APPENDIX A. CLAIM IN SECTION 2

Proof. First note that $H_n(t) = 1 - Q_\varepsilon(2\hat{\theta} - t \mid S_n = s_{obs})$ is a sample-dependent cumulative distribution function on the parameter space. By equation (2.1), we denote both sides as $G(t)$, i.e. $G(t) = \text{pr}^*\{\theta - \hat{\theta}_S \le t \mid S_n = s_{obs}\} = \text{pr}\{\hat{\theta}_S - \theta \le t \mid \theta = \theta_0\}$. Now we can write $H_n(\theta_0) = \text{pr}^*(2\hat{\theta}_{\text{S}} - \theta \le \theta_0 \mid S_n = s_{\text{obs}}) = \text{pr}^*(\theta - \hat{\theta}_{\text{S}} \ge \hat{\theta}_{\text{S}} - \theta_0 \mid S_n = s_{\text{obs}}) = 1 - G(\hat{\theta}_{\text{S}} - \theta_0)$, for $G(t) = \text{pr}(\hat{\theta}_{\text{S}} - \theta \le t \mid \theta = \theta_0)$. The last equality holds by equation (2.1). Since $G(\hat{\theta}_{\text{S}} - \theta_0) \mid \theta_0 \sim Unif(0,1)$ with respect to the sampling variability of $\hat{\theta}_{\text{S}}$, $H_n(\theta_0) = H_n(\theta_0, s_{\text{obs}}) \sim Unif(0,1)$. By definition, $H_n(\cdot)$ is a confidence distribution for θ. □

APPENDIX B. LEMMA 1

Proof. Following the notation established in the claim of Section 2, first note that

$$\begin{aligned}
&\left|\text{pr}\{\theta \in \Gamma_{1-\alpha}(S_n)|\theta = \theta_0\} - (1-\alpha)\right| \\
&= \left|\text{pr}\{W(\theta, S_n) \in A_{1-\alpha}|\theta = \theta_0\} - (1-\alpha)\right| \\
&\le \left|\text{pr}^*\{V(\theta, S_n) \in A_{1-\alpha}|S_n = s_{\text{obs}}\} - (1-\alpha)\right| \\
&\quad + \left|\text{pr}\{W(\theta, S_n) \in A_{1-\alpha}|\theta = \theta_0\}\right. \\
&\quad \left.- \text{pr}^*\{V(\theta, S_n) \in A_{1-\alpha}|S_n = s_{\text{obs}}\}\right|
\end{aligned}$$

and by the definition of $A_{1-\alpha}$ in (4), $|\ \text{pr}^*\{V(\theta, S_n) \in A_{1-\alpha} \mid S_n = s_{\text{obs}}\} - (1-\alpha)\ | = o(\delta')$, almost surely for a pre-selected precision number, $\delta' > 0$. Therefore, by Condition 1, we have $|\ \text{pr}\{\theta \in \Gamma_{1-\alpha}(S_n) \mid \theta = \theta_0\} - (1-\alpha)\ | = \delta$ where $\delta = \max\{\delta_\varepsilon, \delta'\}$. Furthermore, if Condition 1 holds almost surely, then $|\ \text{pr}\{\theta \in \Gamma_{1-\alpha}(S_n) \mid \theta = \theta_0\} - (1-\alpha)\ | = o(\delta)$, almost surely. □

APPENDIX C. THEOREM 1

Proof. If $T = T(\theta, S_n)$ is an approximate pivot for S_n then

$$\text{pr}\{T(\theta, S_n) \in A \mid \theta = \theta_0\} = \int_{t\in A} g(t)dt\,\{1 + o(\delta'')\}, \quad \text{(C.1)}$$

for any Borel set $A \subset \mathcal{S}$. Given θ and t, denote the solution of $t = T(\theta, s)$ by $s_{t,\theta}$. The density functions $g(t)$ and $f_n(s_{t,\theta}|\theta)$ are connected by a Jacobian matrix:

$$f_n(s_{t,\theta}|\theta)|T^{(1)}(\theta, s_{t,\theta})|^{-1} = g(t)\{1 + o(\delta'')\} \quad \text{(C.2)}$$

where $T^{(1)}(\theta, S_n) = (\partial/\partial S_n)T(\theta, S_n)$.

For $\theta' \sim Q_\varepsilon(\cdot \mid s_{obs})$ and corresponding summary S'_n, the joint density of (θ', S'_n) conditional on the observed data, is

$$(\theta', S'_n)|S_n = s_{\text{obs}} \propto r_n(\theta) f_n(S_n \mid \theta) K_\varepsilon(S_n - s_{\text{obs}}).$$

Let $T' = T(\theta', S'_n)$. With a variable transformation from (θ', S'_n) to (θ', T'), the joint density of (θ', T'), conditional on the observed data, is

$$\begin{aligned}
(\theta', T')|s_{\text{obs}} &\propto r_n(\theta)\left[f_n(s_{t,\theta} \mid \theta)|T^{(1)}(\theta, s_{t,\theta})|^{-1}\right] \times \\
&\quad K_\varepsilon(s_{t,\theta} - s_{\text{obs}}) \\
&= r_n(\theta)\left[g(t)\{1 + o(\delta'')\}\right] K_\varepsilon(s_{t,\theta} - s_{\text{obs}}),
\end{aligned}$$

where $s_{t,\theta}$ is the solution of $t = T(\theta, S_n)$ and the equivalence holds by (C.2). Integrating over the parameter space yields

$$\begin{aligned}
T'|S_n = s_{\text{obs}} &\propto \left[g(t)\{1 + o(\delta'')\}\right] \int_{\mathcal{P}} r_n(\theta) K_\varepsilon(s_{t,\theta} - s_{\text{obs}}) d\theta \\
&\propto g(t)\{1 + o(\delta'')\},
\end{aligned}$$

provided (2.5) holds.

Now, consider $W(\theta, S_n) = T(\theta, S_n)$ as a function of the random sample given some fixed, unknown value of θ, by (C.1)

$$\text{pr}\{W(\theta, S_n) \in A \mid \theta = \theta_0\} = \int_{t\in A} g(t)dt\{1 + o(\delta)\}.$$

If we consider $V(\theta, S_n) = T(\theta, S_n)$ and the joint density of (θ', S'_n) pairs then

$$\text{pr}^*\{V(\theta, S_n) \in A \mid S_n = s_{\text{obs}}\} = \int_{t\in A} g(t)dt\{1 + o(\delta'')\}$$

thus satisfying Condition 1. Furthermore, by Lemma 1, $\Gamma_{1-\alpha}(s_{\text{obs}})$ in equation (2.3) is a $(1-\alpha)100\%$ confidence region for θ. □

APPENDIX D. COROLLARY 1

Proof. By Theorem 1, it suffices to show that equation (2.4) is free of t in each case.

(a) Suppose $S_n \sim g_1(S_n - \mu)$. Then $T_1 = T_1(\mu, S_n) = S_n - \mu \sim g_1(t)$ is a pivot for S_n. For any (t, μ) pair $s_{t,\mu} = t + \mu$. With a change of variables $u = t + \mu - s_{obs}$ and with $r_n(\mu) \propto 1$ we have

$$\int_{\mathcal{P}} r_n(\mu) K_\varepsilon(s_{t,\mu} - s_{\text{obs}}) d\mu = \int_{-\infty}^{\infty} K_\varepsilon(u) du,$$

which is free of t.

(b) Suppose $S_n \sim (1/\sigma)g_2(S_n/\sigma)$. Then $T = T(\sigma, S_n) = S_n/\sigma \sim g_2(t)$ is a pivot. For any (t, σ) pair $s_{t,\sigma} = t\sigma$. With $r_n(\sigma) \propto 1/\sigma$ and with a change of variables $u = t\sigma - s_{obs}$, we have

$$\begin{aligned}
\int_{\mathcal{P}} r_n(\sigma) K_\varepsilon(s_{t,\sigma} - s_{\text{obs}}) d\sigma &= \int_0^\infty \frac{1}{\sigma} K_\varepsilon(t\sigma - s_{\text{obs}}) d\sigma \\
&= \int_0^\infty \frac{1}{(u + s_{obs})/t} K_\varepsilon(u) \frac{1}{t} du
\end{aligned}$$

which is free of t.

(c) Since we have already proven parts (a) and (b), part (c) follows provided we select $r_n(\theta) \propto 1/\sigma$ for $\theta = (\mu, \sigma)$.

Finally, to prove the last statement of the corollary first note that the function $H_1(S_n, x) = \int_{-\infty}^{x} g_1(S_n - u)du$ is a CD for μ when $S_n \sim g_1(S_n - \mu)$ because, for a given S,

$H_1(S, x)$ is a distribution function on the parameter space $(-\infty, \infty)$ and given $x = \mu_0$, $H_1(S, x) \sim U(0, 1)$. Similarly, the function $H_2(S^2, x) = 1 - \int_0^x g_2(S/u)du$ is a CD for σ^2 when $S_n \sim (1/\sigma)g_2(S_n/\sigma)$. □

APPENDIX E. REMARK ON DEGENERACY OF ACCEPTANCE RATE

A natural question is whether Theorem 2 holds for a larger ε_n. We claim that the answer is negative, using the following basic normal mean model as a counterexample.

Consider a univariate Gaussian model with mean θ and unit variance, and observations that are IID from the model with $\theta = \theta_0$. Let $r_n(\theta)$ be a normal density with mean μ_n and variance b_n^{-2}, where μ_n and b_n are constant sequences satisfying $b_n(\mu_n - \theta_0) = O(1)$ and $b_n = o(\sqrt{n})$ as $n \to \infty$, and let S_n be the sample mean. One can verify that r_n and S_n satisfy the conditions of Theorem 2. The Gaussian kernel with variance ε_n^2 is used for the acceptance/rejection. Then the density of a linear transformation of $\theta \sim Q_\varepsilon(\theta \mid s_{\mathrm{obs}})$ is Gaussian with a closed form

$$\sqrt{n}(\theta - \hat{\theta}_S) \mid s_{obs} \sim N(0, n\sigma_\varepsilon^2)$$

where $\sigma_\varepsilon^2 = \frac{b_n^{-2}\Delta_n}{1+\Delta_n}$ and $\Delta_n = b_n^2(n^{-1} + \varepsilon^2)$. Also,

$$\sqrt{n}(\hat{\theta}_S - \theta) \mid \theta_0 = \frac{1}{1+\Delta_n}\sqrt{n}(s_{obs} - \theta) + \frac{\sqrt{n}b_n^{-1}\Delta_n}{1+\Delta_n}b_n(\mu_n - \theta).$$

By algebra, the expectation of $\sqrt{n}(\hat{\theta}_S - \theta) \mid \theta_0$ is $o(1)$ only when $\varepsilon_n = o(b_n^{-1/2}n^{-1/4})$, and the variance is $n\sigma_\varepsilon^2 + o(1)$ only when $\varepsilon_n = o(n^{-1/2})$ or $\varepsilon_n^{-1} = o(b_n^2 n^{-1/2})$. Since $b_n = o(\sqrt{n})$, both $\varepsilon_n = o(b_n^{-1/2}n^{-1/4})$ and $\varepsilon_n^{-1} = o(b_n^2 n^{-1/2})$ can not hold simultaneously. Therefore Condition 1 is satisfied only if $\varepsilon_n = o(n^{-1/2})$.

APPENDIX F. THEOREMS 2 AND 3

The proof for Theorem 2 requires establishing Lemmas 1–5 which are given in the Supplementary Material. Theorem 3 is proved after establishing Lemmas 6–7 given in the Supplementary Material. The proofs of these technical lemmas are contained in the Supplementary Material. It is helpful to have a copy of both [2] and [1] (and their supplementary material) on hand as these proofs rely on results from these two publications.

REFERENCES

[1] Li, W. and Fearnhead, P. (2018). Convergence of regression-adjusted approximate Bayesian computation. *Biometrika* **105**(2) 301–318.

[2] Li, W. and Fearnhead, P. (2018). On the Asymptotic Efficiency of approximate Bayesian computation estimators. *Biometrika* **105**(2) 286–299.