

We correct equation 4 as

$$F_k(\theta, M, P_O) = \begin{cases} (X_k - \beta)^+ + (X_k - \bar{\beta})^+ - (X_k - \alpha)^- - (X_k - \bar{\alpha})^+, & \text{Case 1 or 3,} \\ (X_k - \beta)^+ + (X_k - \bar{\beta})^- - (X_k - \alpha)^- - (X_k - \bar{\alpha})^-, & \text{Case 2 or 4,} \end{cases} \quad (1)$$

Next, we provide the proof for theorem 2 as follows, where i represents the i -th generation step.

Let θ_i^r to denote the reversed permutation of $\theta_i \in \Theta$. Noting that P_Θ is uniform on Θ , which means $P_\Theta(\theta_i) = \frac{1}{|\Theta|}$ for each $\theta_i \in \Theta$. Therefore, with message M embedded,

$$\begin{aligned} & 2_{\theta_i \sim P_\Theta} [P_W^M(x_i|a, x_{1:i-1}, \theta_i)] \\ &= 2 \sum_{\theta_i \in \Theta} \frac{1}{|\Theta|} P_W^M(x_i|a, x_{1:i-1}, \theta_i) \\ &= \sum_{\theta_i \in \Theta} \frac{1}{|\Theta|} P_W^M(x_i|a, x_{1:i-1}, \theta_i) + \sum_{\theta_i \in \Theta} \frac{1}{|\Theta|} P_W^M(x_i|a, x_{1:i-1}, \theta_i) \\ &= \sum_{\theta_i \in \Theta} \frac{1}{|\Theta|} P_W^M(x_i|a, x_{1:i-1}, \theta_i) + \sum_{\theta_i^r \in \Theta} \frac{1}{|\Theta|} P_W^M(x_i|a, x_{1:i-1}, \theta_i^r) \\ &= \frac{1}{|\Theta|} \sum_{\theta_i \in \Theta} [P_W^M(x_i|a, x_{1:i-1}, \theta_i) + P_W^M(x_i|a, x_{1:i-1}, \theta_i^r)]. \end{aligned}$$

Next, we will show that

$$P_W^M(x_i|a, x_{1:i-1}, \theta_i) + P_W^M(x_i|a, x_{1:i-1}, \theta_i^r) = 2P_O(x_i|a, x_{1:i-1}).$$

For ease of notation, let t denote the position of x_i in the ordered token set θ_i . Then $|V| + 1 - t$ is the position of x_i in the reversed permutation θ_i^r ,

$$P_W^M(x_i|a, x_{1:i-1}, \theta_i) = F_t(\theta_i, M, P_O) - F_{t-1}(\theta_i, M, P_O)$$

and

$$P_W^M(x_i|a, x_{1:i-1}, \theta_i^r) = F_{|V|+1-t}(\theta_i^r, M, P_O) - F_{|V|-t}(\theta_i^r, M, P_O),$$

Let $X_t = \sum_{j=1}^t P_O(v_j|a, x_{1:i-1}, \theta_i)$ and $X_{|V|+1-t}^r = \sum_{j=1}^{|V|+1-t} P_O(v_j|a, x_{1:i-1}, \theta_i^r)$.

For case 1 or 3:

We know that

$$\begin{aligned} F_{t-1}(\theta_i, M, P_O) &= (X_{t-1} - \beta)^+ + (X_{t-1} - \bar{\beta})^+ - (X_{t-1} - \alpha)^- - (X_{t-1} - \bar{\alpha})^+, \\ F_{|V|+1-t}(\theta_i^r, M, P_O) &= (X_{|V|+1-t}^r - \beta)^+ + (X_{|V|+1-t}^r - \bar{\beta})^+ - (X_{|V|+1-t}^r - \alpha)^- - (X_{|V|+1-t}^r - \bar{\alpha})^+. \end{aligned}$$

Since $\sum_{j=1}^{|V|+1-t} P_O(v_j|a, x_{1:i-1}, \theta_i^r) = 1 - \sum_{j=1}^{t-1} P_O(v_j|a, x_{1:i-1}, \theta_i)$, then $X_{|V|+1-t}^r = 1 - X_{t-1}$. Therefore,

$$\begin{aligned} F_{|V|+1-t}(\theta_i^r, M, P_O) &= (1 - \beta - X_{t-1})^+ + (1 - \bar{\beta} - X_{t-1})^+ - (1 - \alpha - X_{t-1})^- - (1 - \bar{\alpha} - X_{t-1})^+ \\ &= (\bar{\beta} - X_{t-1})^+ + (\beta - X_{t-1})^+ - (\bar{\alpha} - X_{t-1})^- - (\alpha - X_{t-1})^+ \end{aligned}$$

Using the relations $(x)^+ - (-x)^+ = x$ and $(x)^- - (-x)^+ = 0$, we obtain

$$F_{|V|+1-t}(\theta_i^r, M, P_O) - F_{t-1}(\theta_i, M, P_O) = \bar{\beta} - X_{t-1} + \beta - X_{t-1} = 1 - 2X_{t-1}.$$

Similarly, $F_t(\theta_i, M, P_O) - F_{|V|-t}(\theta_i^r, M, P_O) = 1 - 2X_t$. Therefore,

$$\begin{aligned} & P_W^M(x_i|a, x_{1:i-1}, \theta_i) + P_W^M(x_i|a, x_{1:i-1}, \theta_i^r) \\ &= 1 - 2X_{t-1} - (1 - 2X_t) \\ &= 2(X_t - X_{t-1}) = 2P_O(x_i|a, x_{1:i-1}). \end{aligned}$$

Similarly, for case 2 or 4:

We know that

$$\begin{aligned} F_{t-1}(\theta_i, M, P_O) &= (X_{t-1} - \beta)^+ + (X_{t-1} - \bar{\beta})^- - (X_{t-1} - \alpha)^- - (X_{t-1} - \bar{\alpha})^-, \\ F_{|V|+1-t}(\theta_i^r, M, P_O) &= (X_{|V|+1-t}^r - \beta)^+ + (X_{|V|+1-t}^r - \bar{\beta})^- - (X_{|V|+1-t}^r - \alpha)^- - (X_{|V|+1-t}^r - \bar{\alpha})^-. \end{aligned}$$

Since $\sum_{j=1}^{|V|+1-t} P_O(v_j|a, x_{1:i-1}, \theta_i^r) = 1 - \sum_{j=1}^{t-1} P_O(v_j|a, x_{1:i-1}, \theta_i)$, then $X_{|V|+1-t}^r = 1 - X_{t-1}$. Therefore,

$$\begin{aligned} F_{|V|+1-t}(\theta_i^r, M, P_O) &= (1 - \beta - X_{t-1})^+ + (1 - \bar{\beta} - X_{t-1})^- - (1 - \alpha - X_{t-1})^- - (1 - \bar{\alpha} - X_{t-1})^- \\ &= (\bar{\beta} - X_{t-1})^+ + (\beta - X_{t-1})^- - (\bar{\alpha} - X_{t-1})^- - (\alpha - X_{t-1})^- \end{aligned}$$

Using the relations $(x)^- - (-x)^- = -x$ and $(x)^+ - (-x)^- = 0$, we obtain

$$F_{|V|+1-t}(\theta_i^r, M, P_O) - F_{t-1}(\theta_i, M, P_O) = \bar{\alpha} - X_{t-1} + \alpha - X_{t-1} = 1 - 2X_{t-1}.$$

Similarly, $F_t(\theta_i, M, P_O) - F_{|V|-t}(\theta_i^r, M, P_O) = 1 - 2X_t$. Therefore,

$$\begin{aligned} P_W^M(x_i|a, x_{1:i-1}, \theta_i) + P_W^M(x_i|a, x_{1:i-1}, \theta_i^r) \\ &= 1 - 2X_{t-1} - (1 - 2X_t) \\ &= 2(X_t - X_{t-1}) = 2P_O(x_i|a, x_{1:i-1}). \end{aligned}$$