# ASYMPTOTIC RELATIONSHIPS BETWEEN THE D-TEST AND LIKELIHOOD RATIO-TYPE TESTS FOR HOMOGENEITY

Richard Charnigo and Jiayang Sun

*University of Kentucky and Case Western Reserve University*

## Supplementary Material

### S. Regularity Conditions and Proofs

The following regularity conditions are assumed for model (2), in which there is no unknown structural parameter.

**Condition A1** (Support). The probability density function $f(x,\theta)$ has a *common support* in $x$ for all $\theta$ in the parameter space $\Theta$ (i.e., $A := \{x : f(x,\theta) > 0\}$ is independent of $\theta$). The parameter space $\Theta$ is a *compact* subset of the real line, and the true value $\theta_0$ (under $H_0$) is an *interior* point of $\Theta$.

**Condition A2** (Smoothness and integrability). The probability density function $f(x,\theta)$ has three partial derivatives with respect to $\theta$. The function $f(x,\theta)$ and its derivatives are jointly continuous in $x$ and $\theta$. Moreover, there exist functions $M_{\theta_0}(x)$ and $K(x)$ such that

$$\begin{aligned} \left|\frac{\partial^j}{\partial\theta^j}\log f(x,\theta)\right| &\le M_{\theta_0}(x) \quad \forall x \in A \text{ and } |\theta-\theta_0| < c(\theta_0) \quad \text{for } j = 0,1,2,3, \\ \left|\frac{\partial^j}{\partial\theta^j} f(x,\theta)\right| &\le K(x) \quad \forall x \in A \quad \text{for } j = 0,1,2,3 \end{aligned}$$

with $EM_{\theta_0}(X) < \infty$ and $K(x) \in L^1 \cap L^2$ (i.e., $\int |K(x)|^p \; dx < \infty$ for $p = 1,2$).

**Condition A3** (Identifiability). Let $G(\theta) := (1-\alpha)\; 1_{\{\theta\ge\theta_1\}} + \alpha\; 1_{\{\theta\ge\theta_2\}}$. The mixture density

$$\int f(x,\theta)\; dG(\theta) = (1-\alpha) f(x,\theta_1) + \alpha f(x,\theta_2)$$

is *identifiable* in the sense that $\int f(x,\theta)\; dG_1(\theta) = \int f(x,\theta)\; dG_2(\theta)$ for all $x$ implies $G_1 = G_2$. The function $f(x,\theta)$ and its first two derivatives are also *identifiable*: for any distinct $\theta_1, \theta_2 \in \Theta$,

$$\sum_{j=1}^{2} \left\{ a_j f(x,\theta_j) + b_j \frac{\partial}{\partial\theta} f(x,\theta_j) + c_j \frac{\partial^2}{\partial\theta^2} f(x,\theta_j) \right\} = 0 \quad \text{for all } x \tag{S.1}$$

implies that $a_j = b_j = c_j = 0$ for $j = 1, 2$. Chen and Chen (2001) note that (S.1) is satisfied for densities from a one-dimensional exponential family.

**Condition A4** (Tightness). The processes $n^{-1/2} \sum Y_i(\theta)$, $n^{-1/2} \sum Y_i'(\theta)$, $n^{-1/2} \sum Z_i(\theta)$, and $n^{-1/2} \sum Z_i'(\theta)$ are *tight*.

**Condition A5** (Uniform strong law condition of large numbers). There exist a function $g$, for which $E[g(X_i)] < \infty$, and a number $\psi > 0$ such that $|Y_i(\theta)|^{(4+\psi)} \leq g(X_i)$ and $|Y_i'(\theta)|^3 \leq g(X_i)$ for all $\theta \in \Theta$.

We note that the exponential family structure $f(x, \theta) = a(x) \exp[-b(\theta) + t(x)\theta]$ does not by itself imply all of the above regularity conditions. For instance, if $f(x, \theta) := 1_{\{x>0\}} \theta \exp[-x\theta]$, then Condition A5 is not satisfied unless the infimum of $\Theta$ exceeds $(3/4)\theta_0$.

**Proof for Lemma 3.1**: (i) Compute. (ii) Expand $Z_i(\theta)$ about $\theta = \theta_0$ and pass to the limit as $\theta \to \theta_0$. (iii) and (iv) Repeated differentiation of the equation $\exp[b(\theta)] = \int a(x) \exp[t(x)\theta] \, dx$ gives formulas for the expected values of powers of $t(X_i)$, from which the desired results follow.

**Outline of the Proof for Theorem 3.1:** For ease of exposition, we provide a four-step outline of the proof for Theorem 3.1 before presenting details of the proof.

**Step 1** Given arbitrary $\epsilon > 0$, we may divide the $\theta_1\theta_2$-plane into three disjoint regions: $I_1 := \{|\theta_2 - \theta_0| \geq \epsilon\}$; $I_2 := \{|\theta_1 - \theta_0| < \epsilon, \ |\theta_2 - \theta_0| < \epsilon\}$; and, $I_3 := \{|\theta_1 - \theta_0| \geq \epsilon, \ |\theta_2 - \theta_0| < \epsilon\}$. By Lemma 1 of Chen and Chen (2001),

$$P\left((\hat{\theta}_1, \hat{\theta}_2) \in I_3\right) \leq P\left(|\hat{\theta}_1 - \theta_0| \geq \epsilon\right) \to 0$$

under $H_0$. Thus, we may examine maximum likelihood estimation of the mixture parameters over regions $I_1$ and $I_2$.

**Step 2** We maximize the likelihood over regions $I_1$ and $I_2$, employing Taylor expansions of $(1 - \hat{\alpha}) f(x, \hat{\theta}_1) + \hat{\alpha} \ f(x, \hat{\theta}_2) - f(x, \hat{\theta}_0)$ and some lemmas to find that

$$n \ d_n = \begin{cases} C^*(\hat{\theta}_2)\omega_{\hat{\theta}_2} + o_p(1) & \text{in } I_1 \\ C^*(\theta_0)\omega_{\theta_0} + \sqrt{\epsilon} \ O_p(1) & \text{in } I_2 \end{cases}, \quad \text{where} \ \ \omega_\theta := \frac{\left[(\sum W_i(\theta))^+\right]^2}{\sum [W_i(\theta)]^2}.$$

**Step 3** Relating $R_n$ to $\omega_\theta$ in regions $I_1$ and $I_2$ via Lemmas 3 and 4 of Chen and Chen (2001), we deduce that

$$n \ d_n = C^*(\hat{\theta}_2) R_n + \begin{cases} o_p(1) & \text{in } I_1 \\ \sqrt{\epsilon} \ O_p(1) + \{C^*(\theta_0) - C^*(\hat{\theta}_2)\} O_p(1) & \text{in } I_2 \end{cases}. \tag{S.2}$$

**Step 4** Let $\delta > 0$ and $\eta > 0$ be given. Choose $\epsilon > 0$ small enough and $n$ large enough so that the remainder terms in (S.2) are negligible with high probability:

$$P\left(|o_p(1)| > \eta\right) \le \delta/4 \qquad \text{in } I_1$$

and, for all $\theta$ such that $|\theta - \theta_0| < \epsilon$,

$$P\left(|\sqrt{\epsilon}\, O_p(1) + \{C^*(\theta_0) - C^*(\theta)\}O_p(1)| > \eta\right) \le \delta/4 \qquad \text{in } I_2.$$

For sufficiently large $n$, we also have $P\left(J(\bar{t}) \notin \Theta\right) \le \delta/4$ and $P\left(|\hat{\theta}_1 - \theta_0| \ge \epsilon\right) \le \delta/4$, where $J(\theta)$ is the inverse function of $b\,'(\theta)$ and $J(\bar{t}) \in \Theta$ is used in characterizing $\hat{\theta}_0$. Therefore, with probability at least $1 - \delta$,

$$C^*(\hat{\theta}_2)R_n - \eta \le n\ d_n \le C^*(\hat{\theta}_2)R_n + \eta$$

for sufficiently large $n$.

**Details of the Proof for Theorem 3.1:**

We begin with a lemma that characterizes $\hat{\theta}_0$. The lemma's provision that $J(\bar{t}) \in \Theta$ can be removed if one is willing to interpret the first equality as holding with probability approaching 1.

**Lemma S.1.** *Suppose that $f(x, \theta)$ is regular and that the null hypothesis is true. Then*

$$\hat{\theta}_0 = J\left(\bar{t}\right) = \theta_0 + \left(b\,''(\theta_0)\right)^{-1}\left(\bar{t} - b\,'(\theta_0)\right) + o_p(n^{-1/2}) = \theta_0 + O_p(n^{-1/2}),$$

*provided that $J(\bar{t}) \in \Theta$, where $\bar{t} := \sum t(X_i)/n$ and $J(\theta)$ is the inverse function of $b\,'(\theta)$.*

**Proof:** Differentiating $l_n(1/2, \theta, \theta)$ in $\theta$ and setting the result to zero yields $b\,'(\theta) = \bar{t}$. Applying $J$ to both sides of the equality gives $\theta = J(\bar{t})$. A second-order Taylor expansion of $J(\bar{t})$ about $\bar{t} = b\,'(\theta_0)$ yields $J(\bar{t}) = \theta_0 + \left(b\,''(\theta_0)\right)^{-1}\left(\bar{t} - b\,'(\theta_0)\right) + o_p(n^{-1/2})$.

Having characterized $\hat{\theta}_0$, we now proceed with analyses over the regions $I_1$ and $I_2$ identified in the four-step outline above. These analyses culminate in Propositions A.1 and A.2, which yield the crucial relation (S.2) in the four-step outline.

**Analysis over $I_1 : |\theta_2 - \theta_0| \ge \epsilon$**

Define

$$\hat{m}_1 := (1 - \hat{\alpha})(\hat{\theta}_1 - \theta_0) + \hat{\alpha}(\hat{\theta}_2 - \theta_0), \tag{S.3}$$

$$\hat{m}_2 := \hat{\alpha}(\hat{\theta}_2 - \theta_0)^2,$$

$$\hat{m}_1^* := \hat{m}_1 + \hat{m}_2 h(\hat{\theta}_2),$$

$$U_1(m_1^*) := 2m_1^* \sum Y_i(\theta_0) - (m_1^*)^2 \sum (Y_i(\theta_0)^2), \tag{S.4}$$

and

$$U_2(m_2) := 2m_2 \sum W_i(\hat{\theta}_2) - m_2^2 \sum (W_i(\hat{\theta}_2)^2).$$

We state and prove three lemmas relevant to the analysis over $I_1$.

**Lemma S.2.** *Suppose that $f(x,\theta)$ is regular and that the null hypothesis is true. If we restrict maximum likelihood estimation of $\theta_1$, $\theta_2$, and $\alpha$ to $|\theta_2 - \theta_0| \geq \epsilon$, then*

$$\hat{m}_1^* = \frac{\sum Y_i(\theta_0)}{\sum (Y_i(\theta_0)^2)} + o_p(n^{-1/2}) \quad \textit{and} \quad \hat{m}_2 = \frac{\left(\sum W_i(\hat{\theta}_2)\right)^+}{\sum (W_i(\hat{\theta}_2)^2)} + o_p(n^{-1/2}).$$

**Proof**: We can establish the following string of inequalities:

$$\begin{aligned}
2l_n(\hat{\alpha},\hat{\theta}_1,\hat{\theta}_2) - 2l_n(1/2,\theta_0,\theta_0) &\leq U_1(\hat{m}_1^*) + U_2(\hat{m}_2) + o_p(1) \\
&\leq U_1\left(\frac{\sum Y_i(\theta_0)}{\sum (Y_i(\theta_0)^2)}\right) + U_2(\hat{m}_2) + o_p(1) \\
&\leq U_1\left(\frac{\sum Y_i(\theta_0)}{\sum (Y_i(\theta_0)^2)}\right) + U_2\left(\frac{\left(\sum W_i(\hat{\theta}_2)\right)^+}{\sum (W_i(\hat{\theta}_2)^2)}\right) + o_p(1) \\
&\leq 2l_n(\hat{\alpha},\hat{\theta}_1,\hat{\theta}_2) - 2l_n(1/2,\theta_0,\theta_0) + o_p(1).
\end{aligned}$$

The first and last inequalities follow from a Taylor expansion argument in Chen and Chen (2001) for which the restriction $|\theta_2 - \theta_0| \geq \epsilon$ is assumed. By comparing the first and last lines in the string of inequalities, we see that the difference between any two successive lines must be $o_p(1)$. In particular, $U_1(\hat{m}_1^*) - U_1\left(\sum Y_i(\theta_0)/\sum (Y_i(\theta_0)^2)\right) = o_p(1)$. The left side of the preceding is seen to equal $-\left(\hat{m}_1^* - [\sum Y_i(\theta_0)/\sum (Y_i(\theta_0)^2)]\right)^2 \sum (Y_i(\theta_0)^2)$, implying that $\hat{m}_1^* - [\sum Y_i(\theta_0)/\sum (Y_i(\theta_0)^2)] = o_p(n^{-1/2})$. This proves the first half of the lemma.

As for the second half of the lemma, $U_2(\hat{m}_2) - U_2\left(\left(\sum W_i(\hat{\theta}_2)\right)^+ / \sum (W_i(\hat{\theta}_2)^2)\right) = o_p(1)$. The left side is

$$\hat{m}_2 \left[2 \sum \left(W_i(\hat{\theta}_2)\right) \; 1_{\sum W_i(\hat{\theta}_2) < 0}\right] - \left(\hat{m}_2 - \frac{\left(\sum W_i(\hat{\theta}_2)\right)^+}{\sum (W_i(\hat{\theta}_2)^2)}\right)^2 \sum (W_i(\hat{\theta}_2)^2).$$

Since $\hat{m}_2 \left[2 \sum \left(W_i(\hat{\theta}_2)\right) \; 1_{\sum W_i(\hat{\theta}_2) < 0}\right]$ and $-\left(\hat{m}_2 - \left(\sum W_i(\hat{\theta}_2)\right)^+ / \sum (W_i(\hat{\theta}_2)^2)\right)^2 \sum (W_i(\hat{\theta}_2)^2)$ are both nonpositive, they are $o_p(1)$, from which the second half of the lemma follows.

**Lemma S.3.** *Under the same conditions as Lemma S.2,*

$$\hat{\theta}_1 - \theta_0 = O_p(n^{-1/2}).$$

**Proof**: This is a direct consequence of Lemma S.2.

**Lemma S.4.** *Under the same conditions as Lemma S.2,*

$$n\ d_n = C^*(\hat{\theta}_2)\frac{\left(\left(\sum W_i(\hat{\theta}_2)\right)^+\right)^2}{\sum(W_i(\hat{\theta}_2)^2)} + o_p(1).$$

**Proof**: By Taylor expansion, Lemma S.1, and Lemma S.3,

$$\begin{aligned} & (1-\hat{\alpha})f(x,\hat{\theta}_1) + \hat{\alpha}\ f(x,\hat{\theta}_2) - f(x,\hat{\theta}_0) \\ = \ & (\theta_0 - \hat{\theta}_0)\ g_1(x,\theta_0) + (1-\hat{\alpha})(\hat{\theta}_1 - \theta_0)\ g_1(x,\theta_0) + \hat{\alpha}(\hat{\theta}_2 - \theta_0)\ g_1(x,\theta_0) \\ & + \hat{\alpha}(\hat{\theta}_2 - \theta_0)^2\ g_2\left(x, \tilde{\theta}(x,\hat{\theta}_2)\right)/2 + o_p(n^{-1/2})F_1, \end{aligned}$$

where $F_1$ is a remainder dominated by a function in $L^2$ and $\tilde{\theta}(x,\theta)$ is defined by

$$f(x,\theta) = f(x,\theta_0) + (\theta-\theta_0)\ g_1(x,\theta_0) + \frac{(\theta-\theta_0)^2}{2}\ g_2\left(x,\tilde{\theta}(x,\theta)\right).$$

By Lemmas 3.1, S.1, and S.2,

$$\begin{aligned} (\theta_0 - \hat{\theta}_0) + \hat{m}_1 &= -\frac{\sum Y_i(\theta_0)}{nE(Y_i(\theta_0)^2)} + \frac{\sum Y_i(\theta_0)}{\sum(Y_i(\theta_0)^2)} - \hat{m}_2 h(\hat{\theta}_2) + o_p(n^{-1/2}) \quad \text{(S.5)} \\ &= -\hat{m}_2 h(\hat{\theta}_2) + o_p(n^{-1/2}) = -h(\hat{\theta}_2)\frac{\left(\sum W_i(\hat{\theta}_2)\right)^+}{\sum(W_i(\hat{\theta}_2)^2)} + o_p(n^{-1/2}). \end{aligned}$$

Applying Lemma S.2 one more time, we see that

$$\begin{aligned} & (1-\hat{\alpha})f(x,\hat{\theta}_1) + \hat{\alpha}\ f(x,\hat{\theta}_2) - f(x,\hat{\theta}_0) \\ = \ & \frac{\left(\sum W_i(\hat{\theta}_2)\right)^+}{\sum(W_i(\hat{\theta}_2)^2)}\left[-h(\hat{\theta}_2)g_1(x,\theta_0) + g_2\left(x,\tilde{\theta}(x,\hat{\theta}_2)\right)/2\right] + o_p(n^{-1/2})F_2, \quad \text{(S.6)} \end{aligned}$$

from which the desired result follows.

The exponential family membership of $f(x,\theta)$ allows a more precise characterization of the difference $(\theta_0 - \hat{\theta}_0)$ in (S.5) than merely saying that the difference is $O_p(n^{-1/2})$. This makes possible the conclusion in (S.6) that the remainder $F_2$ is accompanied by an $o_p(n^{-1/2})$ quantity rather than by an $O_p(n^{-1/2})$ quantity.

Finally, we obtain the following proposition.

**Proposition S.1.** *Under the same conditions as Lemma S.2,*

$$n\ d_n = C^*(\hat{\theta}_2)R_n + o_p(1).$$

**Proof**: Lemma 3 of Chen and Chen (2001) shows that $R_n = \frac{\left((\sum W_i(\hat{\theta}_2))^+\right)^2}{\sum(W_i(\hat{\theta}_2)^2)} + o_p(1)$. The desired result is therefore a consequence of Lemma S.4.

**Analysis over $I_2 : |\theta_1 - \theta_0| < \epsilon$, $|\theta_2 - \theta_0| < \epsilon$**

Redefine

$$\hat{m}_2 := \ (1-\hat{\alpha})(\hat{\theta}_1 - \theta_0)^2 + \hat{\alpha}(\hat{\theta}_2 - \theta_0)^2, \tag{S.7}$$

$$\hat{m}_1^* := \ \hat{m}_1 + h(\theta_0)\hat{m}_2, \tag{S.8}$$

and

$$U_2(m_2) := \ 2m_2 \sum W_i(\theta_0) - m_2^2 \sum (W_i(\theta_0)^2). \tag{S.9}$$

We state and prove four lemmas relevant to the analysis over $I_2$.

**Lemma S.5.** *Suppose that $f(x,\theta)$ is regular and that the null hypothesis is true. If we restrict maximum likelihood estimation of $\theta_1$, $\theta_2$, and $\alpha$ to $\max\{|\theta_1 - \theta_0|, |\theta_2 - \theta_0|\} < \epsilon$, then*

$$\hat{m}_1^* = \frac{\sum Y_i(\theta_0)}{\sum(Y_i(\theta_0)^2)} + \sqrt{\epsilon}\ O_p(n^{-1/2}) \quad \textit{and} \quad \hat{m}_2 = \frac{(\sum W_i(\theta_0))^+}{\sum(W_i(\theta_0)^2)} + \sqrt{\epsilon}\ O_p(n^{-1/2}).$$

**Proof**: A string of inequalities as in the Proof of Lemma S.2 shows that

$$U_1(\hat{m}_1^*) - U_1\left(\frac{\sum Y_i(\theta_0)}{\sum(Y_i(\theta_0)^2)}\right) = \epsilon\ O_p(1) \quad \text{and} \quad U_2(\hat{m}_2) - U_2\left(\frac{(\sum W_i(\theta_0))^+}{\sum(W_i(\theta_0)^2)}\right) = \epsilon\ O_p(1).$$

We have $\epsilon\ O_p(1)$ above instead of $o_p(1)$ due to the nature of the underlying Taylor expansions. The lemma then follows by arguments similar to those in the last two paragraphs of the Proof of Lemma S.2.

We note that, since $\hat{m}_1$ and $\hat{m}_2$ are themselves $O_p(n^{-1/2})$, the relevance of Lemma S.5 is that it holds for arbitrary $\epsilon > 0$; a suitable $\epsilon$ is prescribed in the four-step outline above.

**Lemma S.6.** *Under the same conditions as Lemma S.5,*

$$\hat{\theta}_1 - \theta_0 = O_p(n^{-1/4}) \quad \text{and} \quad \hat{\alpha}(\hat{\theta}_2 - \theta_0)^2 = O_p(n^{-1/2}).$$

**Proof**: These are direct consequences of Lemma S.5.

**Lemma S.7.** *Under the same conditions as Lemma S.5,*

$$n\ d_n = C^*(\theta_0)\frac{\left((\sum W_i(\theta_0))^+\right)^2}{\sum(W_i(\theta_0)^2)} + \sqrt{\epsilon}\ O_p(1).$$

**Proof**: By Taylor expansion, Lemma S.1, and Lemma S.6,

$$\begin{aligned}
& (1-\hat{\alpha})f(x,\hat{\theta}_1)+\hat{\alpha}\ f(x,\hat{\theta}_2)-f(x,\hat{\theta}_0) \\
= \ & (\theta_0-\hat{\theta}_0)g_1(x,\theta_0)+(1-\hat{\alpha})(\hat{\theta}_1-\theta_0)g_1(x,\theta_0)+\hat{\alpha}(\hat{\theta}_2-\theta_0)g_1(x,\theta_0) \\
& +(1-\hat{\alpha})(\hat{\theta}_1-\theta_0)^2g_2(x,\theta_0)/2+\hat{\alpha}(\hat{\theta}_2-\theta_0)^2g_2(x,\theta_0)/2+\hat{\alpha}(\hat{\theta}_2-\theta_0)^3g_3\left(x,\tilde{\theta}(x,\hat{\theta}_2)\right)/6 \\
& +o_p(n^{-1/2})F_1,
\end{aligned}$$

where $F_1$ is a remainder dominated by a function in $L^2$. Using Lemma S.6 and the fact that $|\hat{\theta}_2-\theta_0|<\epsilon$, we find that $\hat{\alpha}(\hat{\theta}_2-\theta_0)^3g_3\left(x,\tilde{\theta}(x,\hat{\theta}_2)\right)/6=\epsilon\ O_p(n^{-1/2})F_2$. From Lemmas 3.1, S.1, and S.5,

$$\begin{aligned}
(\theta_0-\hat{\theta}_0)+\hat{m}_1 &= -\frac{\sum Y_i(\theta_0)}{nE(Y_i(\theta_0)^2)}+\frac{\sum Y_i(\theta_0)}{\sum(Y_i(\theta_0)^2)}-\hat{m}_2h(\theta_0)+\sqrt{\epsilon}\ O_p(n^{-1/2}) \\
&= -\hat{m}_2h(\theta_0)+\sqrt{\epsilon}\ O_p(n^{-1/2})=-h(\theta_0)\frac{(\sum W_i(\theta_0))^+}{\sum(W_i(\theta_0)^2)}+\sqrt{\epsilon}\ O_p(n^{-1/2}).
\end{aligned}$$

So, applying Lemma S.5 once more,

$$\begin{aligned}
& (1-\hat{\alpha})f(x,\hat{\theta}_1)+\hat{\alpha}\ f(x,\hat{\theta}_2)-f(x,\hat{\theta}_0) \\
= \ & \frac{(\sum W_i(\theta_0))^+}{\sum(W_i(\theta_0)^2)}\left[-h(\theta_0)g_1(x,\theta_0)+g_2(x,\theta_0)/2\right]+\sqrt{\epsilon}\ O_p(n^{-1/2})F_3,
\end{aligned}$$

which completes the proof.

**Lemma S.8.** *Under the same conditions as Lemma S.5,*

$$n\ d_n=C^*(\theta_0)R_n+\sqrt{\epsilon}\ O_p(1).$$

**Proof**: Lemma 4 of Chen and Chen (2001) shows that $R_n=\frac{\left((\sum W_i(\theta_0))^+\right)^2}{\sum(W_i(\theta_0)^2)}+\epsilon\ O_p(1)$. The desired result therefore follows from Lemma S.7.

Finally, we obtain the following proposition.

**Proposition S.2.** *Under the same conditions as Lemma S.5,*

$$n\ d_n=C^*(\hat{\theta}_2)R_n+\sqrt{\epsilon}\ O_p(1)+\{C^*(\theta_0)-C^*(\hat{\theta}_2)\}O_p(1).$$

**Proof**: Noting that $R_n=O_p(1)$, we may apply Lemma S.8.

**Outline of the Proof for Theorem 4.1:** For ease of exposition, we provide a three-step outline of the proof for Theorem 4.1 before presenting details of the proof.

**Step 1** A Taylor expansion of $(1-\hat{\alpha})f(x,\hat{\theta}_1)+\hat{\alpha}\ f(x,\hat{\theta}_2)-f(x,\hat{\theta}_0)$ and lemmas show that

$$n\ d_n = C^*(\theta_0)\omega_{\theta_0} + o_p(1).$$

**Step 2** Relating $M_n$ to $\omega_\theta$ via developments of Chen, Chen, and Kalbfleisch (2001), we have

$$n\ d_n = C^*(\theta_0)M_n + o_p(1). \tag{S.10}$$

**Step 3** Let $\delta > 0$ and $\eta > 0$ be given. Choose $n$ large enough so that the remainder term in (S.10) is negligible with high probability:

$$P\left(|o_p(1)| > \eta\right) \le \delta/2.$$

For sufficiently large $n$, we also have $P\left(J(\bar{t}) \notin \Theta\right) \le \delta/2$. Therefore, with probability at least $1-\delta$,

$$C^*(\theta_0)M_n - \eta \le n\ d_n \le C^*(\theta_0)M_n + \eta$$

for sufficiently large $n$.

**Details of the Proof for Theorem 4.1:**

Let $\hat{m}_1$, $\hat{m}_2$, and $\hat{m}_1^*$ be as in (S.3), (S.7), and (S.8) but without restrictions on the values of $\theta_1$ and $\theta_2$ (except that they belong to $\Theta$). Also, let $U_1(m_1^*)$ and $U_2(m_2)$ be as in (S.4) and (S.9). We now state and prove two lemmas.

**Lemma S.9.** *Suppose that $f(x,\theta)$ is regular and that the null hypothesis is true. Under the indicated Bayesian estimation framework,*

$$\hat{m}_1^* = \frac{\sum Y_i(\theta_0)}{\sum\left(Y_i(\theta_0)^2\right)} + o_p(n^{-1/2}) \quad \textit{and} \quad \hat{m}_2 = \frac{\left(\sum W_i(\theta_0)\right)^+}{\sum\left(W_i(\theta_0)^2\right)} + o_p(n^{-1/2}).$$

**Proof**: In this Bayesian framework, $U_1(\hat{m}_1^*) - U_1\left(\sum Y_i(\theta_0)/\sum(Y_i(\theta_0)^2)\right) = o_p(1)$ without restrictions on $\theta_1$ and $\theta_2$ (and similarly for $U_2$). This is related to the fact that $\hat{\theta}_2 - \theta_0 = o_p(1)$ in the Bayesian framework, which was not the case in the maximum likelihood framework. The rest of the proof is similar to that of Lemma S.2.

**Lemma S.10.** *Under the same conditions as Lemma S.9,*

$$\hat{\theta}_1 - \theta_0 = O_p(n^{-1/4}) \quad \text{and} \quad \hat{\theta}_2 - \theta_0 = O_p(n^{-1/4}).$$

**Proof**: Chen, Chen, and Kalbfleisch (2001) show that $\log\hat{\alpha} = O_p(1)$, so that $\hat{\alpha}$ is bounded away from zero in probability. The results are now consequences of Lemma S.9.

Lemmas S.9 and S.10 lead to the following proposition.

**Proposition S.3.** *Under the same conditions as Lemma S.9,*

$$n\ d_n = C^*(\theta_0)\frac{\left((\sum W_i(\theta_0))^+\right)^2}{\sum(W_i(\theta_0)^2)} + o_p(1).$$

**Proof**: By Taylor expansion and Lemma S.10,

$$\begin{aligned} & (1-\hat{\alpha})f(x,\hat{\theta}_1) + \hat{\alpha}\ f(x,\hat{\theta}_2) - f(x,\hat{\theta}_0) \\ = \ & (\theta_0 - \hat{\theta}_0)g_1(x,\theta_0) + (1-\hat{\alpha})(\hat{\theta}_1 - \theta_0)g_1(x,\theta_0) + \hat{\alpha}(\hat{\theta}_2 - \theta_0)g_1(x,\theta_0) \\ & + (1-\hat{\alpha})(\hat{\theta}_1 - \theta_0)^2 g_2(x,\theta_0)/2 + \hat{\alpha}(\hat{\theta}_2 - \theta_0)^2 g_2(x,\theta_0)/2 + o_p(n^{-1/2})F_1, \end{aligned}$$

where $F_1$ is a remainder dominated by a function in $L^2$. By Lemma S.9, this is

$$\frac{(\sum W_i(\theta_0))^+}{\sum(W_i(\theta_0)^2)}\left[-h(\theta_0)g_1(x,\theta_0) + g_2(x,\theta_0)/2\right] + o_p(n^{-1/2})F_2,$$

from which the desired result follows.

Chen, Chen, and Kalbfleisch (2001) establish that

$$M_n = \frac{((\sum W_i(\theta_0))^+)^2}{\sum(W_i(\theta_0)^2)} + o_p(1),$$

whose limiting distribution is $0.5\chi_0^2 + 0.5\chi_1^2$. Theorem 4.1 is then a consequence of Proposition S.3.

**Outline of the Proof for Theorem 5.1:** For ease of exposition, we provide a three-step outline of the proof for Theorem 5.1 before presenting details of the proof.

**Step 1** Proposition 3 of Charnigo and Sun (2004) indicates that the distribution of $\sigma_0 d_n$ is invariant to $\theta_0$ and $\sigma_0$, so we may as well assume that $\theta_0 = 0$ and $\sigma_0 = 1$. If $\alpha_0 = 0.5$, then by a Taylor expansion of $(1-\hat{\alpha})f_{\hat{\sigma}}(x,\hat{\theta}_1) + \hat{\alpha}\ f_{\hat{\sigma}}(x,\hat{\theta}_2) - f_{\hat{\sigma}_0}(x,\hat{\theta}_0)$ we find that

$$n\ d_n = \frac{35}{256\pi^{1/2}}\frac{((\sum_{i=1}^n V_i)^-)^2}{\sum_{i=1}^n V_i^2} + o_p(1), \tag{S.11}$$

where $V_i := (X_i^4 - 6X_i^2 + 3)/24$ and $x^- := \max\{-x, 0\}$.

**Step 2** If $\alpha_0 < 0.5$, then the Taylor expansion of $(1-\hat{\alpha})f_{\hat{\sigma}}(x,\hat{\theta}_1) + \hat{\alpha}\ f_{\hat{\sigma}}(x,\hat{\theta}_2) - f_{\hat{\sigma}_0}(x,\hat{\theta}_0)$ works out differently and we get

$$n\ d_n = \frac{5}{32\pi^{1/2}}\frac{(\sum_{i=1}^n U_i)^2}{\sum_{i=1}^n U_i^2} + o_p(1), \tag{S.12}$$

where $U_i := (X_i^3 - 3X_i)/6$.

**Step 3** Chen and Li (2008) show that

$$E_n(0.5) = \frac{((\sum_{i=1}^n V_i)^-)^2}{\sum_{i=1}^n V_i^2} + o_p(1) \tag{S.13}$$

and, for $\alpha_0 < 0.5$,

$$E_n(\alpha_0) = \frac{(\sum_{i=1}^n U_i)^2}{\sum_{i=1}^n U_i^2} + 2\log[2\alpha_0] + o_p(1). \tag{S.14}$$

Hence, with $A^*(\alpha)$ as defined in (10) we have

$$n\ d_n = A^*(\alpha_0)\{E_n(\alpha_0) - 2\log[2\alpha_0]\} + o_p(1). \tag{S.15}$$

**Details of the Proof for Theorem 5.1:**

Besides defining $V_i$ and $U_i$ as in the three-step outline above, we put $Z_i := (X_i^2 - 1)/2$,

$$m_j := (1-\alpha)\theta_1^j + \alpha\theta_2^j \ \text{ for } 1 \le j \le 4,$$

$$s_1 := m_1, \quad s_2 := m_2 + \sigma^2 - 1, \quad s_3 := m_3, \quad \text{and} \ \ s_4 := m_4 - 3m_2^2$$

as in Chen and Li (2008). For convenience we set

$$f_{j_1,j_2}(x) := \frac{\partial^{j_1+j_2}}{\partial\theta^{j_1}\partial(\sigma^2)^{j_2}} f_\sigma(x,\theta)|_{\theta=0,\sigma=1} \ \text{ for } 0 \le j_1, j_2 \le 5,$$

$$U_1^\dagger(s_1) := 2s_1 \sum X_i - s_1^2 \sum X_i^2, \quad U_2^\dagger(s_2) := 2s_2 \sum Z_i - s_2^2 \sum Z_i^2,$$
$$U_3^\dagger(s_3) := 2s_3 \sum U_i - s_3^2 \sum U_i^2, \quad \text{and} \ \ U_4^\dagger(s_4) := 2s_4 \sum V_i - s_4^2 \sum V_i^2.$$

Next we present analyses with $\alpha_0 = 0.5$ and $\alpha_0 < 0.5$ respectively. These analyses culminate in Propositions S.4 and S.5, which yield (S.11) and (S.12) in steps 1 and 2 of the three-step outline. Propositions S.4 and S.5 then imply Theorem 5.1 since (S.13) and (S.14) follow from Theorem 2 of Chen and Li (2008), while (S.15) is immediate from (S.11)-(S.14).

**Analysis with $\alpha_0 = 0.5$**

We begin by stating and proving two lemmas.

**Lemma S.11.** *Suppose that model (3) applies and that the null hypothesis is true with $\theta_0 = 0$ and $\sigma_0 = 1$. Under the indicated empirical Bayesian estimation framework, with $\alpha_0 = 0.5$ we have*

$$\hat{s}_1 = \frac{\sum X_i}{\sum X_i^2} + o_p(n^{-1/2}), \quad \hat{s}_2 = \frac{\sum Z_i}{\sum Z_i^2} + o_p(n^{-1/2}), \quad \textit{and} \ \ \hat{s}_4 = \frac{-(\sum V_i)^-}{\sum V_i^2} + o_p(n^{-1/2}).$$

**Proof**: We can establish the following string of inequalities:

$$\begin{aligned}
2l_n^\dagger(\hat\alpha,\hat\theta_1,\hat\theta_2,\hat\sigma) - 2l_n^\dagger(1/2,0,0,1) &\le U_1^\dagger(\hat s_1) + U_2^\dagger(\hat s_2) + U_4^\dagger(\hat s_4) + o_p(1) \\
&\le U_1^\dagger\left(\frac{\sum X_i}{\sum X_i^2}\right) + U_2^\dagger(\hat s_2) + U_4^\dagger(\hat s_4) + o_p(1) \\
&\le U_1^\dagger\left(\frac{\sum X_i}{\sum X_i^2}\right) + U_2^\dagger\left(\frac{\sum Z_i}{\sum Z_i^2}\right) + U_4^\dagger(\hat s_4) + o_p(1) \\
&\le U_1^\dagger\left(\frac{\sum X_i}{\sum X_i^2}\right) + U_2^\dagger\left(\frac{\sum Z_i}{\sum Z_i^2}\right) + U_4^\dagger\left(\frac{-(\sum V_i)^-}{\sum V_i^2}\right) + o_p(1) \\
&\le 2l_n^\dagger(\hat\alpha,\hat\theta_1,\hat\theta_2,\hat\sigma) - 2l_n^\dagger(1/2,0,0,1) + o_p(1).
\end{aligned}$$

The first and last inequalities follow from a Taylor expansion argument in Chen and Li (2008). The difference between any two successive lines must be $o_p(1)$. That $U_1^\dagger(\hat s_1) - U_1^\dagger\left(\sum X_i/\sum X_i^2\right) = o_p(1)$ implies $\hat s_1 = \sum X_i/\sum X_i^2 + o_p(n^{-1/2})$ by an argument identical to that used to prove the first half of Lemma S.2. Likewise, that $U_2^\dagger(\hat s_2) - U_2^\dagger\left(\sum Z_i/\sum Z_i^2\right) = o_p(1)$ implies $\hat s_2 = \sum Z_i/\sum Z_i^2 + o_p(n^{-1/2})$. Finally, that $U_4^\dagger(\hat s_4) - U_4^\dagger\left(-(\sum V_i)^-/\sum V_i^2\right) = o_p(1)$ implies $\hat s_4 = -(\sum V_i)^-/\sum V_i^2 + o_p(n^{-1/2})$ by an argument identical to that used to prove the second half of Lemma S.2.

**Lemma S.12.** *Under the same conditions as Lemma S.11,*

$$(1-\hat\alpha)f_{\hat\sigma}(x,\hat\theta_1) + \hat\alpha\ f_{\hat\sigma}(x,\hat\theta_2) - f_{\hat\sigma_0}(x,\hat\theta_0) = \hat s_4\frac{(x^4-6x^2+3)\exp\left[-\frac{x^2}{2}\right]}{24\ (2\pi)^{1/2}} + o_p(n^{-1/2})F_1,$$

*where $F_1$ is a remainder dominated by a function in $L^2$.*

**Proof**: Theorem 1 of Chen and Li (2008) shows that $\hat\theta_1 = O_p(n^{-1/8})$, $\hat\theta_2 = O_p(n^{-1/8})$, $\hat\alpha - \alpha_0 = O_p(n^{-1/4})$, and $\hat\sigma^2 - 1 = O_p(n^{-1/4})$. Also, $\hat\theta_0 = O_p(n^{-1/2})$ and $\hat\sigma_0^2 - 1 = O_p(n^{-1/2})$. Consequently,

$$\begin{aligned}
(1-\hat\alpha)f_{\hat\sigma}(x,\hat\theta_1) + \hat\alpha\ f_{\hat\sigma}(x,\hat\theta_2) - f_{\hat\sigma_0}(x,\hat\theta_0) &= (1-\hat\alpha)\left\{\sum_{j_1+j_2=0}^{4}\frac{f_{j_1,j_2}(x)\hat\theta_1^{j_1}(\hat\sigma^2-1)^{j_2}}{j_1!\ j_2!} + R_1(x,\hat\theta_1,\hat\sigma)\right\} \\
&\quad +\hat\alpha\left\{\sum_{j_1+j_2=0}^{4}\frac{f_{j_1,j_2}(x)\hat\theta_2^{j_1}(\hat\sigma^2-1)^{j_2}}{j_1!\ j_2!} + R_2(x,\hat\theta_2,\hat\sigma)\right\} \\
&\quad -\left\{\sum_{j_1+j_2=0}^{1}\frac{f_{j_1,j_2}(x)\hat\theta_0^{j_1}(\hat\sigma_0^2-1)^{j_2}}{j_1!\ j_2!} + R_3(x,\hat\theta_0,\sigma_0)\right\},
\end{aligned}$$

where the remainder $R_1(x,\hat\theta_1,\hat\sigma)$ is dominated by

$$\sum_{j_1+j_2=5}\ \sup_{\{a,b:a^2+(b-1)^2\le\hat\theta_1^2+(\hat\sigma^2-1)^2\}}\frac{\left|\frac{\partial^{j_1+j_2}}{\partial\theta^{j_1}\partial(\sigma^2)^{j_2}}f_\sigma(x,\theta)|_{\theta=a,\sigma^2=b}\ \ \hat\theta_1^{j_1}(\hat\sigma^2-1)^{j_2}\right|}{j_1!\ j_2!}$$

and similarly for $R_2(x, \hat{\theta}_2, \hat{\sigma})$ and $R_3(x, \hat{\theta}_0, \hat{\sigma}_0)$. Since $\hat{\theta}_1$ converges in probability to 0 and $\hat{\sigma}^2$ converges in probability to 1, we can replace $\sup_{\{a,b: a^2+(b-1)^2 \le \hat{\theta}_1^2 + (\hat{\sigma}^2-1)^2\}} \left| \frac{\partial^{j_1+j_2}}{\partial\theta^{j_1}\partial(\sigma^2)^{j_2}} f_\sigma(x,\theta)|_{\theta=a,\sigma^2=b} \right|$ by $\sup_{\{a,b: a^2+(b-1)^2 \le c\}} \left| \frac{\partial^{j_1+j_2}}{\partial\theta^{j_1}\partial(\sigma^2)^{j_2}} f_\sigma(x,\theta)|_{\theta=a,\sigma^2=b} \right|$ for a positive constant $c < 1$. Also, because

$$\frac{\partial^{j_1+j_2}}{\partial\theta^{j_1}\partial(\sigma^2)^{j_2}} f_\sigma(x,\theta)|_{\theta=a,\sigma^2=b} = \frac{P_1(x,a,b)\exp\left[-\frac{(x-a)^2}{2b}\right]}{b^{P_2}}$$

for some polynomial $P_1(x,a,b)$ and some constant $P_2$, the supremum of its absolute value over $\{a, b : a^2 + (b-1)^2 \le c\}$ is dominated by a function in $L^2$. Hence, $R_1(x, \hat{\theta}_1, \hat{\sigma})$ is $o_p(n^{-1/2})$ times an $L^2$-dominated function. The same is true of $R_2(x, \hat{\theta}_2, \hat{\sigma})$ and $R_3(x, \hat{\theta}_0, \hat{\sigma}_0)$.

Moreover, all of the terms with $j_1+2j_2 > 4$ in $\sum_{j_1+j_2=0}^{4} f_{j_1,j_2}(x)\hat{\theta}_1^{j_1}(\hat{\sigma}^2-1)^{j_2}/(j_1!\ j_2!)$ are $o_p(n^{-1/2})$ times an $L^2$-dominated function, and similarly for $\sum_{j_1+j_2=0}^{4} f_{j_1,j_2}(x)\hat{\theta}_2^{j_1}(\hat{\sigma}^2-1)^{j_2}/(j_1!\ j_2!)$, so that $(1-\hat{\alpha})f_{\hat{\sigma}}(x,\hat{\theta}_1) + \hat{\alpha}\ f_{\hat{\sigma}}(x,\hat{\theta}_2) - f_{\hat{\sigma}_0}(x,\hat{\theta}_0)$ equals

$$\begin{aligned}
&\hat{m}_1 f_{1,0}(x) + \frac{\hat{m}_2}{2} f_{2,0}(x) + \frac{\hat{m}_3}{6} f_{3,0}(x) + (\hat{\sigma}^2-1) f_{0,1}(x) + (\hat{\sigma}^2-1)\hat{m}_1 f_{1,1}(x) \\
&+\frac{\hat{m}_4}{24} f_{4,0}(x) + (\hat{\sigma}^2-1)\frac{\hat{m}_2}{2} f_{2,1}(x) + \frac{(\hat{\sigma}^2-1)^2}{2} f_{0,2}(x) \qquad \text{(S.16)} \\
&-\left[\hat{\theta}_0 f_{1,0}(x) + (\hat{\sigma}_0^2-1) f_{0,1}(x)\right] + o_p(n^{-1/2}) F_2 \\
= &\left\{\frac{x\exp\left[-\frac{x^2}{2}\right]}{(2\pi)^{1/2}}\right\}[\hat{m}_1 - \hat{\theta}_0] + \left\{\frac{(x^2-1)\exp\left[-\frac{x^2}{2}\right]}{(2\pi)^{1/2}}\right\}\left[\frac{\hat{m}_2}{2} + \frac{(\hat{\sigma}^2-1)}{2} - \frac{(\hat{\sigma}_0^2-1)}{2}\right] \qquad \text{(S.17)} \\
&+\left\{\frac{(x^3-3x)\exp\left[-\frac{x^2}{2}\right]}{(2\pi)^{1/2}}\right\}\left[\frac{\hat{m}_3}{6} + \frac{\hat{m}_1(\hat{\sigma}^2-1)}{2}\right] \qquad \text{(S.18)} \\
&+\left\{\frac{(x^4-6x^2+3)\exp\left[-\frac{x^2}{2}\right]}{(2\pi)^{1/2}}\right\}\left[\frac{\hat{m}_4}{24} + \frac{(\hat{\sigma}^2-1)^2}{8} + \frac{\hat{m}_2(\hat{\sigma}^2-1)}{4}\right] \qquad \text{(S.19)} \\
&+o_p(n^{-1/2}) F_2, \qquad \text{(S.20)}
\end{aligned}$$

where $F_2$ is a remainder dominated by a function in $L^2$.

Line (S.17) can be absorbed into line (S.20) because, by Lemma S.11,

$$\hat{m}_1 - \hat{\theta}_0 = \hat{s}_1 - \hat{\theta}_0 = \frac{\sum X_i}{\sum X_i^2} + o_p(n^{-1/2}) - \frac{\sum X_i}{n} = o_p(n^{-1/2}) \quad \text{and}$$

$$\hat{m}_2 + (\hat{\sigma}^2-1) - (\hat{\sigma}_0^2-1) = \hat{s}_2 - (\hat{\sigma}_0^2-1) = \frac{\sum Z_i}{\sum Z_i^2} + o_p(n^{-1/2}) - \frac{\sum Z_i}{n/2} = o_p(n^{-1/2}).$$

Line (S.18) can also be absorbed into line (S.20): $\hat{m}_1(\hat{\sigma}^2-1)/2 = o_p(n^{-1/2})$ is clear from Theorem 1 of Chen and Li (2008) and Lemma S.11, while the following argument establishes that $\hat{m}_3/6 = o_p(n^{-1/2})$.

Since $\hat{m}_1 = [1/2 + O_p(n^{-1/4})]\ \hat{\theta}_1 + [1/2 + O_p(n^{-1/4})]\ \hat{\theta}_2 = O_p(n^{-1/2})$, we have $\hat{\theta}_1 + \hat{\theta}_2 = O_p(n^{-3/8})$. Also, $\hat{\theta}_1^2 - \hat{\theta}_1\hat{\theta}_2 + \hat{\theta}_2^2 = O_p(n^{-1/4})$, $\hat{\theta}_1^3 = O_p(n^{-3/8})$, and $\hat{\theta}_2^3 = O_p(n^{-3/8})$. Hence,

$$\hat{m}_3 = \left[\frac{1}{2} + O_p(n^{-1/4})\right]\ \hat{\theta}_1^3 + \left[\frac{1}{2} + O_p(n^{-1/4})\right]\hat{\theta}_2^3 = \frac{(\hat{\theta}_1 + \hat{\theta}_2)(\hat{\theta}_1^2 - \hat{\theta}_1\hat{\theta}_2 + \hat{\theta}_2^2)}{2} + o_p(n^{-1/2}) = o_p(n^{-1/2}).$$

All that remains is to show that

$$\hat{s}_4 = \hat{m}_4 + 3(\hat{\sigma}^2 - 1)^2 + 6\hat{m}_2(\hat{\sigma}^2 - 1) + o_p(n^{-1/2}). \tag{S.21}$$

Since $\hat{s}_2 = \hat{m}_2 + (\hat{\sigma}^2 - 1) = O_p(n^{-1/2})$ by Lemma S.11, we may substitute $1 - \hat{\sigma}^2 + O_p(n^{-1/2})$ for $\hat{m}_2$ in the definition of $\hat{s}_4$,

$$\hat{s}_4 = \hat{m}_4 - 3\hat{m}_2^2 = \hat{m}_4 - 3\{1 - \hat{\sigma}^2 + O_p(n^{-1/2})\}^2 = \hat{m}_4 + 3(\hat{\sigma}^2 - 1)^2 - 6(\hat{\sigma}^2 - 1) \times (\hat{\sigma}^2 - 1) + o_p(n^{-1/2}). \tag{S.22}$$

Substituting $O_p(n^{-1/2}) - \hat{m}_2$ for the last $(\hat{\sigma}^2 - 1)$ in (S.22) yields (S.21).

We now arrive at the proposition justifying (S.11).

**Proposition S.4.** *Under the same conditions as Lemma S.11,*

$$n\ d_n = \frac{35}{256\pi^{1/2}} \frac{((\sum_{i=1}^n V_i)^-)^2}{\sum_{i=1}^n V_i^2} + o_p(1).$$

**Proof**: From Lemma S.12 we have

$$n\ d_n = n\ \hat{s}_4^2 \int \frac{(x^4 - 6x^2 + 3)^2 \exp\left[-x^2\right]}{24^2\ (2\pi)}\ dx\ \{1 + o_p(1)\}. \tag{S.23}$$

The integral equals $(35/6144)\pi^{-1/2}$. From Lemma S.11 and the weak law of large numbers we have

$$n\ \hat{s}_4^2 = \frac{((\sum_{i=1}^n V_i)^-)^2}{n^{-1}(\sum_{i=1}^n V_i^2)^2} + o_p(1) = \frac{((\sum_{i=1}^n V_i)^-)^2}{(\sum_{i=1}^n V_i^2)E[V_i^2]} + o_p(1) = \frac{((\sum_{i=1}^n V_i)^-)^2}{(\sum_{i=1}^n V_i^2)/24} + o_p(1). \tag{S.24}$$

Combining (S.23) and (S.24) yields the desired result.

**Analysis with $\alpha_0 < 0.5$**

We begin by stating and proving two lemmas.

**Lemma S.13.** *Suppose that model (3) applies and that the null hypothesis is true with $\theta_0 = 0$ and $\sigma_0 = 1$. Under the indicated empirical Bayesian estimation framework, with $\alpha_0 < 0.5$ we have*

$$\hat{s}_1 = \frac{\sum X_i}{\sum X_i^2} + o_p(n^{-1/2}), \quad \hat{s}_2 = \frac{\sum Z_i}{\sum Z_i^2} + o_p(n^{-1/2}), \quad \textit{and}\ \ \hat{s}_3 = \frac{\sum U_i}{\sum U_i^2} + o_p(n^{-1/2}).$$

**Proof**: We can establish the following string of inequalities:

$$\begin{aligned}
2l_n^\dagger(\hat{\alpha},\hat{\theta}_1,\hat{\theta}_2,\hat{\sigma}) - 2l_n^\dagger(1/2,0,0,1) &\leq U_1^\dagger(\hat{s}_1) + U_2^\dagger(\hat{s}_2) + U_3^\dagger(\hat{s}_3) + 2\log[2\alpha_0] + o_p(1) \\
&\leq U_1^\dagger\left(\frac{\sum X_i}{\sum X_i^2}\right) + U_2^\dagger(\hat{s}_2) + U_3^\dagger(\hat{s}_3) + 2\log[2\alpha_0] + o_p(1) \\
&\leq U_1^\dagger\left(\frac{\sum X_i}{\sum X_i^2}\right) + U_2^\dagger\left(\frac{\sum Z_i}{\sum Z_i^2}\right) + U_3^\dagger(\hat{s}_3) + 2\log[2\alpha_0] + o_p(1) \\
&\leq U_1^\dagger\left(\frac{\sum X_i}{\sum X_i^2}\right) + U_2^\dagger\left(\frac{\sum Z_i}{\sum Z_i^2}\right) + U_3^\dagger\left(\frac{\sum U_i}{\sum U_i^2}\right) + 2\log[2\alpha_0] + o_p(1) \\
&\leq 2l_n^\dagger(\hat{\alpha},\hat{\theta}_1,\hat{\theta}_2,\hat{\sigma}) - 2l_n^\dagger(1/2,0,0,1) + o_p(1).
\end{aligned}$$

The first and last inequalities follow from a Taylor expansion argument in Chen and Li (2008). This Taylor expansion is different from the one in the proof of Lemma S.11 because now $\alpha_0 < 0.5$. However, the rest of the proof — in particular, the strategy of expressing the difference between any two successive lines as $o_p(1)$ — remains the same as for Lemma S.11.

**Lemma S.14.** *Under the same conditions as Lemma S.13,*

$$(1-\hat{\alpha})f_{\hat{\sigma}}(x,\hat{\theta}_1) + \hat{\alpha}\ f_{\hat{\sigma}}(x,\hat{\theta}_2) - f_{\hat{\sigma}_0}(x,\hat{\theta}_0) = \hat{s}_3 \frac{(x^3-3x)\exp\left[-\frac{x^2}{2}\right]}{6\ (2\pi)^{1/2}} + o_p(n^{-1/2})F_1,$$

*where $F_1$ is a remainder dominated by a function in $L^2$.*

**Proof**: Theorem 1 of Chen and Li (2008) shows that $\hat{\theta}_1 = O_p(n^{-1/6})$, $\hat{\theta}_2 = O_p(n^{-1/6})$, $\hat{\alpha} - \alpha_0 = O_p(n^{-1/4})$, and $\hat{\sigma}^2 - 1 = O_p(n^{-1/3})$. These results differ from those quoted in the proof of Lemma S.12 because now $\alpha_0 < 0.5$. Consequently, lines (S.16) and (S.19) in the proof of Lemma S.12 disappear, so we have only lines (S.17) and (S.18) with which to contend.

Line (S.17) is still absorbed into (S.20) by the same argument given in the proof of Lemma S.12. However, line (S.18) is no longer negligible since we do not have $\hat{\alpha} - 0.5 = O_p(n^{-1/4})$ when $\alpha_0 < 0.5$. In fact, by Lemma S.13 and Theorem 1 of Chen and Li (2008) we have

$$\frac{\hat{m}_3}{6} + \frac{\hat{m}_1(\hat{\sigma}^2-1)}{2} = \frac{\hat{s}_3}{6} + O_p(n^{-1/2})O_p(n^{-1/3}) = \frac{\hat{s}_3}{6} + o_p(n^{-1/2}),$$

which completes the proof.

We now arrive at the proposition justifying (S.12).

**Proposition S.5.** *Under the same conditions as Lemma S.13,*

$$n\ d_n = \frac{5}{32\pi^{1/2}} \frac{(\sum_{i=1}^n U_i)^2}{\sum_{i=1}^n U_i^2} + o_p(1).$$

**Proof**: From Lemma S.14 we have

$$n\ d_n = n\ \hat{s}_3^2 \int \frac{(x^3 - 3x)^2 \exp\left[-x^2\right]}{6^2\ (2\pi)}\ dx\ \{1 + o_p(1)\}. \tag{S.25}$$

The integral equals $(5/192)\pi^{-1/2}$. From Lemma S.13 and the weak law of large numbers we have

$$n\ \hat{s}_3^2 = \frac{(\sum_{i=1}^n U_i)^2}{n^{-1}(\sum_{i=1}^n U_i^2)^2} + o_p(1) = \frac{(\sum_{i=1}^n U_i)^2}{(\sum_{i=1}^n U_i^2)E[U_i^2]} + o_p(1) = \frac{(\sum_{i=1}^n U_i)^2}{(\sum_{i=1}^n U_i^2)/6} + o_p(1). \tag{S.26}$$

Combining (S.25) and (S.26) yields the desired result.