A. Proof of Lemma 2.5

To simplify the notation, we omit the ‘BO’ in the superscript. We begin with a few algebraic identities for $p_k$. It is easy to see from (2.18)-(2.19) that

$$p_k = \frac{k + a_0 - 2}{k + 2a_0} p_{k-1} \quad \text{for } k \geq 2. \quad (A.1)$$

Therefore, $\sum_{j=2}^{k}(j + 2a_0)p_j = \sum_{j=2}^{k}(j + a_0 - 2)p_{j-1}$ which implies that

$$p_{k-1} = \frac{k + 2a_0}{a_0 + 1} p_k \quad \text{for } k \geq 2. \quad (A.2)$$

Further by summing both sides of (A.2), we get $\sum_{k=1}^{\infty} kp_k = 2$. Observe that

$$\ell_{\infty}'(a) = \sum_{k=0}^{\infty} \frac{p_{k+1}}{a + k} - \frac{1}{a + 1} = \sum_{k=0}^{\infty} \frac{k + 2a_0}{(a_0 + 1)(a + k)} - \frac{1}{a + 1} \sum_{k=0}^{\infty} \frac{k}{k + a_0} p_{k+2}
= \frac{a - a_0}{(a_0 + 1)(a + 1)} \sum_{k=0}^{\infty} \frac{k + 2a_0}{(a + k)} (k - 1) p_{k+2}
= a - a_0 \sum_{k=0}^{\infty} \frac{k - 1}{a + k} p_{k+1}.$$ 

where the second equality is due to (A.2) and the last one stems from (A.1). In addition,

$$\sum_{k=0}^{\infty} \frac{k - 1}{a + k} p_{k+1} \leq \frac{1}{1 + a} \sum_{k=0}^{\infty} (k - 1) p_{k+1} = 0,$$

where the last equality is due to the fact that $\sum_{k=1}^{\infty} kp_k = 2$. Therefore, $\ell_{\infty}'(a)$ has a unique zero at $a_0$, and $\ell_{\infty}'(a) < 0$ if $a > a_0$ and $\ell_{\infty}'(a) > 0$ if $a < a_0$. These imply that $\ell_{\infty}(\cdot)$ has a unique maximum at $a_0$.

Now we prove (2.21). We have

$$\ell_n'(a) - \ell_{\infty}'(a) = \sum_{k=0}^{\infty} \frac{Z_{k+1}^n/n - p_{k+1}}{a + k} + \left( \frac{1}{a} \sum_{k=1}^{n} \frac{1}{a + 1 - k^{-1}} - \frac{1}{a + 1} \right). \quad (A.3)$$

Standard analysis shows that the second term on the r.h.s. of (A.3) goes to 0 as $n \to \infty$. Note that $(k + 2)Z_{k+1}^n = \sum_{j=2}^{k+2}(k + 2)Z_j^n \leq \sum_{j=2}^{k+2} jZ_j^n \leq 2n$, which implies $Z_{k+1}^n/n \leq 2k^{-1/2}$. Consequently,

$$\sup_{a > \varepsilon} \left| \sum_{k=0}^{\infty} \frac{Z_{k+1}^n/n - p_{k+1}}{a + k} \right| \leq \sum_{k=0}^{\infty} \frac{\sum_{k=0}^{n} Z_{k+1}^n/n - p_{k+1}}{\varepsilon + k}
+ \frac{2}{(2 + k)(a + k)} + \frac{p_{k+1}}{a + k} \quad (A.4)$$

The first term on the r.h.s. of (A.4) converges to 0 a.s. by Theorem 2.4, and the last two terms can be made arbitrarily small for $K$ sufficiently large. Combining the above estimates yields the desired result.

B. Proof of Lemma 2.6

It follows easily from the definition that $(\sum_{k=1}^{n} f_k(a_0); n \geq 1)$ is a martingale. To prove the convergence (2.24), it suffices to use Theorem 3.2 in (Hall & Heyde, 1980) with the following conditions:

- $n^{-1/2} \max_k |f_k(a_0)| \to 0$ in probability.
- $\mathbb{E}(n^{-1} \max_k f_k^2(a_0))$ is bounded in $n$.
- $n^{-1} \sum_{k=1}^{n} f_k^2(a_0) \to \sigma^2$ in probability.

The first two conditions are straightforward since $|f_k(a)| \leq 2/a$. Now we check the last condition. Write

$$\frac{1}{n} \sum_{k=1}^{n} f_k^2(a_0) = \frac{1}{n} \sum_{k=1}^{n} \frac{1}{(a + d_k(v(k)) - 1)^2} + \frac{1}{n} \sum_{k=1}^{n} \frac{1}{(a_0 + 1 - k^{-1})^2}$$

$$- \frac{2}{n} \sum_{k=1}^{n} (a + d_k(v(k)) - 1)(a + 1 - k^{-1})$$

$$\leq S_1 + S_2 - 2S_3.$$ 

Note that

$$S_1 = \sum_{k=0}^{\infty} \frac{Z_{k+1}^n/n}{(a_0 + k)^2} \to \sum_{k=0}^{\infty} \frac{p_{k+1}}{(a_0 + k)^2} \quad a.s.$$

which follows from Theorem 2.4. By standard analysis, $S_2 \to \frac{1}{(a_0+1)^2}$. We decompose $S_3$ into two terms:

$$S_3 = \frac{1}{n} \sum_{k=1}^{n} \frac{1}{a_0 + d_k(v(k)) - 1} \left( \frac{1}{a_0 + 1 - k^{-1}} - \frac{1}{a_0 + 1} \right)$$

$$+ \frac{1}{(a_0 + 1)n} \sum_{k=1}^{n} \frac{1}{a_0 + d_k(v(k)) - 1},$$

where the first term on the r.h.s. is bounded by $\frac{1}{an} \sum_{k=1}^{n} \left( \frac{1}{a_0 + 1 - k^{-1}} - \frac{1}{a_0 + 1} \right) \to 0$, and the second term converges almost surely to $\frac{1}{a_0+1} \sum_{k=0}^{n} \frac{p_{k+1}}{a_0 + k}$. Combining all the above estimates yields the desired result.
C. Proof of Lemma 2.7

Write
\[ \frac{1}{n} \sum_{k=1}^{n} f_k^\ell(a^*) = \frac{1}{n} \sum_{k=1}^{n} f_k^\ell(a_0) + \frac{1}{n} \sum_{k=1}^{n} (f_k^\ell(a^*) - f_k^\ell(a_0)) =: T_1 + T_2. \]

Observe that \( f_k^\ell(a) = -f_k^\ell(\alpha) - 2f_k^\ell(\alpha)_{a+1-k} \). We get
\[ T_1 = -\frac{1}{n} \sum_{k=1}^{n} f_k^\ell(a_0) - \frac{2}{n} \sum_{k=1}^{n} f_k^\ell(a_0) - \frac{1}{n} (a_0 + 1 - k). \]

The first term on the r.h.s. of (C.1) converges to \(-\sigma^2\) as proved in Proposition 2.6. Recall the definition of \( S_2, S_3 \). It is easy to see that
\[ \frac{1}{n} \sum_{k=1}^{n} f_k^\ell(a_0) - \frac{1}{a_0 + 1 - k} = S_3 - S_2 \]

Therefore, \( T_1 \to -\beta \) in probability. By standard analysis, \( |T_2| \leq C|a^* - a_0| \) for some \( C > 0 \). Note that \( a^* \in (a_0, \tilde{a}_n) \). By Theorem 2.1, \( |a^* - a_0| \to 0 \) which implies \( T_2 \to 0 \). The above estimates lead to the desired result.

D. Proof of Lemma 2.9

As discussed in Section 2.3, the consistency of \( \hat{\gamma} \) follows from standard exponential family theory. It suffices to prove that \( \hat{\gamma} \to \gamma^0 \) almost surely.

Let us go back to the limit log-likelihood (2.30). Observe that \( \ell_{\infty}^{BPA}(a) = \ell_{\infty}^{BPA}(\gamma) \) for each \( a > 0 \). By taking the partial derivatives of (2.30) and equating to 0, we get
\[ \frac{\partial}{\partial \gamma_{ij}} \ell_{\infty}^{BPA}(\gamma) = \begin{cases} \frac{\theta^0_n}{\gamma_{ij}} - \frac{\pi_{ik} p^0_i}{\gamma_{ik} p_k} - \frac{\pi_{jk} p^0_j}{\gamma_{jk} p_k} & \text{for } i \neq j, \\ \frac{\theta^0_n}{\gamma_{ii}} - \frac{\pi_{ik} p^0_i}{\gamma_{ik} p_k} & \text{for } i = j. \end{cases} \]

By (2.27), we have \( \nabla \ell_{\infty}^{BPA}(\gamma_0) = 0 \), i.e. \( \gamma_0 \) is a stationary point of \( \ell_{\infty}^{BPA} \). Now it suffices to prove Lemma 2.9 to conclude.

Note that \( \ell_{\infty}^{BPA}(\gamma) \to -\infty \) as \( \gamma \in \partial D \). It suffices to prove that \( \nabla \ell_{\infty}^{BPA}(\gamma) = 0 \) has a unique solution. First \( \partial \ell_{\infty}^{BPA}/\partial \gamma_{ii} = 0 \) gives
\[ \sum_{k=1}^{K} \gamma_{ik} p_k = \frac{\pi_{i0} p^0_i}{\theta^0_{i0} \gamma_{ii}}. \]

By injecting (D.2) into the equation \( \partial \ell_{\infty}^{BPA}/\partial \gamma_{ij} = 0 \), we get
\[ \frac{\theta^0_n}{\gamma_{ii}} = \frac{\theta^0_{i0} p^0_i}{\theta^0_{i0} \gamma_{ii}} + \frac{1}{\gamma_{ij}} \]

Consequently, the values of \( (\gamma_{ij}; i \neq j) \) is uniquely determined by those of \( (\gamma_{ii}; 1 \leq i \leq K) \). By injecting (D.3) into (D.2), we get a system of equations on \((\gamma_{ii}; 1 \leq i \leq K)\):
\[ \sum_{k=1}^{K} \theta_{ik} \left( \frac{\theta^0_{i0} p^0_i}{\theta^0_{i0} \gamma_{ii}} - \frac{1}{\gamma_{ij}} \right) p_k = \frac{\pi_{i0} p^0_i}{\theta^0_{i0} \gamma_{ii}} \]

For \( K = 2 \), it is easy to solve the equations together with the constraints \( \gamma_{11} = 1 \). For \( K \geq 3 \), the explicit solution is not available but we prove that the equations have a unique solution. To illustrate, we consider the generic case \( K = 3 \). All other cases can be proceeded in a similar way.

Let \( x_1 := \frac{\theta^0_{10} p^0_1}{\theta^0_{01} p^0_1} \gamma_{22} \left( \frac{\theta^0_{10} p^0_1}{\theta^0_{01} p^0_1} \gamma_{22} + \frac{\theta^0_{20} p^0_2}{\theta^0_{20} p^0_2} \gamma_{11} \right)^{-1}, \)
\( x_2 := \frac{\theta^0_{10} p^0_1}{\theta^0_{01} p^0_1} \gamma_{23} \left( \frac{\theta^0_{10} p^0_1}{\theta^0_{01} p^0_1} \gamma_{23} + \frac{\theta^0_{20} p^0_2}{\theta^0_{20} p^0_2} \gamma_{13} \right)^{-1}, \) and
\( x_3 := \frac{\theta^0_{10} p^0_1}{\theta^0_{01} p^0_1} \gamma_{11} \left( \frac{\theta^0_{10} p^0_1}{\theta^0_{01} p^0_1} \gamma_{11} + \frac{\theta^0_{20} p^0_2}{\theta^0_{20} p^0_2} \gamma_{13} \right)^{-1} \).

The equations (D.4) give
\[ \begin{align*}
\theta^0_{11} x_1 + \theta^0_{12} x_2 &= \pi^0_1 - \theta^0_{11}, \\
\theta^0_{11} (1 - x_1) + \theta^0_{21} (1 - x_3) &= \pi^0_2 - \theta^0_{22}, \\
\theta^0_{21} (1 - x_2) + \theta^0_{32} x_3 &= \pi^0_3 - \theta^0_{33}.
\end{align*} \]

It suffices to prove that the equations (D.5) have a unique solution. Observe that the system (D.5) has a solution \((x_1^*, x_2^*, x_3^*)\) by taking \( \gamma_{ii} = \gamma^0_{ii} \). Algebraic manipulation shows that the set of solutions to (D.5) has dimension 1, with form
\[ (x_1, x_2, x_3) = (x_1^0, x_2^0, x_3^0) + (1, -\theta^0_{12}/\theta^0_{13}, -\theta^0_{21}/\theta^0_{23}). \]

Consequently,
\[ \begin{align*}
\gamma_{11} &= \frac{\theta^0_{11} (p^0_1)^2}{\theta^0_{22} (p^0_2)^2} 1 - x_0 - \lambda, \\
\gamma_{13} &= \frac{\theta^0_{03} (p^0_3)^2}{\theta^0_{22} (p^0_2)^2} 1 - y_0 + \lambda \theta^0_{12} \theta^0_{13}, \\
\gamma_{22} &= \frac{\theta^0_{11} (p^0_1)^2}{\theta^0_{33} (p^0_3)^2} 1 - z_0 + \lambda \theta^0_{21} \theta^0_{23},
\end{align*} \]

which implies that
\[ \begin{align*}
1 - x_0 - \lambda &= \frac{(1 - y_0 + \lambda \theta^0_{12} \theta^0_{13}) (1 - z_0 + \lambda \theta^0_{21} \theta^0_{23})}{(y_0 - \lambda \theta^0_{21} \theta^0_{23})(z_0 - \lambda \theta^0_{12} \theta^0_{13})}.
\end{align*} \]

Note that the l.h.s. of (D.6) is decreasing in \( \lambda \) while the r.h.s. is increasing in \( \lambda \). Thus, \( \lambda = 0 \) is the only solution which proves the uniqueness.