Biometrika (2017), **xx**, x, *pp*. 1–18

Printed in Great Britain

Supplementary Material for "On the Asymptotic Efficiency of Approximate Bayesian Computation Estimators"

BY WENTAO LI

School of Mathematics, Statistics and Physics, Newcastle University, Newcastle upon Tyne NE1 7RU, U.K.

wentao.li@newcastle.ac.uk

AND PAUL FEARNHEAD

Department of Mathematics and Statistics, Lancaster University, Lancaster LA1 4YF, U.K.

p.fearnhead@lancaster.ac.uk

Technical lemmas and proofs of the main results are presented. In the following, the data are considered to be random, and $O(\cdot)$ and $\Theta(\cdot)$ denote the limiting behaviour when n goes to ∞. For two sets A and B, the sum of integrals $\int_A f(x)\,dx + \int_B f(x)\,dx$ is written as $(\int_A + \int_B) f(x)\,dx$. Recall that $T_{\rm obs} = a_n A(\theta_0)^{-1/2}\{s_{\rm obs} - s(\theta_0)\}$ and by Condition 4, $T_{\rm obs} \to N(0, I_d)$ in distribution, where I_d is the identity matrix with dimension d. For a constant $d \times p$ matrix A, let the minimum and maximum eigenvalues of $A^T A$ be $\lambda^2_{\min}(A)$ and $\lambda^2_{\max}(A)$. Obviously for any p-dimension vector x, $\lambda_{\min}(A)\|x\| \leq \|Ax\| \leq \lambda_{\max}(A)\|x\|$. For two matrices A and B, we say A is bounded by B if $\lambda_{\max}(A) \leq \lambda_{\min}(B)$.

1. PROOF OF RESULTS FROM SECTION 3

1·1. *Overview and Notation*

We first give an overview of the proof to Theorem 1. The convergence of the maximum likelihood estimator based on the summary follows almost immediately from Creel & Kristensen (2013). The minor extensions we used are summarized in Lemmas 1 and 2 below.

The main challenge with Theorem 1 are the results about the posterior mean of approximate Bayesian computation. For the convergence of posterior means of approximate Bayesian computation we need to consider convergence of integrals over the parameter space, $\mathbb{R}^p$. We will divide $\mathbb{R}^p$ into $B_\delta = \{\theta : \|\theta - \theta_0\| < \delta\}$ and B_δ^c for some $\delta < \delta_0$, and introduce the notation $\pi(h) = \int h(\theta)\pi(\theta) f_{\rm ABC}(s_{\rm obs} \mid \theta)\,d\theta$. The posterior mean of approximate Bayesian computation is $h_{\rm ABC} = \pi(h)/\pi(1)$. We can write $\pi(h)$, say, as $\pi(h) = \pi_{B_\delta}(h) + \pi_{B_\delta^c}(h)$, where

$$\pi_{B_\delta}(h) = \int_{B_\delta} h(\theta)\pi(\theta) f_{\rm ABC}(s_{\rm obs} \mid \theta)\,d\theta, \quad \pi_{B_\delta^c}(h) = \int_{B_\delta^c} h(\theta)\pi(\theta) f_{\rm ABC}(s_{\rm obs} \mid \theta)\,d\theta.$$

As $n \to \infty$ the posterior distribution of approximate Bayesian computation concentrates around θ_0. The first step of our proof is to show that, as a result, the contribution that comes from integrating over B_δ^c can be ignored. Hence we need consider only $\pi_{B_\delta}(h)/\pi_{B_\delta}(1)$.

Second, we perform a Taylor expansion of $h(\theta)$ around θ_0. Let $Dh(\theta)$ and $Hh(\theta)$ denote the vector of first derivatives and the matrix of second derivatives of $h(\theta)$ respectively. Then

$$h(\theta) = h(\theta_0) + Dh(\theta_0)^T(\theta - \theta_0) + \frac{1}{2}(\theta - \theta_0)^T Hh(\theta_*)(\theta - \theta_0),$$

for some θ_*, that depends on θ and that satisfies $||\theta_* - \theta_0|| < ||\theta - \theta_0||$. We plug this into $\pi_{B_\delta}(h)$, but re-express the integrals in term of the rescaled random vector

$$t(\theta) = a_{n,\varepsilon}(\theta - \theta_0),$$

and let $t(B_\delta)$ be the set $\{\phi : \phi = t(\theta) \text{ for some } \theta \in B_\delta\}$. This gives

$$\frac{\pi_{B_\delta}(h)}{\pi_{B_\delta}(1)} = h(\theta_0) + a_{n,\varepsilon}^{-1} Dh(\theta_0)^T \frac{\pi_{B_\delta}(t)}{\pi_{B_\delta}(1)} + \frac{1}{2} a_{n,\varepsilon}^{-2} \frac{\pi_{B_\delta}\{t^T Hh(\theta_t) t\}}{\pi_{B_\delta}(1)}, \tag{1}$$

where we write t for $t(\theta)$, and θ_t is the value θ_* from remainder term in the Taylor expansion for $h(\theta)$. We use the notation θ_t to emphasize its dependence on t, and note that θ_t belongs to B_δ.

Let $\widetilde{f}_{\text{ABC}}(s_{\text{obs}} \mid \theta) = \int \widetilde{f}_n(s_{\text{obs}} + \varepsilon_n v \mid \theta) K(v)\, dv$, which is the likelihood approximation that we get if we replace the true likelihood by its Gaussian limit, and define $\widetilde{\pi}_{B_\delta}(h) = \int_{B_\delta} h(\theta)\pi(\theta)\widetilde{f}_{\text{ABC}}(s_{\text{obs}} \mid \theta)\, d\theta$. Our third step is to re-write (1) as

$$\frac{\pi_{B_\delta}(h)}{\pi_{B_\delta}(1)} = h(\theta_0) + a_{n,\varepsilon}^{-1} Dh(\theta_0)^T \frac{\widetilde{\pi}_{B_\delta}(t)}{\widetilde{\pi}_{B_\delta}(1)} + a_{n,\varepsilon}^{-1} Dh(\theta_0)^T \left\{ \frac{\widetilde{\pi}_{B_\delta}(t)}{\widetilde{\pi}_{B_\delta}(1)} - \frac{\pi_{B_\delta}(t)}{\pi_{B_\delta}(1)} \right\}$$
$$+ \frac{1}{2} a_{n,\varepsilon}^{-2} \frac{\pi_{B_\delta}\{t^T Hh(\theta_t) t\}}{\pi_{B_\delta}(1)}.$$

We bound the size of the last two terms, so that asymptotically h_{ABC} behaves as

$$h(\theta_0) + a_{n,\varepsilon}^{-1} Dh(\theta_0)^T \frac{\widetilde{\pi}_{B_\delta}(t)}{\widetilde{\pi}_{B_\delta}(1)}.$$

If we introduce the density $g_n(t, v)$, defined as $g_n(t, v, \tau)$ in Section 4·3 of the main text but with $\tau = 0$, so

$$g_n(t, v) \propto \begin{cases} N\Big\{Ds(\theta_0)t; a_n\varepsilon_n v + A(\theta_0)^{1/2} T_{\text{obs}}, A(\theta_0)\Big\} K(v), & a_n\varepsilon_n \to c < \infty, \\ N\Big\{Ds(\theta_0)t; v + \frac{1}{a_n\varepsilon_n} A(\theta_0)^{1/2} T_{\text{obs}}, \frac{1}{a_n^2\varepsilon_n^2} A(\theta_0)\Big\} K(v), & a_n\varepsilon_n \to \infty, \end{cases}$$

then we can show that

$$\frac{\widetilde{\pi}_{B_\delta}(t)}{\widetilde{\pi}_{B_\delta}(1)} \approx \frac{\int_{t(B_\delta)} \int_{\mathbb{R}^d} t g_n(t, v)\, dt dv}{\int_{t(B_\delta)} \int_{\mathbb{R}^d} g_n(t, v)\, dt dv},$$

with a remainder that can be ignored. Putting this together, we get that asymptotically h_{ABC} is

$$h(\theta_0) + a_{n,\varepsilon}^{-1} Dh(\theta_0)^T \frac{\int_{t(B_\delta)} \int_{\mathbb{R}^d} t g_n(t, v)\, dt dv}{\int_{t(B_\delta)} \int_{\mathbb{R}^d} g_n(t, v)\, dt dv},$$

and the proof finishes by calculating the form of this.

A recurring theme in the proofs for the bounds on the various remainders is the need to bound expectations of polynomials of either the rescaled parameter t, or a rescaled difference in the summary statistic from s_{obs}, or both. Later we will present a lemma, stated in terms of a general polynomial, that is used repeatedly to obtain the bounds we need.

To define this we need to introduce a set of suitable polynomials. For any integer l and vector x, if a scalar function of x has the expression $\sum_{i=0}^{l} \alpha_i(x, n)^T x^i$, where for each i, x^i denotes the vector with all monomials of x with degree i as elements and $\alpha_i(x, n)$ is a vector of functions of x and n, we denote it by $P_l(x)$. Let $\mathbb{P}_{l,x}$ be the set

$$\{P_l(x) : \text{for all } i \leq l,\ \text{as } n \to \infty,\ \alpha_i(x, n) = O_p(1) \text{ holds uniformly in } x\}$$

To simplify the notations, for two vectors x_1 and x_2, $P_l\{(x_1^T, x_2^T)^T\}$ and $\mathbb{P}_{l,(x_1^T,x_2^T)^T}$ are written as $P_l(x_1, x_2)$ and $\mathbb{P}_{l,(x_1,x_2)}$. Where the specific form of the polynomial does not matter, and we only use the fact that it lies in $\mathbb{P}_{l,x}$, we will often simplify expressions by writing it as $P_l(x)$.

1·2. *Proof of Theorem 1*

For the maximum likelihood estimator based on the summary, Creel & Kristensen (2013) gives the central limit theorem for $\hat{\theta}_{\text{MLES}}$ when $a_n = n^{1/2}$ and $\mathcal{P}$ is compact. According to the proof in Creel & Kristensen (2013), extending the result to the general a_n is straightforward. Additionally, we give the extension for general $\mathcal{P}$.

LEMMA 1. *Assume Conditions 1,4-6. Then* $a_n(\hat{\theta}_{\text{MLES}} - \theta_0) \to N\{0, I^{-1}(\theta_0)\}$ *in distribution as* $n \to \infty$.

Given Condition 3, by Lemma 1 and the delta method (Lehmann, 2004), the convergence of the maximum likelihood estimator for general $h(\theta)$ holds as follows.

LEMMA 2. *Assume the conditions of Lemma 1 and Condition 3. Then* $a_n\{h(\hat{\theta}_{\text{MLES}}) - h(\theta_0)\} \to N\{0, Dh(\theta_0)^T I^{-1}(\theta_0) Dh(\theta_0)\}$ *in distribution as* $n \to \infty$.

The following lemmas are used for the result about the posterior mean of approximate Bayesian computation, proofs of these are given in Section 1·3. Our first lemma is used to justify ignoring integrals over B_δ^c.

LEMMA 3. *Assume Conditions 2, 3–6. Then for any* $\delta < \delta_0$, $\pi_{B_\delta^c}(h) = O_p(e^{-a_{n,\varepsilon}^{\alpha_\delta} c_\delta})$ *for some positive constants* c_δ *and* α_δ *depending on* δ.

The following lemma is used to calculate the form of

$$\frac{\int_{t(B_\delta)} \int_{\mathbb{R}^d} t g_n(t, v)\, dt dv}{\int_{t(B_\delta)} \int_{\mathbb{R}^d} g_n(t, v)\, dt dv},$$

which is the leading term for $\{h_{\text{ABC}} - h(\theta_0)\}$.

LEMMA 4. *Assume Condition 2. Let* c *be a constant vector,* $\{k_n\}$ *be a series converging to* $k_\infty \in (0, \infty]$ *and* $\{b_n'\}$ *be a series converging to a non-negative constant. Let* $b_n = \mathbb{1}_{\{k_\infty = \infty\}} + b_n' \mathbb{1}_{\{k_\infty < \infty\}}$. *Then for any* $d \times p$ *constant matrix* A *and any* $d \times d$ *constant matrix* B,

$$\int_{\mathbb{R}^p} \int_{\mathbb{R}^d} t \frac{N(At; B_n v + \frac{1}{k_n} c, \frac{1}{k_n^2} I_d) K(v)}{\int_{\mathbb{R}^p} \int_{\mathbb{R}^d} N(At; B_n v + \frac{1}{k_n} c, \frac{1}{k_n^2} I_d) K(v)\, dt dv} dt dv = \frac{1}{k_n} \left\{(A^T A)^{-1} A^T c + R(A, B_n, k_n, c)\right\},$$

where $B_n = b_n B$, *the expression of* $R(c; A, B_n, k_n)$ *is stated in the proof. Specifically,* $R(A, B_n, k_n, c) = o(1)$ *when* $B_n = o(1)$ *and* $O(1)$ *otherwise.*

Our final two lemmas are used to bound the remainder terms in the expansion for h_{ABC} we presented in Section 1·1.

LEMMA 5. *Assume Conditions 1, 2 and 4 hold. If $\varepsilon_n = o({a_n}^{-1/2})$, there exists a $\delta < \delta_0$ such that*

$$\widetilde{\pi}_{B_\delta}(1) = a_{n,\varepsilon}^{d-p}\Big\{\pi(\theta_0)\int_{t(B_\delta)}\int_{\mathbb{R}^d} g_n(t,v)\,dvdt + O_p(a_{n,\varepsilon}^{-1}) + O_p(a_n^2\varepsilon_n^4)\Big\},$$

$$\int_{t(B_\delta)}\int_{\mathbb{R}^d} g_n(t,v)\,dtdv = \Theta_p(1),$$

$$\frac{\widetilde{\pi}_{B_\delta}(t)}{\widetilde{\pi}_{B_\delta}(1)} = \frac{\int_{t(B_\delta)}\int_{\mathbb{R}^d} tg_n(t,v)\,dtdv}{\int_{t(B_\delta)}\int_{\mathbb{R}^d} g_n(t,v)\,dtdv} + O_p(a_{n,\varepsilon}^{-1}) + O_p(a_n^2\varepsilon_n^4), \qquad (2)$$

and $\widetilde{\pi}_{B_\delta}\{P_2(t)\}/\widetilde{\pi}_{B_\delta}(1) = O_p(1)$ for any $P_2(t) \in \mathbb{P}_{2,t}$.

LEMMA 6. *Assume the conditions of Lemma 5 and Conditions 3 and 5. Then if $\varepsilon_n = o({a_n}^{-1/2})$, there exists a $\delta < \delta_0$ such that*

$$\frac{\pi_{B_\delta}(h)}{\pi_{B_\delta}(1)} = h(\theta_0) + a_{n,\varepsilon}^{-1}Dh(\theta_0)^T\left\{\frac{\widetilde{\pi}_{B_\delta}(t)}{\widetilde{\pi}_{B_\delta}(1)} + O_p(\alpha_n^{-1})\right\} + \frac{1}{2}a_{n,\varepsilon}^{-2}\left[\frac{\widetilde{\pi}_{B_\delta}\{t^THh(\theta_t)t\}}{\widetilde{\pi}_{B_\delta}(1)} + O_p(\alpha_n^{-1})\right], \qquad (3)$$

Now we are ready to prove Theorem 1.

Proof of Theorem 1. The convergence of the maximum likelihood estimator based on the summary is given by Lemma 1 and Lemma 2.

We now focus on the convergence for the posterior mean of approximate Bayesian computation. The convergence of the posterior mean given the summaries follows from a similar, but simpler, argument and is omitted.

We can bound $t^TH(\theta_t)t$ for θ in B_δ by the quadratic $t^TH_{max}t$, where H_{max} is an upper bound on $H(\theta_t)$ for θ_t in B_δ. This means that

$$\widetilde{\pi}_{B_\delta}\{t^THh(\theta_t)t\} = O(1).$$

Together with Lemmas 3, 5 and 6, we then have the expansion

$$h_{\text{ABC}} = h(\theta_0) + a_{n,\varepsilon}^{-1}Dh(\theta_0)^T\left\{\frac{\int_{t(B_\delta)\times\mathbb{R}^d} tg_n(t,v)\,dtdv}{\int_{t(B_\delta)\times\mathbb{R}^d} g_n(t,v)\,dtdv} + O_p(a_{n,\varepsilon}^{-1}) + O_p(a_n^2\varepsilon_n^4) + O_p(\alpha_n^{-1})\right\}.$$

The analytical form of the integral in the above expansion, which we will denote by $E_{g_n}(t)$, can be obtained by applying Lemma 4 with $A = A(\theta_0)^{-1/2}DS(\theta_0)$, $c = T_{\text{obs}}$,

$$B_n = \begin{cases} a_n\varepsilon_n A(\theta_0)^{-1/2}, & c_\varepsilon < \infty, \\ A(\theta_0)^{-1/2}, & c_\varepsilon = \infty, \end{cases} \qquad k_n = \begin{cases} 1, & c_\varepsilon < \infty, \\ a_n\varepsilon_n, & c_\varepsilon = \infty. \end{cases}$$

It can be seen that $E_{g_n}(t)$ is $\Theta_p(k_n^{-1})$, and the remainder term, $O_p(a_{n,\varepsilon}^{-1}) + O_p(a_n^2\varepsilon_n^4) + O_p(\alpha_n^{-1})$, is $o_p(1)$ as $\varepsilon_n = o({a_n}^{-3/5})$ and $\alpha_n^{-1} = o({a_n}^{-2/5})$. Then since $a_{n,\varepsilon}^{-1}k_n^{-1} = a_n^{-1}$, we have

$$\begin{aligned} &a_n\{h_{\text{ABC}} - h(\theta_0)\} \\ =&Dh(\theta_0)^T\left[\left\{Ds(\theta_0)^TA(\theta_0)^{-1}Ds(\theta_0)\right\}^{-1}Ds(\theta_0)^TA(\theta_0)^{-1/2}T_{\text{obs}} + R_n(a_n\varepsilon_n, T_{\text{obs}})\right] + o_p(1), \end{aligned} \qquad (4)$$

where $R_n(a_n\varepsilon_n, T_{\rm obs})$ is $Dh(\theta_0)^T R(A, B_n, k_n, c)$ with $R(A, B_n, k_n, c)$ defined in Lemma 4. We can interpret $R_n(a_n\varepsilon_n, T_{\rm obs})$ as the extra variation brought by ε_n: $a_n[h_{\rm ABC} - E\{h(\theta) \mid s_{\rm obs}\}]$.

By the delta method, the first term in the right hand side of (4) converges to $I(\theta_0)^{-1/2}Z$. For the second term, since $A(A^TA)^{-1}A^T$ is a projection matrix, by eigen decompositition

$$I - A(A^TA)^{-1}A^T = U\begin{pmatrix} 0 & 0 \\ 0 & I_{d-p}\end{pmatrix}U^T,\ (A^TA)^{-1/2}A^T = \begin{pmatrix} I_p & 0 \\ 0 & 0\end{pmatrix}U^T,$$

where U is an orthogonal matrix. For a vector x, let $x_{k_1:k_2}$ be the $(k_2 - k_1 + 1)$-dimension vector containing the k_1th–k_2th coordinates of x. Let $v' = U^TA(\theta_0)^{-1/2}v$, and $T'_{\rm obs} = U^TT_{\rm obs}$. Then $R_n(a_n\varepsilon_n, T_{\rm obs})$ can be written as

$$R_n(a_n\varepsilon_n, T_{\rm obs})$$
$$= Dh(\theta_0)^T(A^TA)^{-1/2}a_n\varepsilon_n \frac{\int v'_{1:p} N\{v'_{(p+1):d}; -\frac{1}{a_n\varepsilon_n}T'_{{\rm obs},(p+1):d}, \frac{1}{a_n^2\varepsilon_n^2}I_{d-p}\}K\{A(\theta_0)^{1/2}Uv'\}\, dv'}{\int N\{v'_{(p+1):d}; -\frac{1}{a_n\varepsilon_n}T'_{{\rm obs},(p+1):d}, \frac{1}{a_n^2\varepsilon_n^2}I_{d-p}\}K\{A(\theta_0)^{1/2}Uv'\}\, dv'}. \tag{5}$$

Denote the weak limit of $R_n(a_n\varepsilon_n, T_{\rm obs})$ as $R(c_\varepsilon, Z)$. When $d = p$, obviously $R_n(a_n\varepsilon_n, T_{\rm obs}) = 0$ and therefore $R(c_\varepsilon, Z) = 0$. When $d > p$, if $\varepsilon_n = o(1/a_n)$, $R_n(a_n\varepsilon_n, T_{\rm obs}) = o_p(1)$ by Lemma 4 and therefore $R(c_\varepsilon, Z) = 0$. When the covariance matrix of $K(\cdot)$ is $c^2A(\theta_0)$, for constant $c > 0$, $K(v) \propto \overline{K}\{c\|A(\theta_0)^{-1/2}v\|^2\}$. Then $K\{A(\theta_0)^{1/2}Uv'\}$ in (5) can be replaced by $\overline{K}(c\|v'\|^2)$ and for fixed $v'_{(p+1):d}$, the integrand in the numerator, as a function of $v'_{1:p}$, is symmetric around zero. Therefore $R_n(a_n\varepsilon_n, T_{\rm obs}) = 0$ and $R(c_\varepsilon, Z) = 0$.

Otherwise, $R_n(a_n\varepsilon_n, z)$ is not necessarily zero. Since for any n, $R_n(a_n\varepsilon_n, z)$ as a function of z is symmetric around 0, $R(c_\varepsilon, z)$ is also symmetric and $R(c_\varepsilon, Z)$ has mean zero. Since $I^{-1}(\theta_0)$ is the Cramer-Rao lower bound, ${\rm var}\{I(\theta_0)^{-1/2}Z + R(c_\varepsilon, Z)\} \geq I^{-1}(\theta_0)$.

For (i), the asymptotic normality holds for $h(\hat\theta)$ by Lemma 2. □

1·3. *Proof of Lemmas*

Here we give the proofs of lemmas from Section 1·2.

Proof of Lemma 3. It is sufficient to show that for any δ, $\sup_{\theta\in B_\delta^c} f_{\rm ABC}(s_{\rm obs} \mid \theta) = O_p(e^{-a_{n,\varepsilon}^{\alpha_\delta}c_\delta})$. By dividing $\mathbb{R}^d$ into $\{v : \|\varepsilon_n v\| \leq \delta'/3\}$ and its complement, we have

$$\sup_{\theta\in B_\delta^c} f_{\rm ABC}(s_{\rm obs} \mid \theta) = \sup_{\theta\in B_\delta^c}\int_{\mathbb{R}^d} f_n(s_{\rm obs} + \varepsilon_n v \mid \theta)K(v)\, dv$$
$$\leq \sup_{\theta\in B_\delta^c\setminus\mathcal{P}_0^c}\left\{\sup_{\|s-s_{\rm obs}\|\leq\delta'/3} f_n(s \mid \theta)\right\} + \sup_{\theta\in\mathcal{P}_0^c}\left\{\sup_{\|s-s_{\rm obs}\|\leq\delta'/3} f_n(s \mid \theta)\right\} + \bar{K}(\lambda_{\min}(\Lambda)\varepsilon_n^{-1}\delta'/3)\varepsilon_n^{-d},$$

where $\lambda_{\min}(\Lambda)$ is positive. In the above, as $n \to \infty$, the third term is exponentially decreasing by Conditions 2(iv). For the second term, by Condition 4, with probability 1,

$$\|s - s(\theta)\| = \|\{s(\theta_0) - s(\theta)\} + \{s_{\rm obs} - s(\theta_0)\} + \varepsilon_n v\|$$
$$\geq \delta' - \delta'/3 - \delta'/3 = \delta'/3.$$

Recall that $W_n(s) = a_nA(\theta)^{-1/2}\{s - s(\theta)\}$. Then by Condition 6, the second term is exponentially decreasing. For the first term, when $\theta \in B_\delta^c\setminus\mathcal{P}_0^c$ and $\|s - s_{\rm obs}\| \leq \delta'/3$, $\|W_n(s)\| \geq a_n\delta' r$ for some constant r. By Condition 5 and 6, $f_{W_n}(w \mid \theta)$ is bounded by the sum of a normal density and $\alpha_n^{-1} r_{\max}(w)$, which are both exponentially decreasing, so $\sup_{\theta \in B_\delta^c \setminus \mathcal{P}_0^c} \sup_{\|s-s_{\mathrm{obs}}\| \leq \delta'/3} f_n(s \mid \theta)$ is also exponentially decreasing. Finally, the sum of all the above is $O(e^{-a_{n,\varepsilon}^{\alpha_\delta} c_\delta})$ by noting that $a_{n,\varepsilon} \leq \min(\varepsilon_n^{-1}, a_n)$. □

The following additional lemma will be used repeatedly to bound error terms that appear in Lemmas 5 and 6.

LEMMA 7. *Assume Condition 2. For $t \in \mathbb{R}^p$ and $v \in \mathbb{R}^d$, let $\{A_n(t)\}$ be a series of $d \times p$ matrix functions, $\{C_n(t)\}$ be a series of $d \times d$ matrix functions, Q be a positive definite matrix and $g_1(v)$ and $g_2(v)$ be probability densities in $\mathbb{R}^d$. Let c be a random vector, $\{k_n\}$ be a series converging to $k_\infty \in (0, \infty]$ and $\{b'_n\}$ be a series converging to a non-negative constant. Let $b_n = \mathbb{1}_{\{k_\infty = \infty\}} + b'_n \mathbb{1}_{\{k_\infty < \infty\}}$. If*

(i) $g_1(v)$ and $g_2(v)$ are bounded in $\mathbb{R}^d$;

(ii) $g_1(v)$ and $g_2(v)$ depend on v only through $\|v\|$ and are decreasing functions of $\|v\|$;

(iii) there exists an integer l such that $\int \prod_{k=1}^{l+p} v_{i_k} g_j(v)\, dv < \infty$, $j = 1, 2$, for any coordinates $(v_{i_1}, \cdots, v_{i_l})$ of v;

(iv) there exists a positive constant m such that for any $t \in \mathbb{R}^p$ and n, $\lambda_{\min}\{A_n(t)\}$ and $\lambda_{\min}\{C_n(t)\}$ are greater than m;

then for any $P_l(t, v) \in \mathbb{P}_{l,(t,v)}$,

$$\int_{\mathbb{R}^p} \int_{\mathbb{R}^d} P_l(t, v) k_n^d g_1[k_n C_n(t)\{A_n(t)t - b_n v - k_n^{-1} c\}] g_2(Qv)\, dv dt = O_p(1),$$
$$\int_{\mathbb{R}^p} \int_{\mathbb{R}^d} k_n^d g_1[k_n C_n(t)\{A_n(t)t - b_n v - k_n^{-1} c\}] g_2(Qv)\, dv dt = \Theta_p(1).$$

Proof. For simplicity, here $\int$ denotes the integration over the whole Euclidean space. According to (ii), $g_1(v)$ can be written as $\bar{g}_1(\|v\|)$. When $k_\infty < \infty$, assume $k_n = 1$ without loss of generality. For any $P_l(t, v) \in \mathbb{P}_{l,(t,v)}$, by Cauchy–Schwarz inequality, there exists a $P_l(\|t\|, \|v\|) \in \mathbb{P}_{l,(\|t\|,\|v\|)}$ with coefficient functions taking positive values such that $|P_l(t, v)|$ is bounded by $P_l(\|t\|, \|v\|)$ almost surely. Therefore for the first equality, it is sufficient to consider the equality where $P_l(t, v)$ is replaced by $P_l(\|t\|, \|v\|)$ and the coefficient functions of $P_l(\|t\|, \|v\|)$ are positive almost surely. For each n, divide $\mathbb{R}^p$ into $V = \{t : \|A_n(t)t\|/2 \geq \|b'_n v + c\|\}$ and V^c. In V, $\|C_n(t)\{A_n(t)t - b'_n v - c\}\| \geq m^2 \|t\|/2$; in V^c, $\|t\| \leq 2m^{-1} \|b'_n v + c\|$. With probability tending to 1,

$$\int P_l(\|t\|, \|v\|) g_1[C_n(t)\{A_n(t)t - b'_n v - c\}] g_2(Qv)\, dv dt \leq$$
$$\int P_l(\|t\|, \|v\|) \bar{g}_1(m^2 \|t\|/2) g_2(Qv)\, dv dt + \sup_{v \in \mathbb{R}^d} g_1(v) \int \int_{V^c} dt\, P_l(2m^{-1} \|b'_n v + c\|, \|v\|) g_2(Qv)\, dv.$$

In the above, $\int_{V^c} dt$ is the volume of V^c in $\mathbb{R}^p$ and is proportional to $\|b'_n v + c\|^p$. By (iii), the right hand side of the above inequality is $O_p(1)$.

When $k_\infty = \infty$, let $v^* = k_n\{A(t)t - v - k_n^{-1}c\}$. Then for any $P_l(t,v) \in \mathbb{P}_{l,(t,v)}$, with probability 1,

$$\left|\int P_l(t,v) k_n^d g_1[k_n C_n(t)\{A(t)t - v - k_n^{-1}c\}] g_2(Qv)\, dvdt\right|$$
$$= \left|\int P_l(t,v^*) g_2[Q\{A(t)t - k_n^{-1}v^* - k_n^{-1}c\}] g_1(C_n(t)v^*)\, dv^*dt\right|,$$
$$\leq \int P_l(\|t\|,\|v^*\|) g_2[Q\{A(t)t - k_n^{-1}v^* - k_n^{-1}c\}] \overline{g}_1(m\|v^*\|)\, dv^*dt$$

for some $P_l(t,v^*) \in \mathbb{P}_{l,(t,v^*)}$ and $P_l(\|t\|,\|v^*\|) \in \mathbb{P}_{l,(\|t\|,\|v^*\|)}$. The right hand side of the above inequality is similar to the integral when $k_\infty < \infty$ with $g_1(\cdot)$ and $g_2(\cdot)$ replaced by $g_2(\cdot)$ and $\overline{g}_1(\cdot)$ respectively. Therefore it is $O_p(1)$ by the same reasoning.

For $P_l(t,v) = 1$, by considering only the integral in a compact region, it is easy to see the target integral is larger than 0. Therefore the lemma holds. □

Proof of Lemma 4. Let $P = A^TA$. By matrix algebra,

$$N\Big(At; B_nv + \frac{1}{k_n}c, \frac{1}{k_n^2}I_d\Big)K(v) = N\Big\{t; P^{-1}A^T\left(B_nv + \frac{1}{k_n}c\right), \frac{1}{k_n^2}P^{-1}\Big\} r(v; A, B_n, k_n, c),$$

where

$$r(v; A, B_n, k_n, c) = \frac{k_n^{d-p}}{(2\pi)^{(d-p)/2}} \exp\Big\{-\frac{k_n^2}{2}\left(B_nv + \frac{c}{k_n}\right)^T (I - AP^{-1}A^T)\left(B_nv + \frac{c}{k_n}\right)\Big\}K(v).$$

Then the target integral can be expanded as

$$\int t\frac{N(At; B_nv + \frac{1}{k_n}c, \frac{1}{k_n^2}I_d)K(v)}{\int N(At; B_nv + \frac{1}{k_n}c, \frac{1}{k_n^2}I_d)K(v)\, dtdv}\, dtdv = \int P^{-1}A^T\left(\frac{1}{k_n}c + B_nv\right)\frac{r(v; A, B_n, k_n, c)}{\int r(v; A, B_n, k_n, c)\, dv}\, dv$$
$$= \frac{1}{k_n}\left\{(A^TA)^{-1}A^Tc + R(A, B_n, k_n, c)\right\},$$

where

$$R(A, B_n, k_n, c) = (A^TA)^{-1}A^TB_n\int k_nv\frac{r(v; A, B_n, k_n, c)}{\int r(v; A, B_n, k_n, c)\, dv}\, dv.$$

The remainder term $R(A, B_n, k_n, c)$ depends on the mean of the probability density proportional to $r(v; A, B_n, k_n, c)$ in the directions of $(A^TA)^{-1}A^TB$. If B_n does not degenerate to 0 as $n \to \infty$, then in the directions orthogonal to those of $(I - A(A^TA)^{-1}A^T)^{1/2}B$, $r(v; A, B_n, k_n, c)$ is symmetric around 0; in the directions of $(I - A(A^TA)^{-1}A^T)^{1/2}B$, $r(v; A, B_n, k_n, c)$ is a product of a normal density whose mean is $O(1/k_n)$ and a rescaled $K(v)$, which is symmetric around 0, so its mean value is $O(1/k_n)$. Therefore when the spaces expanded by $(A^TA)^{-1}A^TB$ and $\{I - A(A^TA)^{-1}A^T\}B$ are orthogonal, $R(A, B_n, k_n, c) = 0$; when it is not the case, $R(A, B_n, k_n, c) = O(1)$.

If $B_n = o(1)$ as $n \to \infty$, which implies $k_n \to c \in (0,\infty)$, it is easy to see that $\int k_nvr(v; A, B_n, k_n, c)\, dv / \int r(v; A, B_n, k_n, c)\, dv$ is upper bounded as $n \to \infty$ and hence $R(A, B_n, k_n, c)$ is $o(1)$. □

In the following lemmas, to deal with the case where $K(x) = \bar{K}(||x||_\Lambda)$ with Λ not the identity, we use the property that such a $K(x)$ can be bounded above by a function that depends only on $||x||$. We refer to this bound as $K(\cdot)$ rescaled to have identity covariance matrix.

Proof of Lemma 5. First consider $\widetilde{\pi}_{B_\delta}(1)$. With the transformation $t = t(\theta)$,

$$\widetilde{\pi}_{B_\delta}(1) = a_{n,\varepsilon}^{-p} \int_{t(B_\delta)} \int_{\mathbb{R}^d} \pi(\theta_0 + a_{n,\varepsilon}^{-1} t) \widetilde{f}_n(s_{\rm obs} + \varepsilon_n v \mid \theta_0 + a_{n,\varepsilon}^{-1} t) K(v)\, dv dt. \tag{6}$$

We can obtain an expansion of $\widetilde{\pi}_{B_\delta}(1)$ by expanding $\widetilde{f}_n(s_{\rm obs} + \varepsilon_n v \mid \theta_0 + a_{n,\varepsilon}^{-1} t)K(v)$ as follows. The expansion needs to be discussed separately for two cases, depending on whether the limit of $a_n\varepsilon_n$ is finite or infinite.

When $a_n\varepsilon_n \to c_\varepsilon < \infty$, $a_{n,\varepsilon} = a_n$. We apply a Taylor expansion to $s(\theta_0 + a_n^{-1}t)$ and $A(\theta_0 + a_n^{-1}t)^{-1/2}$ and have

$$\widetilde{f}_n(s_{\rm obs} + \varepsilon_n v \mid \theta_0 + a_n^{-1} t) = \frac{a_n^d}{|A(\theta_0 + a_n^{-1}t)|^{1/2}}$$
$$\times N\left(\left\{A(\theta_0)^{-1/2} + a_n^{-1} r_A(t,\epsilon_2)\right\}\left[A(\theta_0)^{1/2}T_{\rm obs} + a_n\varepsilon_n v - \{Ds(\theta_0) + a_n^{-1}r_s(t,\epsilon_1)\}t\right]; 0, I_d\right), \tag{7}$$

where $r_s(t,\epsilon_1)$ is the $d \times p$ matrix whose ith row is $t^T H s_i\{\theta_0 + \epsilon_1(t)\}$, $r_A(t,\epsilon_2)$ is the $d \times d$ matrix $\sum_{k=1}^p \frac{d}{d\theta_k} A\{\theta_0 + \epsilon_2(t)\}^{-1/2} t_k$, and $\epsilon_1(t)$ and $\epsilon_2(t)$ are from the remainder terms of the Taylor expansions and satisfy $\|\epsilon_1(t)\| \leq \delta$ and $\|\epsilon_2(t)\| \leq \delta$. For a $d \times d$ matrix τ_2, let $g_n(t, v; \tau_1, \tau_2)$ be the function $g_n(t, v; \tau_1)$, defined in Section 4·3 of the main text, with $A(\theta_0)$ replaced by $\{A(\theta_0)^{-1/2} + \tau_2\}^{-2}$. Applying a Taylor expansion to the normal density in (7), we have

$$\widetilde{f}_n(s_{\rm obs} + \varepsilon_n v \mid \theta_0 + a_n^{-1} t) K(v)$$
$$= \frac{a_n^d |A(\theta_0)|^{1/2}}{|A(\theta_0 + a_n^{-1}t)|^{1/2}} \left[g_n(t,v) + a_n^{-1} P_3(t,v) g_n\{t, v; e_{n1} r_s(t,\epsilon_1), e_{n1} r_A(t,\epsilon_2)\}\right], \tag{8}$$

where $P_3(t,v)$ is the function

$$\frac{1}{2|A(\theta_0)^{-1/2} + r_2(a_n^{-1}t)|}$$
$$\times \left.\frac{d}{dx}\left\|\left\{A(\theta_0)^{-1/2} + x r_A(t,\epsilon_2)\right\}\left[A(\theta_0)^{1/2}T_{\rm obs} + a_n\varepsilon_n v - \{Ds(\theta_0) + x r_s(t,\epsilon_1)\}t\right]\right\|^2\right|_{x=e_{n1}},$$

and e_{n1} is from the remainder term of Taylor expansion and satisfies $|e_{n1}| \leq a_n^{-1}$. Since $\|e_{n1}t\| \leq \delta$ and $r_s(t,\epsilon_1)$ and $r_A(t,\epsilon_2)$ belong $\mathbb{P}_{1,t}$, this $P_3(t,v)$ belongs to $\mathbb{P}_{3,(t,v)}$. Furthermore, since $r_s(t,\epsilon_1)$ and $r_A(t,\epsilon_2)$ have no constant term, for any small σ, $e_{n1}r_s(t,\epsilon_1)$ and $e_{n1}r_A(t,\epsilon_2)$ can be bounded by σI_d and σI_p uniformly in n and t, if δ is small enough.

When $a_n\varepsilon_n \to \infty$, $a_{n,\varepsilon} = \varepsilon_n^{-1}$. Let $v^*(v) = A(\theta_0)^{1/2}T_{\rm obs} + a_n\varepsilon_n v - a_n\varepsilon_n Ds(\theta_0)t$. Under the transformation $v^* = v^*(v)$, the expansion of $\widetilde{f}_n(s_{\rm obs} + \varepsilon_n v \mid \theta_0 + \varepsilon_n t)$ obtained by applying a Taylor expansion to $s(\theta_0 + \varepsilon_n t)$ and $A(\theta_0 + \varepsilon_n t)^{-1/2}$ is

$$\widetilde{f}_n(s_{\rm obs} + \varepsilon_n v \mid \theta_0 + \varepsilon_n t)$$
$$= \frac{a_n^d}{|A(\theta_0 + a_n^{-1}t)|^{1/2}} N\left[\left\{A(\theta_0)^{-1/2} + a_n\varepsilon_n^2 \frac{r_A(t,\epsilon_4)}{a_n\varepsilon_n}\right\}\left\{v^* - a_n\varepsilon_n^2 r_s(t,\epsilon_3)t\right\}; 0, I_d\right],$$

where $\epsilon_3(t)$ and $\epsilon_4(t)$ are from the remainder terms of the Taylor expansion and satisfy $\|\epsilon_3(t)\| \leq \delta$ and $\|\epsilon_4(t)\| \leq \delta$. Let $g_n^*(t, v^*; \tau_1, \tau_2)$ be the function

$$g_n^*(t, v^*; \tau_1, \tau_2) \\ = N\left[v^*; a_n\varepsilon_n\tau_1 t, \{A(\theta_0)^{-1/2} + \tau_2\}^{-2}\right] K\left\{Ds(\theta_0)t + \frac{1}{a_n\varepsilon_n}v^* - \frac{1}{a_n\varepsilon_n}A(\theta_0)^{1/2}T_{\text{obs}}\right\},$$

so that $(a_n\varepsilon_n)^d g_n^*(t, v^*; \tau_1, \tau_2)$ is $g_n(t, v; \tau_1, \tau_2)$ with transformed variable $v^* = v^*(v)$, and $g_n^*(t, v^*) = g_n^*(t, v^*; 0, 0)$. Denote a $k_1 \times k_2$ matrix with element being $P_l(t)$ by $P_l^{(k_1\times k_2)}(t)$. Then by applying a Taylor expansion to the normal density in the expansion above,

$$\widetilde{f}_n(s_{\text{obs}} + \varepsilon_n v \mid \theta_0 + \varepsilon_n t)K(v) \\ = \frac{\varepsilon_n^{-d}|A(\theta_0)|^{1/2}}{|A(\theta_0 + \varepsilon_n t)|^{1/2}}\Big[g_n^*(t, v^*) + a_n\varepsilon_n^2\left\{P_2^{(d\times 1)}(t)v^* + \frac{1}{a_n\varepsilon_n}v^{*T}P_1^{(d\times d)}(t)v^*\right\}g_n^*(t, v^*) \\ + (a_n\varepsilon_n^2)^2 P_4(t, v^*)g_n^*\{t, v^*; e_{n2}r_s(t, \epsilon_3), e_{n2}r_A(t, \epsilon_4)\}\Big](a_n\varepsilon_n)^d, \tag{9}$$

where $P_2^{(d\times 1)}(t)$ is the function $t^T r_s(t, \epsilon_3)^T A(\theta_0)^{-1/2}/2$, $P_1^{(d\times d)}(t)$ is the function $-A(\theta_0)^{-1/2}r_A(t, \epsilon_4)$, $e_{n2} = e'_{n2}/(a_n\varepsilon_n)$, e'_{n2} is from the remainder term of the Taylor expansion and satisfies $|e'_{n2}| \leq a_n\varepsilon_n^2$, and $P_4(t, v^*)$ is a linear combination of $\{d\rho(w)/dw\}^2$ and $d^2\rho(w)/dw^2$ at $w = e'_{n2}$ with $\rho(w)$ being the function

$$\left\|\left\{A(\theta_0)^{-1/2} + w\frac{r_A(t, \epsilon_4)}{a_n\varepsilon_n}\right\}\left\{v^* - wr_s(t, \epsilon_3)t\right\}\right\|^2.$$

Obviously elements of $P_2^{(d\times 1)}(t)$ and $P_1^{(d\times d)}(t)$ belong to $\mathbb{P}_{2,t}$ and $\mathbb{P}_{1,t}$ respectively. Since $\|e_{n2}t\| \leq \delta$, the function $P_4(t, v^*)$ belongs to $\mathbb{P}_{4,(t,v^*)}$ and, similar to before, $e_{n2}r_s(t, \epsilon_3)$ and $e_{n2}r_A(t, \epsilon_4)$ can be bounded by σI_d and σI_p uniformly in n and t for any small σ, if δ is small enough.

For $\pi(\theta_0 + a_{n,\varepsilon}^{-1}t)$ in the integral of $\widetilde{\pi}_{B_\delta}(1)$ in (6), a Taylor expansion gives that

$$\frac{\pi(\theta_0 + a_{n,\varepsilon}^{-1}t)}{|A(\theta_0 + a_{n,\varepsilon}^{-1}t)|^{1/2}} = \frac{\pi(\theta_0)}{|A(\theta_0)|^{1/2}} + a_{n,\varepsilon}^{-1}D_\theta\frac{\pi\{\theta_0 + \epsilon_5(t)\}}{|A\{\theta_0 + \epsilon_5(t)\}|^{1/2}}t, \quad |\epsilon_5(t)| \leq \delta. \tag{10}$$

As mentioned before, δ can be selected such that $Ds(\theta_0) + e_{n1}r_s(t, \epsilon_1)$ and $Ds(\theta_0) + e_{n2}r_s(t, \epsilon_3)$ are lower bounded by $m_1 I_p$ and $A(\theta_0)^{-1/2} + e_{n1}r_A(t, \epsilon_2)$ and $A(\theta_0)^{-1/2} + e_{n2}r_A(t, \epsilon_4)$ are lowered bounded by $m_2 I_d$ for some positive constant m_1 and m_2. We choose δ satisfying these and, since $\|a_{n,\varepsilon}^{-1}t\| \leq \delta$, this means $\pi(\theta_0 + a_{n,\varepsilon}^{-1}t)/|A(\theta_0 + a_{n,\varepsilon}^{-1}t)|^{1/2}$ is bounded uniformly in t and n.

By plugging (8)–(10) into (6), it can be seen that the leading term of $\widetilde{\pi}_{B_\delta}(1)$ is $a_{n,\varepsilon}^{d-p}\pi(\theta_0)\int_{t(B_\delta)\times\mathbb{R}^d} g_n(t,v)\,dtdv$. The remainder terms are given in the following,

$$
\begin{aligned}
& a_{n,\varepsilon}^{p-d}\widetilde{\pi}_{B_\delta}(1) - \pi(\theta_0)\int_{t(B_\delta)\times\mathbb{R}^d} g_n(t,v)\,dtdv \\
=& a_{n,\varepsilon}^{-1}\int_{t(B_\delta)\times\mathbb{R}^d} |A(\theta_0)|^{1/2} D\frac{\pi(\theta_0+\epsilon_5)}{|A(\theta_0+\epsilon_5)|^{1/2}} t g_n(t,v)\,dvdt \\
&+ a_n^{-1}\int_{t(B_\delta)\times\mathbb{R}^d} P_3(t,v) g_n\{t,v;e_{n1}r_s(t,\epsilon_1),e_{n1}r_A(t,\epsilon_2)\}\,dvdt\ \mathbb{1}_{\{\lim a_n\varepsilon_n<\infty\}} \\
&+ a_n\varepsilon_n^2\int_{t(B_\delta)} P_2^{(d\times 1)}(t)\int_{\mathbb{R}^d} v^* g_n^*(t,v^*)\,dv^* dt\ \mathbb{1}_{\{\lim a_n\varepsilon_n=\infty\}} \\
&+ \varepsilon_n\int_{t(B_\delta)\times\mathbb{R}^d} v^{*T} P_1^{(d\times d)}(t) v^* g_n^*(t,v^*)\,dv^* dt\ \mathbb{1}_{\{\lim a_n\varepsilon_n=\infty\}} \\
&+ a_n^2\varepsilon_n^4\int_{t(B_\delta)\times\mathbb{R}^d} P_4(t,v^*) g_n^*\{t,v^*;e_{n2}r_s(t,\epsilon_3),e_{n2}r_A(t,\epsilon_4)\}\,dv^* dt\ \mathbb{1}_{\{\lim a_n\varepsilon_n=\infty\}},
\end{aligned}
\tag{11}
$$

where $P_3(t,v)$, $P_2^{(d\times 1)}(t)$, $P_1^{(d\times d)}(t)$ and $P_4(t,v^*)$ are products of $\pi(\theta_0+a_{n,\varepsilon}^{-1}t)/|A(\theta_0+a_{n,\varepsilon}^{-1}t)|^{1/2}$ and corresponding terms in expansions (8) and (9). In the above, there are five remainder terms. For the integrals in the first two terms, it is easy to write them in the form of the first integral in Lemma 7 and conditions therein are satisfied, where $g_1(\cdot)$ is the standard normal density and $g_2(\cdot)$ is $K(v)$ rescaled to have identity covariance. Then the first two terms are $O_p(a_{n,\varepsilon}^{-1})$ and $O_p(a_n^{-1})$. The integral in the fourth term can also be written in this form where $g_1(\cdot)$ is the rescaled $K(v)$ and $g_2(\cdot)$ is the standard normal density. The integral in the fifth term needs to use the transformation $v^{**}=v^*-a_n\varepsilon_n e_{n2}r_s(t,\epsilon_3)t$, after which it can be written in a similar form, as $P_5\{t,v^{**}+a_n\varepsilon_n e_{n2}r_s(t,\epsilon_3)t\}\in\mathbb{P}_{5,(t,v^{**})}$ by the expression of $P_4(t,v^*)$ in (9). Thus the fourth and fifth term are $O_p(\varepsilon_n)$ and $O_p(a_n^2\varepsilon_n^4)$.

The third term is somewhat different as the center of $g_n^*(t,v^*)$ in the direction of v^* degenerates to zero as $n\to\infty$. Let ψ_k be the d-dimension unit vector with 1 at the kth coordinate. Then

$$
\begin{aligned}
\int_{-\infty}^{\infty} v_k^* g_n^*(t,v^*)\,dv_k^* &= \int_0^{\infty} v_k^*\{g_n^*(t,v^*)-g_n^*(t,v^*-2v_k^*\psi_k)\}\,dv_k^* \\
&= \int_0^{\infty} v_k^* N\{v^*;0,A(\theta_0)\}[K\{v(v^*)\}-K\{v(v^*-2v_k^*\psi_k)\}]\,dv_k^*,
\end{aligned}
$$

which by a Taylor expansion is bounded by $(a_n\varepsilon_n)^{-1}c$ for some constant c. Hence the third term is $O_p(\varepsilon_n)$. Combining the orders of all remainder terms, the expansion of $\widetilde{\pi}_{B_\delta}(1)$ in the lemma holds.

For any $P_2(t)\in\mathbb{P}_{2,t}$, $\widetilde{\pi}_{B_\delta}\{P_2(t)\}$ can be expanded similarly to $\widetilde{\pi}_{B_\delta}(1)$ in (11), simply by multplying $P_2(t)$ into every integral in (11). This gives that

$$
\widetilde{\pi}_{B_\delta}\{P_2(t)\} = a_{n,\varepsilon}^{d-p}\Big\{\pi(\theta_0)\int_{t(B_\delta)\times\mathbb{R}^d} P_2(t)g_n(t,v)\,dtdv + O_p(a_{n,\varepsilon}^{-1}) + O_p(a_n^2\varepsilon_n^4)\Big\}.
$$

Then since $\int_{t(B_\delta)\times\mathbb{R}^d} g_n(t,v)\,dtdv=\Theta_p(1)$ by the second result of Lemma 7, $\widetilde{\pi}_{B_\delta}\{P_2(t)\}/\widetilde{\pi}_{B_\delta}(1)=O_p(1)$ and (2) holds by taking $P_2(t)=t$. □

Proof of Lemma 6. Let $r_n(s \mid \theta)$ be the scaled remainder $\alpha_n\{f_n(s \mid \theta) - \widetilde{f}_n(s \mid \theta)\}$. The error of using $\widetilde{\pi}_{B_\delta}\{P_l(t)\}$ to approximate $\pi_{B_\delta}\{P_l(t)\}$ is

$$\pi_{B_\delta}\{P_l(t)\} - \widetilde{\pi}_{B_\delta}\{P_l(t)\} = \alpha_n^{-1} \int_{B_\delta} \int P_l\{t(\theta)\}\pi(\theta) r_n(s_{\text{obs}} + \varepsilon_n v \mid \theta) K(v)\, dv d\theta.$$

If this approximation error satisfies

$$\frac{\pi_{B_\delta}\{P_l(t)\} - \widetilde{\pi}_{B_\delta}\{P_l(t)\}}{\widetilde{\pi}_{B_\delta}(1)} = O_p(\alpha_n^{-1}), \tag{12}$$

then, since $a_{n,\varepsilon}^{p-d}\widetilde{\pi}_{B_\delta}(1) = \Theta_p(1)$ by Lemma 5,

$$\pi_{B_\delta}(1) = \widetilde{\pi}_{B_\delta}(1)\{1 + O_p(\alpha_n^{-1})\}, \quad \frac{\pi_{B_\delta}\{P_l(t)\}}{\pi_{B_\delta}(1)} = \frac{\widetilde{\pi}_{B_\delta}\{P_l(t)\}}{\widetilde{\pi}_{B_\delta}(1)} + O_p(\alpha_n^{-1}). \tag{13}$$

By plugging (12) into (1),

$$\frac{\pi_{B_\delta}(h)}{\pi_{B_\delta}(1)} = h(\theta_0) + a_{n,\varepsilon}^{-1} Dh(\theta_0)^T \left\{ \frac{\widetilde{\pi}_{B_\delta}(t)}{\widetilde{\pi}_{B_\delta}(1)} + O_p(\alpha_n^{-1}) \right\} + \frac{1}{2} a_{n,\varepsilon}^{-2} \left[\frac{\widetilde{\pi}_{B_\delta}\{t^T Hh(\theta_t)t\}}{\widetilde{\pi}_{B_\delta}(1)} + O_p(\alpha_n^{-1}) \right]. \tag{14}$$

Verification of (12) is given by the following argument. With the transformation $t = t(\theta)$ we have

$$\pi_{B_\delta}\{P_l(t)\} - \widetilde{\pi}_{B_\delta}\{P_l(t)\} = \alpha_n^{-1} a_{n,\varepsilon}^{-p} \int_{t(B_\delta)} \int P_l(t)\pi(\theta_0 + a_{n,\varepsilon}^{-1}t) r_n(s_{\text{obs}} + \varepsilon_n v \mid \theta_0 + a_{n,\varepsilon}^{-1}t) K(v)\, dv dt.$$

Let $r_{W_n}(w \mid \theta) = \alpha_n\{f_{W_n}(w \mid \theta) - \widetilde{f}_{W_n}(w \mid \theta)\}$, and we have

$$r_n(s \mid \theta) = a_n^d |A(\theta)|^{-1/2} r_{W_n}[a_n A(\theta)^{-1/2}\{s - s(\theta)\} \mid \theta].$$

For the value of δ, we choose the smaller value of the one from Lemma 5 and the one such that $Ds(\theta)$ is lower bounded and $A(\theta)^{-1/2}$ is upper bounded by MI_d in B_δ for some $M > 0$. Since $r_{W_n}(w \mid \theta)$ is upper bounded by $r_{\max}(w)$ according to Condition 5, by applying a Taylor expansion to $s(\theta_0 + a_{n,\varepsilon}^{-1}t)$ we have

$$\begin{aligned} |\pi_{B_\delta}\{P_l(t)\} - \widetilde{\pi}_{B_\delta}\{P_l(t)\}| \le \alpha_n^{-1} a_{n,\varepsilon}^{d-p} \sup_{\theta \in B_\delta} |\pi(\theta) A(\theta)^{-1/2}| \int_{t(B_\delta)} \int |P_l(t)| (a_n a_{n,\varepsilon}^{-1})^d \\ r_{\max}\left[a_n a_{n,\varepsilon}^{-1} M \left\{ Ds(\theta_0 + \epsilon_t)t - a_{n,\varepsilon}\varepsilon_n v - \frac{1}{a_n a_{n,\varepsilon}^{-1}} A(\theta_0)^{1/2} T_{\text{obs}} \right\} \right] K(v)\, dv dt, \end{aligned}$$

where ϵ_t is from the remainder term of the Taylor expansion and satisfies $|\epsilon_t| \le \delta$. Since $\widetilde{\pi}_{B_\delta}(1) = \Theta_p(a_{n,\varepsilon}^{d-p})$ by Lemma 5, it is sufficient to show that the above integral is $O_p(1)$. This is immediate by noting that when either $\lim a_n\varepsilon_n \to \infty$ or $\lim a_n\varepsilon_n \to c_\varepsilon < \infty$, the above integral can be written in the form of the first integral in Lemma 7 and conditions therein are satisfied, where $g_1(\cdot)$ and $g_2(\cdot)$ are $r_{\max}(\cdot)$ and $K(\cdot)$ rescaled to have identity covariance matrix. □

2. Proof of Results from Section 4

2·1. *Proof of Proposition 2*

The proof of Proposition 2 follows the standard asymptotic argument of importance sampling. In the following we use the convention that for a vector x, the matrix xx^T is denoted by x^2.

Proof of Proposition 2. Algorithm 1 generates independent, indentically distributed triples, $(\phi_i, \theta_i, s_n^{(i)})$, where $(\theta_i, s_n^{(i)})$ is generated from $g_n(\theta)f(s_n \mid \theta)$, and, conditional on $s_n = s_n^{(i)}$, ϕ_i is generated from a Bernoulli distribution with probability $K_{\varepsilon_n}(s_n - s_{\text{obs}})$.

Now $\hat{h}$ can be expressed as a ratio of sample means of functions of these independent, identically distributed random variables. Thus we can use the standard delta method (Lehmann, 2004) for ratio statistics to show that the central limit theorem holds. Further we obtain that the limiting distribution has mean

$$\frac{E\{h(\theta_1)w_1\phi_1\}}{E(w_1\phi_1)} = \frac{E\{h(\theta_1)w_1K_{\varepsilon_n}(s_n^{(1)} - s_{\text{obs}})\}}{E\{w_1K_{\varepsilon_n}(s_n^{(1)} - s_{\text{obs}})\}} = \frac{\int h(\theta)\pi(\theta)f_n(s_n \mid \theta)K_{\varepsilon_n}(s_n - s_{\text{obs}})\, ds_n\, d\theta}{\int \pi(\theta)f_n(s_n \mid \theta)K_{\varepsilon_n}(s_n - s_{\text{obs}})\, ds_n\, d\theta},$$

which is equal to h_{ABC}. Its variance is

$$\begin{aligned}
&\frac{1}{E^2(w_1\phi_1)}\text{var}\{h(\theta_1)w_1\phi_1\} + \frac{E^2\{h(\theta_1)w_1\phi_1\}}{E^4(w_1\phi_1)}\text{var}(w_1\phi_1) - 2\frac{E\{h(\theta_1)w_1\phi_1\}}{E^3(w_1\phi_1)}\text{cov}\{h(\theta_1)w_1\phi_1, w_1\phi_1\}^T \\
=&p_{\text{acc},\pi}^{-2}\left[E\{h(\theta_1)^2w_1^2\phi_1\} - h_{\text{ABC}}^2p_{\text{acc},\pi}^2 + h_{\text{ABC}}^2\left\{E(w_1^2\phi_1) - p_{\text{acc},\pi}^2\right\}\right. \\
&\left.-2h_{\text{ABC}}\left\{E\{h(\theta_1)w_1^2\phi_1\} - h_{\text{ABC}}p_{\text{acc},\pi}^2\right\}^T\right] \\
=&p_{\text{acc},\pi}^{-2}E[\{h(\theta_1)^2 - 2h_{\text{ABC}}h(\theta_1)^T + h_{\text{ABC}}^2\}w_1^2K_{\varepsilon_n}(s_n^{(1)} - s_{\text{obs}})] \\
=&p_{\text{acc},\pi}^{-1}E_{\pi_{\text{ABC}}}\left\{(h(\theta) - h_{\text{ABC}})^2\frac{\pi(\theta)}{q_n(\theta)}\right\}.
\end{aligned}$$

In the above expression we used $p_{\text{acc},\pi} = E(w_1\phi_1)$. It is easy to verify that

$$\Sigma_{\text{ABC},n} = p_{\text{acc},\pi}^{-1}E_{\pi_{\text{ABC}}}\left\{(h(\theta) - h_{\text{ABC}})^2\frac{\pi(\theta)}{q_n(\theta)}\right\}, \tag{15}$$

as required. □

2·2. *Proof of Theorem 2*

For simplicity, a consider one-dimensional function $h(\theta)$. For multi-dimensional functions, the extension is trivial by considering each element of $\Sigma_{\text{IS},n}$ seperately. Denote $\{h(\theta) - h_{\text{ABC}}\}^2$ by $G_n(\theta)$. In Theorem 2(i), $\Sigma_{\text{IS},n}$ is just the ABC posterior variance of $h(\theta)$, and the derivation of its order is similar to that of h_{ABC} in Section 1 of this supplementary material. The result is stated in the following lemma.

LEMMA 8. *Assume the conditions of Theorem 1. Then* $\text{var}_{\pi_{\text{ABC}}}\{h(\theta)\} = O_p(a_{n,\varepsilon}^{-2})$.

Proof. Using the notation of Section 1, $\text{var}_{\pi_{\text{ABC}}}[h(\theta)] = \pi(G_n)/\pi(1)$. It follows immediately from Lemma 3 that

$$\text{var}_{\pi_{\text{ABC}}}\{h(\theta)\} = \frac{\pi_{B_\delta}(G_n)}{\pi_{B_\delta}(1)}\{1 + o_p(1)\}.$$

Applying a first order Taylor expansion of $h(\theta)$ around $\theta = \theta_0$ gives

$$\frac{\pi_{B_\delta}(G_n)}{\pi_{B_\delta}(1)} = G_n(\theta_0) + 2a_{n,\varepsilon}^{-1}\{h(\theta_0) - h_{\text{ABC}}\}\frac{\pi_{B_\delta}\{Dh(\theta_t)^Tt\}}{\pi_{B_\delta}(1)} + a_{n,\varepsilon}^{-2}\frac{\pi_{B_\delta}\{t^TDh(\theta_t)Dh(\theta_t)^Tt\}}{\pi_{B_\delta}(1)}, \tag{16}$$

where θ_t is from the remainder term and belongs to B_δ. In the above decomposition, $G_n(\theta_0)$ and $a_{n,\varepsilon}^{-1}\{h(\theta_0) - h_{\text{ABC}}\}$ are $O_p(a_{n,\varepsilon}^{-2})$ by Theorem 1. Since $Dh(\theta_t)^Tt$ and $t^TDh(\theta_t)Dh(\theta_t)^Tt$ belong to $\mathbb{P}_{2,t}$, the two ratios in the above are $O_p(1)$ by Lemma 5 and Lemma 6. □

The following lemma states that moments of $K(v)^\gamma$ exist for any postive constant γ.

LEMMA 9. *Assume Condition 2. For any constant $\gamma \in (0,\infty)$ and coordinates $(v_{i_1},\cdots,v_{i_l})$ of v with $l \le p+6$, $\int \prod_{k=1}^{l} v_{i_k} K(v)^\gamma\, dv < \infty$.*

Proof. By Condition 2 (iv), for some positive constant M there exists $x_0 \in (0,\infty)$ such that when $\|v\| > x_0$, $K(v) < Me^{-c_1\|v\|^{\alpha_1}}$. Then consider the integration in two regions $\{v : \|v\| \le x_0\}$ and $\{v : \|v\| > x_0\}$ separately. In the first region, since $K(v) \le 1$, we have

$$\int_{\|v\|\le x_0} \prod_{k=1}^{l} v_{i_k} K(v)^\gamma\, dv \le x_0^l V_{x_0},$$

where V_{x_0} is the volume of the d-dimension sphere with radius x_0, and is finite. In the second region,

$$\int_{\|v\|>x_0} \prod_{k=1}^{l} v_{i_k} K(v)^\gamma\, dv \le M \int_{\|v\|>x_0} \|v\|^l e^{-c_1\gamma\|v\|^{\alpha_1}}\, dv.$$

The right hand side of this is proportional to $\exp\{-c_1\gamma x_0^{\alpha_1/(l+d)}\}$ by integrating in spherical coordinates. □

Proof of Theorem 2. For (i), since $p_{\mathrm{acc},\pi} = \varepsilon_n^d \pi(1)$ and $\pi(1) = \Theta_p(a_{n,\varepsilon}^{d-p})$ by Lemmas 3, 5 and 6, then $p_{\mathrm{acc},\pi} = \Theta_p(\varepsilon_n^d a_{n,\varepsilon}^{d-p})$. Together with Lemma 8, (i) holds.

For (ii), if we can show that $p_{\mathrm{acc},q} = \Theta_p(\varepsilon_n^d a_{n,\varepsilon}^d)$, then the order of $\Sigma_{\mathrm{IS},n}$ is obvious from (15) and the definition of $\Sigma_{\mathrm{ABC},n}$. Similar to the expansion of $\pi(1)$ from Lemma 3 and (13),

$$\begin{aligned} p_{\mathrm{acc},q} &= \varepsilon_n^d \int \pi_{\mathrm{ABC}}(\theta \mid s_{\mathrm{obs}}, \varepsilon_n) f_{\mathrm{ABC}}(s_{\mathrm{obs}} \mid \theta)\, d\theta \\ &= \varepsilon_n^d \left\{ \frac{\int_{B_\delta} \pi(\theta) \widetilde{f}_{\mathrm{ABC}}(s_{\mathrm{obs}} \mid \theta)^2\, d\theta}{\widetilde{\pi}_{B_\delta}(1)} + O_p(\alpha_n^{-1}) \right\} \{1 + o_p(1)\}. \end{aligned}$$

The integral in the above differs from $\widetilde{\pi}_{B_\delta}(1)$ by the square power of $\widetilde{f}_{\mathrm{ABC}}(s_{\mathrm{obs}} \mid \theta)$ in the integrand. We will show that this integral has order $\Theta_p(a_{n,\varepsilon}^{2d-p})$, from which $p_{\mathrm{acc},q} = \Theta_p(\varepsilon_n^d a_{n,\varepsilon}^d)$ trivially holds. Let $g_n^{**}(t, v; \tau_1, \tau_2)$ be the function

$$g_n^{**}(t, v; \tau_1, \tau_2) = N[v; 0, \{A(\theta_0)^{-1/2} + \tau_2\}^{-2}] K \left[\{Ds(\theta_0) + \tau_1\} t + \frac{1}{a_n\varepsilon_n} v^* - \frac{1}{a_n\varepsilon_n} A(\theta_0)^{1/2} T_{\mathrm{obs}} \right],$$

and $g_n^{**}(t, v; \tau_1, \tau_2) = g_n^*(t, v + a_n\varepsilon_n\tau_1 t; \tau_1, \tau_2)$. Here expansions (8) and (9) of $\widetilde{f}_n(s_{\mathrm{obs}} + \varepsilon_n v \mid \theta_0 + a_{n,\varepsilon}^{-1} t) K(v)$ are to be used in the form of

$$\frac{a_{n,\varepsilon}^d |A(\theta_0)|^{1/2}}{|A(\theta_0 + a_{n,\varepsilon}^{-1} t)|^{1/2}} \begin{cases} \left\{ g_n(t,v) + a_n^{-1} P_3(t,v) g_{n,r}(t,v) \right\}, & \lim_{n\to\infty} a_n\varepsilon_n < \infty, \\ \left\{ g_n^*(t,v^*) + a_n\varepsilon_n^2 P_3(t,v^*) g_n^*(t,v^*) \right. \\ \left. + (a_n\varepsilon_n^2)^2 P_4(t,v^{**}) g_{n,r}^*(t,v^{**}) \right\} (a_n\varepsilon_n)^d, & \lim_{n\to\infty} a_n\varepsilon_n = \infty, \end{cases} \tag{17}$$

where $P_3(t, v^*) \in \mathbb{P}_{3,(t,v^*)}$, $g_{n,r}(t,v) = g_n\{t, v; e_{n1} r_s(t,\epsilon_1), e_{n1} r_A(t,\epsilon_2)\}$, $g_{n,r}^*(t, v^{**})$ is $g_n^{**}\{t, v^{**}; e_{n2} r_s(t,\epsilon_3), e_{n2} r_A(t,\epsilon_4)\}$ and $P_4(t, v^{**})$ is $P_4(t, v^*)$ with the transformation $v^{**} = v^* - a_n\varepsilon_n e_{n2} r_s(t,\epsilon_3) t$, and the expansion of $\pi(\theta)/|A(\theta)|$ similar to (10) is to be used.

By the expression of $P_4(t,v^*)$ in (9), it can be seen that $P_4(t,v^{**}) \in \mathbb{P}_{4,(t,v^{**})}$. Basic inequalities $(a+\varepsilon b)^2 \le \varepsilon a^2 + (\varepsilon+\varepsilon^2)b^2$ and $(a+\varepsilon b+\varepsilon^2 c)^2 \le (\varepsilon+\varepsilon^2)a^2 + (\varepsilon+\varepsilon^2+\varepsilon^3)b^2 + (\varepsilon^2+\varepsilon^3+\varepsilon^4)c^2$ for any real constants a, b, c and ε, from the fact that $2ab \le a^2+b^2$, are also to be used. Then by the above expansions and inequalities, an expansion of the target integral similar to (11) can be obtained, with the leading term $a_{n,\varepsilon}^{2d-p}\pi(\theta_0)\int_{t(B_\delta)}\{\int g_n(t,v)\,dv\}^2\,dt$ and remainder term with the following upper bound

$$\begin{aligned}
&\left|a_{n,\varepsilon}^{p-2d}\int_{B_\delta}\pi(\theta)\widetilde{f}_{\mathrm{ABC}}(s_{\mathrm{obs}}\mid\theta)^2\,d\theta - \pi(\theta_0)\int_{t(B_\delta)}\left\{\int g_n(t,v)\,dv\right\}^2 dt\right| \\
\le & a_{n,\varepsilon}^{-1}\int_{t(B_\delta)}|A(\theta_0)|D_\theta\frac{\pi(\theta_0+\epsilon_6)}{|A(\theta_0+\epsilon_6)|}t\left\{\int g_n(t,v)\,dv\right\}^2 dt \\
&+M\int_{t(B_\delta)}\left[a_n^{-1}\left\{\int g_n(t,v)\,dv\right\}^2 + (a_n^{-1}+a_n^{-2})\left\{\int P_3(t,v)g_{n,r}(t,v)\,dv\right\}^2\right]dt\mathbb{1}_{\{\lim a_n\varepsilon_n<\infty\}} \\
&+M\int_{t(B_\delta)}\Bigg[\{a_n\varepsilon_n^2+(a_n\varepsilon_n^2)^2\}\left\{\int g_n^*(t,v^*)\,dv^*\right\}^2 \\
&\quad+\{a_n\varepsilon_n^2+(a_n\varepsilon_n^2)^2+(a_n\varepsilon_n^2)^3\}\left\{\int P_3(t,v^*)g_n^*(t,v^*)\,dv^*\right\}^2 \\
&\quad+\{(a_n\varepsilon_n^2)^2+(a_n\varepsilon_n^2)^3+(a_n\varepsilon_n^2)^4\}\left\{\int P_4(t,v^{**})g_{n,r}^*(t,v^{**})\,dv^{**}\right\}^2\Bigg]dt\mathbb{1}_{\{\lim a_n\varepsilon_n=\infty\}},
\end{aligned}$$

where M is the upper bound of $\pi(\theta)|A(\theta_0)|/|A(\theta)|$ for $\theta\in B_\delta$ with δ chosen so that M exists. Then if we can show that for any $P_4(t,v)\in\mathbb{P}_{5,(t,v)}$, $d\times p$ matrix function $r_{n1}(t)$ and $d\times d$ matrix function $r_{n2}(t)$ which can be bounded by σI_d and σI_p uniformly in n and t for any small δ if δ is small enough, (a)$\int_{t(B_\delta)}\left\{\int_{\mathbb{R}^d} g_n(t,v)\,dv\right\}^2 dt$ is $\Theta_p(1)$; (b) $\int_{t(B_\delta)}\left[\int_{\mathbb{R}^d}P_4(t,v)g_n\{t,v;r_{n1}(t),r_{n2}(t)\}\,dv\right]^2 dt$ is $O_p(1)$ when $\lim_{n\to\infty}a_n\varepsilon_n<\infty$; (c) $\int_{t(B_\delta)}\left[\int_{\mathbb{R}^d}P_4(t,v)g_n^{**}\{t,v;r_{n1}(t),r_{n2}(t)\}\,dv\right]^2 dt$ is $O_p(1)$ when $\lim_{n\to\infty}a_n\varepsilon_n=\infty$, the lemma would hold.

Here δ is selected such that $Ds(\theta_0)+r_{n1}(t)$ is bounded bounded by m_1I_p and $m_2I_d\le A(\theta_0)^{-1/2}+r_{n2}(t)\le M_2I_d$, for some positive constants m_1, m_2 and M_2, uniformly in n and t. For the purpose of bounding integrals, we can assume that $A(\theta_0)=I_d$ and $r_{n2}(t)=0$ without loss of generality by the following inequality when $\lim_{n\to\infty}a_n\varepsilon_n<\infty$,

$$g_n\{t,v;r_{n1}(t),r_{n2}(t)\}\le\frac{M_2^d}{(2\pi)^{d/2}}\exp\left[-\frac{m_2^2}{2}\|a_n\varepsilon_nv+A(\theta_0)^{1/2}T_{\mathrm{obs}}-\{Ds(\theta_0)+r_{n1}(t)\}t\|^2\right]K(v),$$

and a similar one for $g_n^{**}\{t,v;r_{n1}(t),r_{n2}(t)\}$.

Consider any $P_4(t,v)\in\mathbb{P}_{4,(t,v)}$. When $\lim_{n\to\infty}a_n\varepsilon_n<\infty$, let $E_1=\{v:\|a_n\varepsilon_nv\|^2\le\beta_1\|\{Ds(\theta_0)+r_{n1}(t)\}t-A(\theta_0)^{1/2}T_{\mathrm{obs}}\|^2\}$ for some $\beta_1\in(0,1)$. Then for any $\beta_2\in(0,1)$ we have

$$\begin{aligned}
&\int_{\mathbb{R}^d} P_4(t,v) g_n\{t,v;r_{n1}(t),r_{n2}(t)\}\,dv \\
\leq &\left(\int_{E_1} + \int_{E_1^c}\right) P_4(t,v)\frac{M_2^d}{(2\pi)^{d/2}} \exp\left[-\frac{m_2^2}{2}\|a_n\varepsilon_n v - \{Ds(\theta_0)+r_{n1}(t)\}t + A(\theta_0)^{1/2}T_{\text{obs}}\|^2\right] K(v)\,dv \\
\leq &P_4(t)\left(\exp\left[-\frac{m_2^2(1-\beta_1)}{2}\|\{Ds(\theta_0)+r_{n1}(t)\}t - A(\theta_0)^{1/2}T_{\text{obs}}\|^2\right]\right. \\
&\left.+\overline{K}^{\beta_2}\left[\frac{\lambda_{\min}^2(\Lambda)\beta_1}{a_n^2\varepsilon_n^2}\|\{Ds(\theta_0)+r_{n1}(t)\}t - A(\theta_0)^{1/2}T_{\text{obs}}\|^2\right]\right), \qquad (18)
\end{aligned}$$

where $P_4(t) \in \mathbb{P}_{4,t}$ and the above inequality uses Lemma 9. Then using $(a+b)^2 \leq 2(a^2+b^2)$,

$$\begin{aligned}
&\int_{t(B_\delta)}\left[\int_{\mathbb{R}^d} P_4(t,v) g_n\{t,v;r_{n1}(t),r_{n2}(t)\}\,dv\right]^2 dt \\
\leq &\int_{t(B_\delta)} P_8(t)\exp\left[-m_2^2(1-\beta_1)\|\{Ds(\theta_0)+r_{n1}(t)\}t - A(\theta_0)^{1/2}T_{\text{obs}}\|^2\right] dt \\
&+ \int_{t(B_\delta)} P_8(t)\overline{K}^{2\beta_2}\left[\frac{\lambda_{\min}^2(\Lambda)\beta_1}{a_n^2\varepsilon_n^2}\|\{Ds(\theta_0)+r_{n1}(t)\}t - A(\theta_0)^{1/2}T_{\text{obs}}\|^2\right] dt,
\end{aligned}$$

where $P_8(t) \in \mathbb{P}_{8,t}$.

When $a_n\varepsilon_n \to \infty$, let $E_2 = \{v : \|(a_n\varepsilon_n)^{-1}v\|^2 \leq \beta_1\|\{Ds(\theta_0)+r_{n1}(t)\}t - (a_n\varepsilon_n)^{-1}A(\theta_0)^{1/2}T_{\text{obs}}\|^2\}$ for some $\beta_1 \in (0,1)$. Then for any $\beta_2 \in (0,1)$ we have

$$\begin{aligned}
&\int_{\mathbb{R}^d} P_4(t,v) g_n^{**}\{t,v;r_{n1}(t),r_{n2}(t)\}\,dv \\
\leq &\left(\int_{E_2} + \int_{E_2^c}\right) P_4(t,v) K\left[\frac{1}{a_n\varepsilon_n}v + \{Ds(\theta_0)+r_{n1}(t)\}t - \frac{1}{a_n\varepsilon_n}A(\theta_0)^{1/2}T_{\text{obs}}\right] \qquad (19) \\
&\times \frac{M_2^d}{(2\pi)^{d/2}}\exp\left(-\frac{m_2^2}{2}\|v\|^2\right) dv \\
\leq &P_4(t)\left(\overline{K}\left[\lambda_{\min}^2(\Lambda)(1-\beta_1)\|\{Ds(\theta_0)+r_{n1}(t)\}t - \frac{1}{a_n\varepsilon_n}A(\theta_0)^{1/2}T_{\text{obs}}\|^2\right]\right. \\
&\left.+\exp\left[-\frac{a_n^2\varepsilon_n^2\beta_1 m_2^2\beta_2}{2}\|\{Ds(\theta_0)+r_{n1}(t)\}t - \frac{1}{a_n\varepsilon_n}A(\theta_0)^{1/2}T_{\text{obs}}\|^2\right]\right), \qquad (20)
\end{aligned}$$

where $P_4(t) \in \mathbb{P}_{4,t}$. Then using $(a+b)^2 \leq 2(a^2+b^2)$,

$$\begin{aligned}
&\int_{t(B_\delta)}\left[\int_{\mathbb{R}^d} P_4(t,v) g_n^{**}\{t,v;r_{n1}(t),r_{n2}(t)\}\,dv\right]^2 dt \\
\leq &\int_{t(B_\delta)} P_8(t)\overline{K}^2\left[\frac{\lambda_{\min}^2(\Lambda)(1-\beta_1)}{2}\|\{Ds(\theta_0)+r_{n1}(t)\}t - \frac{1}{a_n\varepsilon_n}A(\theta_0)^{1/2}T_{\text{obs}}\|^2\right] dt \\
&+ \int_{t(B_\delta)} P_8(t)\exp\left[-\frac{a_n^2\varepsilon_n^2\beta_1 m_2^2\beta_2}{2}\|\{Ds(\theta_0)+r_{n1}(t)\}t - \frac{1}{a_n\varepsilon_n}A(\theta_0)^{1/2}T_{\text{obs}}\|^2\right] dt
\end{aligned}$$

Applying Lemma 7 on these upper bounds, (b) and (c) hold.

For (a), to see that the limit of $\int_{t(B_\delta)}\{\int_{\mathbb{R}^d} g_n(t,v)\,dv\}^2\,dt$ is lower bounded away from zero, just use the positivity of the limit of the integrand and Fatou's lemma to interchange the order of limit and integral. □

2·3. *Proof of Theorem 3*

Now let $w_n(\theta)$ be the importance weight $\pi(\theta)/q_n(\theta)$, define $\pi_{B_\delta,\mathrm{IS}}(h) = \int_{B_\delta} h(\theta)\pi(\theta)f_{\mathrm{ABC}}(s_{\mathrm{obs}}\mid\theta)w_n(\theta)\,d\theta$ and define $\pi_{B_\delta^c,\mathrm{IS}}(h)$ correspondingly. Then by (15), we have

$$\Sigma_{ABC,n} = p_{\mathrm{acc},\pi}^{-1}\frac{\pi_{B_\delta,\mathrm{IS}}(G_n)+\pi_{B_\delta^c,\mathrm{IS}}(G_n)}{\pi_{B_\delta}(1)+\pi_{B_\delta^c}(1)}. \tag{21}$$

Proof of Theorem 3. For p_{acc,q_n}, we only need to consider the case when $\beta=0$. Recall that $t(\theta)=a_{n,\varepsilon}(\theta-\theta_0)$. By the transformation $t=t(\theta)$, since $a_{n,\varepsilon}\sigma_n=1$, $q_n(\theta)=a_{n,\varepsilon}^p|\Sigma|^{-1/2}q\{\Sigma^{-1/2}(t-c_\mu)\}$. Then, similar to the expansion of $\pi(1)$ from Lemma 3,

$$\begin{aligned}p_{\mathrm{acc},q_n} &= \varepsilon_n^d\int q_n(\theta)f_{\mathrm{ABC}}(s_{\mathrm{obs}}\mid\theta)\,d\theta\\ &= \varepsilon_n^d|\Sigma|^{-1/2}\int_{t(B_\delta)} q\{\Sigma^{-1/2}(t-c_\mu)\}\widetilde{f}_{\mathrm{ABC}}(s_{\mathrm{obs}}\mid\theta_0+a_{n,\varepsilon}^{-1}t)\,dt\{1+o_p(1)\}.\end{aligned}$$

The above integral differs from $\widetilde{\pi}_{B_\delta}(1)$ by replacing $\pi(\theta_0+a_{n,\varepsilon}^{-1}t)$ with the density $q\{\Sigma^{-1/2}(t-c_\mu)\}$ which does not degenerate to a constant as $n\to\infty$. We will show that this integral has order $\Theta_p(1)$. Plugging in the expansion (17) of $\widetilde{f}_{\mathrm{ABC}}(s_{\mathrm{obs}}\mid\theta_0+a_{n,\varepsilon}^{-1}t)$ into p_{acc,q_n}, we can obtain an expansion similar to (11), differing in that parts from expanding $\pi(\theta_0+a_{n,\varepsilon}^{-1}t)/|A(\theta_0+a_{n,\varepsilon}^{-1}t)|^{1/2}$ are replaced by the Taylor expansion

$$\frac{q\{\Sigma^{-1/2}(t-c_\mu)\}}{|A(\theta_0+a_{n,\varepsilon}^{-1}t)|^{1/2}} = q\{\Sigma^{-1/2}(t-c_\mu)\}\left[1+a_{n,\varepsilon}^{-1}D_\theta\frac{1}{|A\{\theta_0+\epsilon_6(t)\}|^{1/2}}t\right],$$

where $\|\epsilon_6(t)\|\le\delta$. The explicit form is ommitted here to avoid repetition. It can be seen that $p_{\mathrm{acc},q_n}=\Theta_p(a_{n,\varepsilon}^d\varepsilon_n^d)$ if (a) $\int_{\mathbb{R}^d\times t(B_\delta)} q\{\Sigma^{-1/2}(t-c_\mu)\}g_n(t,v)\,dvdt = \Theta_p(1)$; (b) $\int_{\mathbb{R}^d\times t(B_\delta)} P_3(t,v)q\{\Sigma^{-1/2}(t-c_\mu)\}g_n\{t,v;r_{n1}(t),r_{n2}(t)\}\,dvdt = O_p(1)$ when $\lim_{n\to\infty}a_n\varepsilon_n<\infty$; and (c) $\int_{\mathbb{R}^d\times t(B_\delta)} P_3(t,v)q\{\Sigma^{-1/2}(t-c_\mu)\}g_n^{**}\{t,v;r_{n1}(t),r_{n2}(t)\}\,dvdt = O_p(1)$ when $\lim_{n\to\infty}a_n\varepsilon_n=\infty$, where $r_{n1}(t)$ and $r_{n2}(t)$ are defined as in the proof of Theorem 2. Since $q\{\Sigma^{-1/2}(t-c_\mu)\}$ is uniformly upper bounded for $t\in\mathbb{R}^p$, (b) and (c) hold and the integral in (a) is $O_p(1)$ following the arguments for the similar cases in the proof of Theorem 2. By the positivity of the limit of the integrand and Fatou's lemma, the limit of the integral in (a) is lower bounded away from 0. Therefore $p_{\mathrm{acc},q_n}=\Theta_p(a_{n,\varepsilon}^d\varepsilon_n^d)$ holds.

As $\Sigma_{\mathrm{IS},n}$ is equal to $p_{\mathrm{acc},q_n}\Sigma_{ABC,n}$, by (15) we have

$$\Sigma_{\mathrm{IS},n} = \frac{p_{\mathrm{acc},q_n}}{p_{\mathrm{acc},\pi}}\frac{\pi_{B_\delta,\mathrm{IS}}(G_n)+\pi_{B_\delta^c,\mathrm{IS}}(G_n)}{\pi_{B_\delta}(1)+\pi_{B_\delta^c}(1)} = \frac{p_{\mathrm{acc},q_n}}{p_{\mathrm{acc},\pi}}\frac{\pi_{B_\delta,\mathrm{IS}}(G_n)}{\pi_{B_\delta}(1)}\{1+o_p(1)\},$$

where the second equality holds by noting that $\omega_n(\theta)\le\beta^{-1}$. Given the obtained orders of p_{acc,q_n} and $p_{\mathrm{acc},\pi}$, $\Sigma_{\mathrm{IS},n}=O_p(a_{n,\varepsilon}^{-2})$ if $\pi_{B_\delta,\mathrm{IS}}(G_n)/\pi_{B_\delta}(1)=O_p(a_{n,\varepsilon}^{-p-2})$. Similar to (16), we have the following expansion

$$\frac{\pi_{B_\delta,\mathrm{IS}}(G_n)}{\pi_{B_\delta}(1)} = G(\theta_0)\frac{\pi_{B_\delta,\mathrm{IS}}(1)}{\pi_{B_\delta}(1)}$$
$$+ 2a_{n,\varepsilon}^{-1}\{h(\theta_0) - h_{\mathrm{ABC}}\}\frac{\pi_{B_\delta,\mathrm{IS}}\{Dh(\theta_t)^T t\}}{\pi_{B_\delta}(1)} + a_{n,\varepsilon}^{-2}\frac{\pi_{B_\delta,\mathrm{IS}}\{t^T Dh(\theta_t)Dh(\theta_t)^T t\}}{\pi_{B_\delta}(1)},$$

and we only need $\pi_{B_\delta,\mathrm{IS}}\{P_2(t)\}/\pi_{B_\delta}(1) = O_p(a_{n,\varepsilon}^{-p})$ for any $P_2(t) \in \mathbb{P}_{2,t}$. Since $w_n(\theta) \leq (1-\beta)^{-1}w_{n,0}(\theta)$, where $w_{n,0}(\theta)$ is the weight when $\beta = 0$, it is sufficient to consider the case $\beta = 0$.

Similar to the proof of Theorem 1, first the normal counterpart $\widetilde{\pi}_{B_\delta,\mathrm{IS}}\{P_2(t)\}/\widetilde{\pi}_{B_\delta}(1)$ of $\pi_{B_\delta,\mathrm{IS}}\{P_2(t)\}/\pi_{B_\delta}(1)$, where $f_{\mathrm{ABC}}(s_{\mathrm{obs}} \mid \theta)$ is replaced by $\widetilde{f}_{\mathrm{ABC}}(s_{\mathrm{obs}} \mid \theta)$, is considered, then it is shown that their difference can be ignored. Using the transformation $t = t(\theta)$ and plugging in expansion (17) of $\widetilde{f}_{\mathrm{ABC}}(s_{\mathrm{obs}} \mid \theta_0 + a_{n,\varepsilon}^{-1}t)$ into $\widetilde{\pi}_{B_\delta,\mathrm{IS}}\{P_2(t)\}$, we obtain an expansion similar to (11), differing in that parts from expanding $\pi(\theta_0 + a_{n,\varepsilon}^{-1}t)/|A(\theta_0 + a_{n,\varepsilon}^{-1}t)|^{1/2}$ are replaced by the Taylor expansion

$$\frac{1}{q_n(\theta)}\frac{\pi(\theta_0 + a_{n,\varepsilon}^{-1}t)^2}{|A(\theta_0 + a_{n,\varepsilon}^{-1}t)|^{1/2}}$$
$$= \frac{1}{a_{n,\varepsilon}^{p}|\Sigma|^{-1/2}q\{\Sigma^{-1/2}(t-c_\mu)\}}\left[\pi(\theta_0)^2 + a_{n,\varepsilon}^{-1}D_\theta\frac{\pi\{\theta_0 + \epsilon_7(t)\}^2}{|A\{\theta_0 + \epsilon_7(t)\}|^{1/2}}t\right],$$

where $\|\epsilon_7(t)\| \leq \delta$. The explicit form is omitted here to avoid repetition. Then it can be seen that if we can show that

$$(d)\ \int_{t(B_\delta)}\frac{\int_{\mathbb{R}^d}P_5(t,v)g_n\{t,v;r_{n1}(t),r_{n2}(t)\}\,dv}{q\{\Sigma^{-1/2}(t-c_\mu)\}}\,dt = O_p(1) \text{ when } \lim_{n\to\infty}a_n\varepsilon_n < \infty,$$
$$(e)\ \int_{t(B_\delta)}\frac{\int_{\mathbb{R}^d}P_5(t,v)g_n^{**}\{t,v;r_{n1}(t),r_{n2}(t)\}\,dv}{q\{\Sigma^{-1/2}(t-c_\mu)\}}\,dt = O_p(1) \text{ when } \lim_{n\to\infty}a_n\varepsilon_n = \infty,$$

where $r_{n1}(t)$ and $r_{n2}(t)$ are defined as in the proof of Theorem 2, $\widetilde{\pi}_{B_\delta,\mathrm{IS}}\{P_2(t)\} = O_p(a_{n,\varepsilon}^{d-2p})$ and $\widetilde{\pi}_{B_\delta,\mathrm{IS}}\{P_2(t)\}/\widetilde{\pi}_{B_\delta}(1) = O_p(a_{n,\varepsilon}^{-p})$ by Lemma 5. By (18) and the following equality for $d\times p$ full column-rank matrix A and vector c,

$$\|At - c\| = \|P^{1/2}(t - P^{-1}Ac)\|^2 + c^T(I - AP^{-1}A^T)c,$$

where $P = A^TA$ and $P^{1/2}P^{1/2} = P$, for (d) we have

$$\frac{\int_{\mathbb{R}^d}P_5(t,v)g_n\{t,v;r_{n1}(t),r_{n2}(t)\}\,dv}{q\{\Sigma^{-1/2}(t-c_\mu)\}}$$
$$\leq P_5(t)\frac{\exp\left\{-\frac{m_1^2m_2^2\gamma}{2}\|t - P(\theta_0,t)T_{\mathrm{obs}}\|^2\right\}}{q\{\Sigma^{-1/2}(t-c_\mu)\}}\exp\left[-\frac{m_2^2\Delta}{2}\|\{Ds(\theta_0)+r_{n1}(t)\}t - A(\theta_0)^{1/2}T_{\mathrm{obs}}\|^2\right]$$
$$+ P_5(t)\frac{\overline{K}^\alpha\left\{\frac{\lambda_{\min}^2(\Lambda)(1-\gamma-\Delta)m_1^2}{a_n^2\varepsilon_n^2}\|t - P(\theta_0,t)T_{\mathrm{obs}}\|^2\right\}}{q\{\Sigma^{-1/2}(t-c_\mu)\}}$$
$$\times\overline{K}^\Delta\left[\frac{\lambda_{\min}^2(\Lambda)(1-\gamma-\Delta)}{a_n^2\varepsilon_n^2}\|\{Ds(\theta_0)+r_{n1}(t)\}t - A(\theta_0)^{1/2}T_{\mathrm{obs}}\|^2\right],$$

where $P(\theta_0,t) = [\{Ds(\theta_0)+r_{n1}(t)\}^T\{Ds(\theta_0)+r_{n1}(t)\}]^{-1}\{Ds(\theta_0)+r_{n1}(t)\}^TA(\theta_0)^{1/2}$, both $P_5(t)$ belong to $\mathbb{P}_{5,t}$ and Δ is chosen such that $\gamma+\Delta\in(0,1)$ and $\alpha+\Delta\in(0,1)$ for γ

and α in Condition 7. Then since both ratios on the right hand side of the above inequality are $O_p(1)$ by Condition 7, by Lemma 7 and Lemma 9, (d) holds. Similarly by (20), for (e) we have

$$\begin{aligned}&\frac{\int_{\mathbb{R}^d} P_5(t,v)g_n^{**}\{t,v;r_{n1}(t),r_{n2}(t)\}\,dv}{q\{\Sigma^{-1/2}(t-c_\mu)\}}\\ \leq &P_5(t)\frac{\exp\left\{-\frac{m_1^2m_2^2\gamma}{2}\|t-\frac{1}{a_n\varepsilon_n}P(\theta_0,t)T_{\rm obs}\|^2\right\}}{q\{\Sigma^{-1/2}(t-c_\mu)\}}\\ &\times\exp\left[-\frac{(a_n^2\varepsilon_n^2\beta_1\beta_2-\gamma)m_2^2}{2}\|\{Ds(\theta_0)+r_{n1}(t)\}t-A(\theta_0)^{1/2}T_{\rm obs}\|^2\right]\\ &+P_5(t)\frac{\overline{K}^{\alpha}\left\{\lambda_{\rm min}^2(\Lambda)(1-\beta_1)m_1^2\|t-\frac{1}{a_n\varepsilon_n}P(\theta_0,t)T_{\rm obs}\|^2\right\}}{q\{\Sigma^{-1/2}(t-c_\mu)\}}\\ &\times\overline{K}^{1-\alpha}\left[\lambda_{\rm min}^2(\Lambda)(1-\beta_1)\|\{Ds(\theta_0)+r_{n1}(t)\}t-A(\theta_0)^{1/2}T_{\rm obs}\|^2\right],\end{aligned}$$

where both $P_5(t)$ belong to $\mathbb{P}_{5,t}$. Thus by Condition 7, Lemma 7 and Lemma 9, (e) holds. Therefore $\widetilde{\pi}_{B_\delta,\rm IS}\{P_2(t)\}/\widetilde{\pi}_{B_\delta}(1)=O_p(a_{n,\varepsilon}^{-p})$.

To show that $\pi_{B_\delta,\rm IS}\{P_2(t)\}/\pi_{B_\delta}(1)=O_p(a_{n,\varepsilon}^{-p})$, similar to the discussion of (13), it is sufficient to show that

$$\frac{\pi_{B_\delta,\rm IS}\{P_2(t)\}-\widetilde{\pi}_{B_\delta,\rm IS}\{P_2(t)\}}{\widetilde{\pi}_{B_\delta}(1)}=O_p(\alpha_n^{-1}a_{n,\varepsilon}^{-p}). \tag{22}$$

With the transformation $t=t(\theta)$ we have $\pi_{B_\delta,\rm IS}\{P_2(t)\}-\widetilde{\pi}_{B_\delta,\rm IS}\{P_2(t)\}$ is equal to

$$\alpha_n^{-1}a_{n,\varepsilon}^{-2p}\int_{t(B_\delta)}\int P_2(t)\pi(\theta_0+a_{n,\varepsilon}^{-1}t)^2\frac{r_n(s_{\rm obs}+\varepsilon_n v\mid\theta_0+a_{n,\varepsilon}^{-1}t)K(v)}{|\Sigma|^{-1/2}q\{\Sigma^{-1/2}(t-c_\mu)\}}\,dvdt.$$

Then by following the arguments of the proof of Lemma 6, we have

$$\begin{aligned}&|\pi_{B_\delta,\rm IS}\{P_2(t)\}-\widetilde{\pi}_{B_\delta,\rm IS}\{P_2(t)\}|\leq\alpha_n^{-1}a_{n,\varepsilon}^{d-2p}\sup_{\theta\in B_\delta}|\pi(\theta)^2A(\theta)^{-1/2}|\\ &\times\int_{t(B_\delta)}\int|P_2(t)|\frac{(a_na_{n,\varepsilon}^{-1})^d r_{\rm max}\left[a_na_{n,\varepsilon}^{-1}M\left\{Ds(\theta_0+\epsilon_t)t-a_{n,\varepsilon}\varepsilon_n v-\frac{1}{a_na_{n,\varepsilon}^{-1}}A(\theta_0)^{1/2}T_{\rm obs}\right\}\right]K(v)}{q\{\Sigma^{-1/2}(t-c_\mu)\}}\,dvdt.\end{aligned}$$

The ratio above is similar to the ratio of $g_n\{t,v;r_1(t),r_2(t)\}/q\{\Sigma^{-1/2}(t-c_\mu)\}$ except that the normal density is replaced by $r_{\rm max}(\cdot)$. Then by Condition 7, previous arguments for proving (iv) and (v) can be followed. Hence $\pi_{B_\delta,\rm IS}\{P_2(t)\}-\widetilde{\pi}_{B_\delta,\rm IS}\{P_2(t)\}=O_p(\alpha_n a_{n,\varepsilon}^{d-2p})$ and (22) holds. Therefore $\Sigma_{{\rm IS},n}=O_p(a_{n,\varepsilon}^{-2})$. □

REFERENCES

CREEL, M. & KRISTENSEN, D. (2013). Indirect likelihood inference (revised). UFAE and IAE working papers, Unitat de Fonaments de l'Analisi Economica (UAB) and Institut d'Analisi Economica (CSIC).

LEHMANN, E. L. (2004). *Elements of large-sample theory*. Springer Science & Business Media.