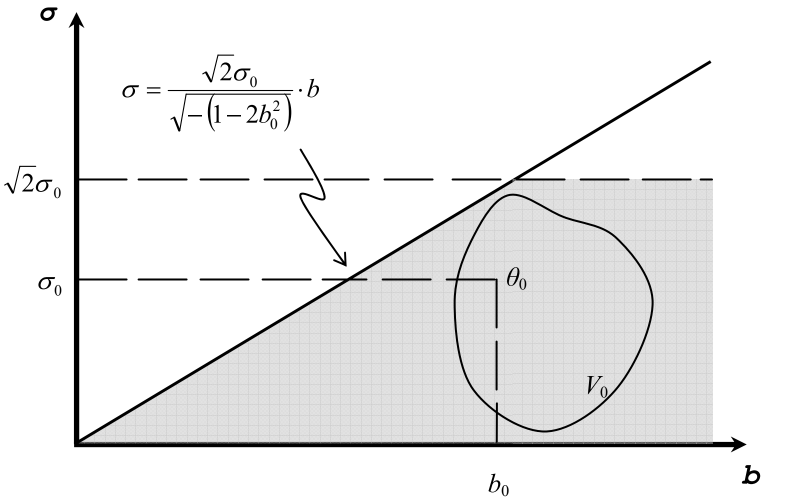

Figure 3 - Parameter restrictions ensuring the existence of the Neyman's measure of distance in the Gaussian AR(1) model (F4). If $b_0 > 0$, the Neyman's measure of distance $\mathcal{N}(\theta)$ exists for all the parameter values $\theta = (\sigma, b)$ in the shaded area. The case $b_0 < 0$ is symmetric.

References

Aït-Sahalia, Y., 1994, "The Delta Method for Nonparametric Kernel Functionals," working paper, Princeton University.

Aït-Sahalia, Y., 1996, "Testing Continuous-Time Models of the Spot Interest Rate," *Review of Financial Studies*, 9, 385–426.

Altissimo, F. and A. Mele, 2008, "Simulated Nonparametric Estimation of Dynamic Models."

Andrews, D.W.K., 1992, "Generic Uniform Convergence," *Econometric Theory*, 8, 241-257.

Andrews, D.W.K., 1995, "Nonparametric Kernel Estimation for Semiparametric Models," *Econometric Theory*, 11, 560-596.

Arcones, M.A. and B. Yu, 1994, "Central Limit Theorems for Empirical and U-Processes of Stationary Mixing Sequences," *Journal of Theoretical Probability*, 7, 47-71.

Arnold, L., 1992, *Stochastic Differential Equations*: *Theory and Applications.* Malabar, Florida: Krieger Publishing Company.

Carrasco, M., L.P. Hansen and X. Chen, 1999, "Time Deformation and Dependence," working paper, University of Rochester.

Chen, X., L.P. Hansen and M. Carrasco, 1999, "Nonlinearity and Temporal Dependence," working paper, University of Rochester.

Devroye, L. and L. Györfy, 1985, *Nonparametric Density Estimation–The* L_1 *View.* Wiley Series in Probability and Mathematical Statistics. New York: John Wiley and Sons.

Dudley, R.M., 1990, "Nonlinear Functionals of Empirical Measures and the Bootstrap," in: Eberlein, E., J. Kuelbs and M.B. Marcus (Eds.), *Probability in Banach Spaces 7: Proceedings (Progress in Probability).* Boston: Birkhäuser.

Duffie, D. and K.J. Singleton, 1993, "Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica*, 61, 929-952.

Glick, N., 1974, "Consistency Conditions for Probability Estimators and Integrals of Density Estimators," *Utilitas Mathematica*, 6, 61-74.

Kloeden, P.E. and E. Platen, 1999, *Numerical Solutions of Stochastic Differential Equations*, Berlin: Springer Verlag.

Meyn, S.P. and R.L. Tweedie, 1993, "Stability of Markovian Processes III: Foster-Lyapunov Criteria for Continuous-Time Processes," *Advances in Applied Probability*, 25, 518-548.

Newey, W.K., 1991, "Uniform Convergence in Probability and Stochastic Equicontinuity," *Econometrica*, 59, 1161-1167.

Newey, W. K. and D. L. McFadden, 1994, "Large Sample Estimation and Hypothesis Testing," in Engle, R.F. and D. L. McFadden (eds.), *Handbook of Econometrics*, vol. 4, chapter 36, 2111-2245. Amsterdam: Elsevier.

Politis, D.N. and J.P. Romano, 1994, "Limit Theorems for Weakly Dependent Hilbert Space Valued Random Variables with Applications to the Stationary Bootstrap," *Statistica Sinica*, 4, 461-476.

Tjøstheim D., 1990, "Non-Linear Time Series and Markov Chains," *Advances in Applied Probability,* 22, 587-611.

White, H., 1994, *Estimation, Inference and Specification Analysis*, Cambridge: Cambridge University Press.

Wooldridge, J.M., 1994, "Estimation and Inference for Dependent Processes," in: Engle, R.F. and D.L. McFadden (Eds.), *Handbook of Econometrics*, vol. 4, chapter 45, 2639-2738. Amsterdam: Elsevier.