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

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

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

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

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

$$\begin{aligned}
\ell'_\infty(a) &= \sum_{k \geq 0} \frac{p_{>k+1}}{a + k} - \frac{1}{a + 1} \\
&= \sum_{k \geq 0} \frac{(k + 2 + 2a_0) p_{k+2}}{(a_0 + 1)(a + k)} \\
&\qquad - \frac{1}{a + 1} \sum_{k \geq 0} \frac{k + 2 + 2a_0}{k + a_0} p_{k+2} \\
&= \frac{a - a_0}{(a_0 + 1)(a + 1)} \sum_{k \geq 0} \frac{(k + 2 + 2a_0)(k - 1)}{(k + a_0)(k + a)} p_{k+2} \\
&= \frac{a - a_0}{(a_0 + 1)(a + 1)} \sum_{k \geq 0} \frac{k - 1}{k + a} p_{k+1}.
\end{aligned}$$

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

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

where the last equality is due to the fact that $\sum_{k \geq 1} k p_k = 2$. Therefore, $\ell'_\infty(\cdot)$ 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

$$\begin{aligned}
\ell'_n(a) - \ell'_\infty(a) &= \sum_{k \geq 0} \frac{Z^n_{>k+1}/n - p_{>k+1}}{a + k} \\
&\quad + \left(\frac{1}{n} \sum_{k=1}^{n} \frac{1}{a + 1 - k^{-1}} - \frac{1}{a + 1} \right).
\end{aligned} \tag{A.3}$$

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

$$\begin{aligned}
\sup_{a > \varepsilon} \left| \sum_{k \geq 0} \frac{Z^n_{>k+1}/n - p_{>k+1}}{a + k} \right| &\leq \sum_{k=0}^{K} \frac{|Z^n_{>k+1}/n - p_{>k+1}|}{\varepsilon + k} \\
&\quad + \sum_{k > K} \frac{2}{(2 + k)(a + k)} + \sum_{k > K} \frac{p_{>k+1}}{a + k}.
\end{aligned} \tag{A.4}$$

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

$$\begin{aligned}
\frac{1}{n} \sum_{k=1}^{n} f_k^2(a_0) &= \frac{1}{n} \sum_{k=1}^{n} \frac{1}{(a_0 + d_k(v^{(k)}) - 1)^2} + \frac{1}{n} \sum_{k=1}^{n} \frac{1}{(a_0 + 1 - k^{-1})^2} \\
&\quad - \frac{2}{n} \sum_{k=1}^{n} \frac{1}{(a_0 + d_k(v^{(k)}) - 1)(a_0 + 1 - k^{-1})} \\
&:= S_1 + S_2 - 2S_3.
\end{aligned}$$

Note that

$$S_1 = \sum_{k \geq 0} \frac{Z^n_{>k+1}/n}{(a_0 + k)^2} \longrightarrow \sum_{k \geq 0} \frac{p_{>k+1}}{(a_0 + k)^2} \quad a.s.$$

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

$$\begin{aligned}
S_3 &= \frac{1}{n} \sum_{k=1}^{n} \frac{1}{a_0 + d_k(v^{(k)}) - 1} \left(\frac{1}{a_0 + 1 - k^{-1}} - \frac{1}{a_0 + 1} \right) \\
&\quad + \frac{1}{(a_0 + 1)n} \sum_{k=1}^{n} \frac{1}{a_0 + d_k(v^{(k)}) - 1},
\end{aligned}$$

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

C. Proof of Lemma 2.7

Write

$$\frac{1}{n}\sum_{k=1}^{n} f_k'(a^\star) = \frac{1}{n}\sum_{k=1}^{n} f_k'(a_0) + \frac{1}{n}\sum_{k=1}^{n}\left(f_k'(a^\star) - f_k'(a_0)\right) := T_1 + T_2.$$

Observe that $f_k'(a) = -f_k^2(a) - 2f_k(a)\frac{1}{a+1-k^{-1}}$. We get

$$T_1 = -\frac{1}{n}\sum_{k=1}^{n} f_k^2(a_0) - \frac{2}{n}\sum_{k=1}^{n} f_k(a_0)\frac{1}{a_0+1-k^{-1}}. \tag{C.1}$$

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

$$\frac{1}{n}\sum_{k=1}^{n} f_k(a_0)\frac{1}{a_0+1-k^{-1}} = S_3 - S_2 \longrightarrow \frac{1}{a+1}\sum_{k\geq 0}\frac{p_{>k+1}}{a_0+k} - \frac{1}{(a_0+1)^2}.$$

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

D. Proof of Lemma 2.9

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

Let us go back to the limit log-likelihood (2.30). Observe that ℓ_∞^{BPA} is homogeneous of order 1, i.e. $\ell_\infty^{BPA}(a\boldsymbol{\gamma}) = \ell_\infty^{BPA}(\boldsymbol{\gamma})$ for each $a > 0$. By taking the partial derivatives of (2.30) and equating to 0, we get

$$\frac{\partial}{\partial\gamma_{ij}}\ell_\infty^{BPA}(\boldsymbol{\gamma}) = \begin{cases} \frac{\theta_{ij}^0}{\gamma_{ij}} - \frac{\pi_i^0 p_j^0}{\sum_{k=1}^K \gamma_{ik}p_k^0} - \frac{\pi_j^0 p_i^0}{\sum_{k=1}^K \gamma_{jk}p_k^0} & \text{for} \quad i\neq j, \\ \frac{\theta_{ii}^0}{\gamma_{ii}} - \frac{\pi_i^0 p_i^0}{\sum_{k=1}^K \gamma_{ik}p_k^0} & \text{for} \quad i = j. \end{cases} \tag{D.1}$$

By (2.27), we have $\nabla\ell_\infty^{BPA}(\boldsymbol{\gamma}_0) = \mathbf{0}$, i.e. $\boldsymbol{\gamma}_0$ is a stationary point of ℓ_∞^{BPA}. Now it suffices to prove Lemma 2.9 to conclude.

Note that $\ell_\infty^{BPA}(\boldsymbol{\gamma}) \to -\infty$ as $\boldsymbol{\gamma} \in \partial\mathcal{D}$. It suffices to prove that $\nabla\ell_\infty^{BPA}(\boldsymbol{\gamma}) = \mathbf{0}$ has a unique solution. First $\partial\ell_\infty^{BPA}/\partial\gamma_{ii} = 0$ gives

$$\sum_{k=1}^K \gamma_{ik}p_k^0 = \frac{\pi_i^0 p_i^0}{\theta_{ii}^0}\gamma_{ii}. \tag{D.2}$$

By injecting (D.2) into the equation $\partial\ell_\infty^{BPA}/\partial\gamma_{ij} = 0$, we get

$$\frac{\theta_{ij}^0}{\gamma_{ij}} = \frac{\theta_{ii}^0 p_j^0}{p_i^0}\frac{1}{\gamma_{ii}} + \frac{\theta_{jj}^0 p_i^0}{p_j^0}\frac{1}{\gamma_{jj}}. \tag{D.3}$$

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 $(\gamma_{ii};\, 1\leq i\leq K)$:

$$\sum_{k=1}^K \theta_{ik}^0\left(\frac{\theta_{ii}^0 p_j^0}{p_i^0}\frac{1}{\gamma_{ii}} + \frac{\theta_{kk}^0 p_i^0}{p_k^0}\frac{1}{\gamma_{kk}}\right)^{-1} p_k^0 = \frac{\pi_i^0 p_i^0}{\theta_{ii}^0}\gamma_{ii} \tag{D.4}$$

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 $x_1 := \frac{\theta_{11}^0 p_2^0}{p_1^0}\gamma_{22}\left(\frac{\theta_{11}^0 p_2^0}{p_1^0}\gamma_{22} + \frac{\theta_{22}^0 p_1^0}{p_2^0}\gamma_{11}\right)^{-1}$, $x_2 := \frac{\theta_{11}^0 p_3^0}{p_1^0}\gamma_{33}\left(\frac{\theta_{11}^0 p_3^0}{p_1^0}\gamma_{33} + \frac{\theta_{33}^0 p_1^0}{p_3^0}\gamma_{11}\right)^{-1}$, and $x_3 := \frac{\theta_{33}^0 p_2^0}{p_3^0}\gamma_{22}\left(\frac{\theta_{33}^0 p_2^0}{p_3^0}\gamma_{22} + \frac{\theta_{22}^0 p_3^0}{p_2^0}\gamma_{33}\right)^{-1}$. The equations (D.4) give

$$\begin{cases} \theta_{12}^0 x_1 + \theta_{13}^0 x_2 = \pi_1^0 - \theta_{11}^0, \\ \theta_{21}^0(1-x_1) + \theta_{23}^0(1-x_3) = \pi_2^0 - \theta_{22}^0, \\ \theta_{31}^0(1-x_2) + \theta_{32}^0 x_3 = \pi_3^0 - \theta_{33}^0. \end{cases} \tag{D.5}$$

It suffices to prove that the equations (D.5) have a unique solution. Observe that the system (D.5) has a solution (x_1^0, x_2^0, x_3^0) by taking $\gamma_{ii} = \gamma_{ii}^0$. Algebraic manipulation shows that the set of solutions to (D.5) has dimension 1, with form

$$(x_1, x_2, x_3) = (x_1^0, x_2^0, x_3^0) + \lambda(1, -\theta_{12}^0/\theta_{13}^0, -\theta_{21}^0/\theta_{23}^0).$$

Consequently,

$$\frac{\gamma_{11}}{\gamma_{22}} = \frac{\theta_{11}^0(p_2^0)^2}{\theta_{22}^0(p_1^0)^2}\frac{1-x_0-\lambda}{x_0+\lambda}, \quad \frac{\gamma_{11}}{\gamma_{13}} = \frac{\theta_{11}^0(p_3^0)^2}{\theta_{33}^0(p_1^0)^2}\frac{1-y_0+\lambda\theta_{12}^0\theta_{13}^0}{y_0-\lambda\theta_{12}^0\theta_{13}^0}$$

$$\frac{\gamma_{33}}{\gamma_{22}} = \frac{\theta_{33}^0(p_2^0)^2}{\theta_{22}^0(p_3^0)^2}\frac{1-z_0+\lambda\theta_{21}^0/\theta_{23}^0}{z_0-\lambda\theta_{21}^0/\theta_{23}^0},$$

which implies that

$$\frac{1-x_0-\lambda}{x_0+\lambda} = \frac{(1-y_0+\lambda\theta_{12}^0\theta_{13}^0)(1-z_0+\lambda\theta_{21}^0/\theta_{23}^0)}{(y_0-\lambda\theta_{12}^0\theta_{13}^0)(z_0-\lambda\theta_{21}^0/\theta_{23}^0)}. \tag{D.6}$$

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