Appendix

Lemma 3.1. *Suppose $\pi_\theta(u|s)$ is a compatible conditional PDF of $u \in \mathbb{R}$ whose cumulative distribution function (CDF) is $\Pi_\theta(u|s)$. Then, the following equations hold:*

$$\mathbb{E}_u[\mathbb{1}_{u\le\alpha}\nabla_\theta \log \pi_\theta(u|s)] = \mathbb{E}_u[\mathbb{1}_{u\le\alpha}\nabla_\theta \log \Pi_\theta(\alpha|s)],$$
$$\mathbb{E}_u[\mathbb{1}_{\beta\le u}\nabla_\theta \log \pi_\theta(u|s)] = \mathbb{E}_u[\mathbb{1}_{\beta\le u}\nabla_\theta \log(1-\Pi_\theta(\beta|s))].$$

Proof. Noting that $\pi_\theta(u|s)$ allows the exchange of derivative and integral, we obtain

$$\begin{aligned}
\mathbb{E}_u[\mathbb{1}_{u\le\alpha}\nabla_\theta \log \pi_\theta(u|s)] &= \int_{-\infty}^{\alpha} \pi_\theta(u|s)\nabla_\theta \log \pi_\theta(u|s) du \\
&= \int_{-\infty}^{\alpha} \nabla_\theta \pi_\theta(u|s) du \\
&= \nabla_\theta \int_{-\infty}^{\alpha} \pi_\theta(u|s) du \\
&= \nabla_\theta \Pi_\theta(\alpha|s) \\
&= \Pi_\theta(\alpha|s)\nabla_\theta \log \Pi_\theta(\alpha|s) \\
&= \mathbb{E}_u[\mathbb{1}_{u\le\alpha}\nabla_\theta \log \Pi_\theta(\alpha|s)].
\end{aligned}$$

A similar calculation shows

$$\mathbb{E}_u[\mathbb{1}_{\beta\le u}\nabla_\theta \log \pi_\theta(u|s)] = \mathbb{E}_u[\mathbb{1}_{\beta\le u}\nabla_\theta \log(1-\Pi_\theta(\beta|s))],$$

where we used $\int_{\beta}^{\infty} \pi_\theta(u|s)du = 1-\Pi_\theta(\beta|s)$ instead of $\int_{-\infty}^{\alpha} \pi_\theta(u|s)du = \Pi_\theta(\alpha|s)$. □

Lemma 3.2. *Suppose $\pi_\theta(u|s)$ is a compatible conditional PDF of $u \in \mathbb{R}$ whose CDF is $\Pi_\theta(u|s)$. Then, the following inequalities hold:*

$$\mathbb{V}_u[\mathbb{1}_{u\le\alpha}\nabla_\theta \log \pi_\theta(u|s)] \ge \mathbb{V}_u[\mathbb{1}_{u\le\alpha}\nabla_\theta \log \Pi_\theta(\alpha|s)],$$
$$\mathbb{V}_u[\mathbb{1}_{\beta\le u}\nabla_\theta \log \pi_\theta(u|s)] \ge \mathbb{V}_u[\mathbb{1}_{\beta\le u}\nabla_\theta \log(1-\Pi_\theta(\beta|s))].$$

The equalities hold only when $\nabla_\theta \log \pi_\theta(u|s)$ is constant over $u \le \alpha$ and $\beta \le u$, respectively.

Proof. Because both $\mathbb{1}_{u\le\alpha}\nabla_\theta \log \pi_\theta(u|s)$ and $\mathbb{1}_{u\le\alpha}\nabla_\theta \log \Pi_\theta(\alpha|s)$ have the same expected values from Lemma 3.1, the difference of their variances is written as follows:

$$\mathbb{V}_u[\mathbb{1}_{u\le\alpha}\nabla_\theta \log \pi_\theta(u|s)] - \mathbb{V}_u[\mathbb{1}_{u\le\alpha}\nabla_\theta \log \Pi_\theta(\alpha|s)] = \mathbb{E}_u[\mathbb{1}_{u\le\alpha}(\nabla_\theta \log \pi_\theta(u|s))^2] - \mathbb{E}_u[\mathbb{1}_{u\le\alpha}(\nabla_\theta \log \Pi_\theta(\alpha|s))^2].$$

The difference above is nonnegative because

$$\begin{aligned}
\mathbb{E}_u[\mathbb{1}_{u\le\alpha}(\nabla_\theta \log \pi_\theta(u|s))^2] &= \int_{-\infty}^{\alpha} \pi_\theta(u|s)(\nabla_\theta \log \pi_\theta(u|s))^2 du \\
&= \Pi_\theta(\alpha|s) \int_{-\infty}^{\infty} \mathbb{1}_{u\le\alpha} \frac{\pi_\theta(u|s)}{\Pi_\theta(\alpha|s)} (\nabla_\theta \log \pi_\theta(u|s))^2 du \\
&\ge \Pi_\theta(\alpha|s) \Big(\int_{-\infty}^{\infty} \mathbb{1}_{u\le\alpha} \frac{\pi_\theta(u|s)}{\Pi_\theta(\alpha|s)} \nabla_\theta \log \pi_\theta(u|s) du \Big)^2 \\
&= \Pi_\theta(\alpha|s) \Big(\frac{1}{\Pi_\theta(\alpha|s)} \int_{-\infty}^{\alpha} \pi_\theta(u|s) \nabla_\theta \log \pi_\theta(u|s) du \Big)^2 \\
&= \Pi_\theta(\alpha|s) \Big(\frac{1}{\Pi_\theta(\alpha|s)} \int_{-\infty}^{\alpha} \nabla_\theta \pi_\theta(u|s) du \Big)^2 \\
&= \Pi_\theta(\alpha|s) \Big(\frac{1}{\Pi_\theta(\alpha|s)} \nabla_\theta \int_{-\infty}^{\alpha} \pi_\theta(u|s) du \Big)^2 \\
&= \Pi_\theta(\alpha|s) \Big(\frac{1}{\Pi_\theta(\alpha|s)} \nabla_\theta \Pi_\theta(\alpha|s) \Big)^2 \\
&= \Pi_\theta(\alpha|s)(\nabla_\theta \log \Pi_\theta(\alpha|s))^2 \\
&= \mathbb{E}_u[\mathbb{1}_{u\le\alpha}(\nabla_\theta \log \Pi_\theta(\alpha|s))^2],
\end{aligned}$$

where the equality holds only when $\nabla_\theta \log \pi_\theta(u|s)$ is constant over $u \le \alpha$.

A similar calculation shows

$$\mathbb{V}_u[\mathbb{1}_{\beta\le u} \nabla_\theta \log \pi_\theta(u|s)] \ge \mathbb{V}_u[\mathbb{1}_{\beta\le u} \nabla_\theta \log(1 - \Pi_\theta(\beta|s))],$$

where the equality holds only when $\nabla_\theta \log \pi_\theta(u|s)$ is constant over $\beta \le u$. □

Lemma 3.4. *Suppose $\pi_\theta(\mathbf{u}|s)$ is a conditional PDF of $\mathbf{u} \in \mathbb{R}^d$ ($d \ge 2$) whose CDF is $\Pi_\theta(\mathbf{u}|s)$. The conditional PDF and CDF of u_i are denoted by $\pi_\theta^{(i)}$ and $\Pi_\theta^{(i)}$, respectively. Suppose each $\pi_\theta^{(i)}$ is compatible and the elements of $\mathbf{u}$ are conditionally independent given s. Let $f(s, \mathbf{u})$ be a real-valued function such that $f(s, \mathbf{u}) = f(s, \mathrm{clip}(\mathbf{u}, \alpha, \beta))$. Define $\psi(s, \mathbf{u}) = \sum_i \psi^{(i)}(s, u_i)$, where $\psi^{(i)}(s, u) = \nabla_\theta \log \pi_\theta^{(i)}(u|s)$. Similarly, define $\overline{\psi}(s, \mathbf{u}) = \sum_i \overline{\psi}^{(i)}(s, u_i)$, where*

$$\overline{\psi}^{(i)}(s, u) = \begin{cases} \nabla_\theta \log \Pi_\theta^{(i)}(\alpha|s) & \text{if } u \le \alpha \\ \nabla_\theta \log \pi_\theta^{(i)}(u|s) & \text{if } \alpha < u < \beta \\ \nabla_\theta \log(1 - \Pi_\theta^{(i)}(\beta|s)) & \text{if } \beta \le u \end{cases}.$$

Then, the following equality and inequality hold:

$$\mathbb{E}_{\mathbf{u}}[f(s, \mathbf{u})\overline{\psi}(s, \mathbf{u})] = \mathbb{E}_{\mathbf{u}}[f(s, \mathbf{u})\psi(s, \mathbf{u})], \tag{11}$$

$$\mathbb{V}_{\mathbf{u}}[f(s, \mathbf{u})\overline{\psi}(s, \mathbf{u})] \le \mathbb{V}_{\mathbf{u}}[f(s, \mathbf{u})\psi(s, \mathbf{u})]. \tag{12}$$

The equality of the variances holds only when $\psi^{(i)}(s, u)$ is constant over both $u \le \alpha$ and $\beta \le u$ for all $1 \le i \le d$.

Proof. Applying Lemma 3.3 to each u_i yields

$$\mathbb{E}_{\mathbf{u}}[f(s, \mathbf{u})\overline{\psi}^{(i)}(s, u_i)] = \mathbb{E}_{\mathbf{u}}[f(s, \mathbf{u})\psi^{(i)}(s, u_i)], \tag{14}$$

$$\mathbb{V}_{\mathbf{u}}[f(s, \mathbf{u})\overline{\psi}^{(i)}(s, u_i)] \le \mathbb{V}_{\mathbf{u}}[f(s, \mathbf{u})\psi^{(i)}(s, u_i)]. \tag{15}$$

Because each action is conditionally independent, we can decompose the expectations as

$$\mathbb{E}_{\mathbf{u}}[f(s, \mathbf{u})\psi(s, \mathbf{u})] = \sum_{i=1}^{d} \mathbb{E}_{\mathbf{u}}[f(s, \mathbf{u})\psi^{(i)}(s, u_i)],$$

$$\mathbb{E}_{\mathbf{u}}[f(s, \mathbf{u})\overline{\psi}(s, \mathbf{u})] = \sum_{i=1}^{d} \mathbb{E}_{\mathbf{u}}[f(s, \mathbf{u})\overline{\psi}^{(i)}(s, u_i)].$$

From (14), these two are equal, and hence (11) holds.

The variances can also be decomposed as

$$\mathbb{V}_{\mathbf{u}}[f(s,\mathbf{u})\psi(s,\mathbf{u})] = \sum_{i=1}^{d} \mathbb{V}_{\mathbf{u}}[f(s,\mathbf{u})\psi^{(i)}(s,u_i)] + \sum_{1\le i<j\le d} \mathrm{Cov}[f(s,\mathbf{u})\psi^{(i)}(s,u_i), f(s,\mathbf{u})\psi^{(j)}(s,u_j)], \quad (16)$$

$$\mathbb{V}_{\mathbf{u}}[f(s,\mathbf{u})\overline{\psi}(s,\mathbf{u})] = \sum_{i=1}^{d} \mathbb{V}_{\mathbf{u}}[f(s,\mathbf{u})\overline{\psi}^{(i)}(s,u_i)] + \sum_{1\le i<j\le d} \mathrm{Cov}[f(s,\mathbf{u})\overline{\psi}^{(i)}(s,u_i), f(s,\mathbf{u})\overline{\psi}^{(j)}(s,u_j)], \quad (17)$$

where

$$\begin{aligned}&\mathrm{Cov}[f(s,\mathbf{u})\psi^{(i)}(s,u_i), f(s,\mathbf{u})\psi^{(j)}(s,u_j)]\\ &\quad= \mathbb{E}_{\mathbf{u}}[(f(s,\mathbf{u}))^2\psi^{(i)}(s,u_i)\psi^{(j)}(s,u_j)] - \mathbb{E}_{\mathbf{u}}[f(s,\mathbf{u})\psi^{(i)}(s,u_i)]\mathbb{E}_{\mathbf{u}}[f(s,\mathbf{u})\psi^{(j)}(s,u_j)], \quad (18)\\ &\mathrm{Cov}[f(s,\mathbf{u})\overline{\psi}^{(i)}(s,u_i), f(s,\mathbf{u})\overline{\psi}^{(j)}(s,u_j)]\\ &\quad= \mathbb{E}_{\mathbf{u}}[(f(s,\mathbf{u}))^2\overline{\psi}^{(i)}(s,u_i)\overline{\psi}^{(j)}(s,u_j)] - \mathbb{E}_{\mathbf{u}}[f(s,\mathbf{u})\overline{\psi}^{(i)}(s,u_i)]\mathbb{E}_{\mathbf{u}}[f(s,\mathbf{u})\overline{\psi}^{(j)}(s,u_j)]. \quad (19)\end{aligned}$$

The first term of (17) is smaller than or equal to that of (16) from (15). Thus, to prove (12), it is sufficient to show that the second terms of (16) and (17) are equal.

The second terms of (18) and (19) are equal from (14). Using the law of total variance, the first terms of (18) and (19) can be written as

$$\begin{aligned}\mathbb{E}_{\mathbf{u}}[(f(s,\mathbf{u}))^2\psi^{(i)}(s,u_i)\psi^{(j)}(s,u_j)] &= \mathbb{E}_{\mathbf{u}_{\setminus i,j}}\left[\mathbb{E}_{u_i}\left[\mathbb{E}_{u_j}[(f(s,\mathbf{u}))^2\psi^{(i)}(s,u_i)\psi^{(j)}(s,u_j)|\mathbf{u}_{\setminus j}]\middle|\mathbf{u}_{\setminus i,j}\right]\right]\\ &= \mathbb{E}_{\mathbf{u}_{\setminus i,j}}\left[\mathbb{E}_{u_i}\left[\psi^{(i)}(s,u_i)\mathbb{E}_{u_j}[(f(s,\mathbf{u}))^2\psi^{(j)}(s,u_j)|\mathbf{u}_{\setminus j}]\middle|\mathbf{u}_{\setminus i,j}\right]\right], \quad (20)\\ \mathbb{E}_{\mathbf{u}}[(f(s,\mathbf{u}))^2\overline{\psi}^{(i)}(s,u_i)\overline{\psi}^{(j)}(s,u_j)] &= \mathbb{E}_{\mathbf{u}_{\setminus i,j}}\left[\mathbb{E}_{u_i}\left[\overline{\psi}^{(i)}(s,u_i)\mathbb{E}_{u_j}[(f(s,\mathbf{u}))^2\overline{\psi}^{(j)}(s,u_j)|\mathbf{u}_{\setminus j}]\middle|\mathbf{u}_{\setminus i,j}\right]\right], \quad (21)\end{aligned}$$

where $\mathbf{u}_{\setminus j}$ denotes a vector $\mathbf{u}$ with the j-th element excluded, and $\mathbf{u}_{\setminus i,j}$ denotes a vector $\mathbf{u}$ with the i-th and j-th elements excluded. Noting the fact that $(f(s,\mathbf{u}))^2$ is a function of u_j conditioned on s and $\mathbf{u}_{\setminus j}$, we can have the following equation by applying Lemma 3.3.

$$\mathbb{E}_{u_j}[(f(s,\mathbf{u}))^2\overline{\psi}^{(j)}(s,u_j)|\mathbf{u}_{\setminus j}] = \mathbb{E}_{u_j}[(f(s,\mathbf{u}))^2\psi^{(j)}(s,u_j)|\mathbf{u}_{\setminus j}].$$

Similarly, we can use the fact that $\mathbb{E}_{u_j}[(f(s,\mathbf{u}))^2\psi^{(j)}(s,u_j)|\mathbf{u}_{\setminus j}]$ is a function of u_i conditioned on s and $\mathbf{u}_{\setminus i,j}$ to show

$$\mathbb{E}_{u_i}\left[\overline{\psi}^{(i)}(s,u_i)\mathbb{E}_{u_j}[(f(s,\mathbf{u}))^2\overline{\psi}^{(j)}(s,u_j)|\mathbf{u}_{\setminus j}]\middle|\mathbf{u}_{\setminus i,j}\right] = \mathbb{E}_{u_i}\left[\psi^{(i)}(s,u_i)\mathbb{E}_{u_j}[(f(s,\mathbf{u}))^2\psi^{(j)}(s,u_j)|\mathbf{u}_{\setminus j}]\middle|\mathbf{u}_{\setminus i,j}\right]. \quad (22)$$

From (22), we can see (20) and (21) are equal. This implies that the first terms of (18) and (19) are equal, and the second terms of (16) and (17) are equal. Therefore, (12) is satisfied. The equality of (12) holds only when $\nabla_\theta \log \pi_\theta{}^{(i)}(u|s)$ is constant over both $u \le \alpha$ and $\beta \le u$ for all $1 \le i \le d$. □