The decision to send σ_G or not matters only if the team achieves G *and* one agent comes to know this fact. We define the random variable, K_G, to be the earliest time at which an agent knows this fact. We denote agent A_G as the agent who knows of the achievement at time K_G. If $A_G = i$, for some agent, i, and $K_G = t_0$, then agent i has some pre-communication belief state, $b_{i \bullet \Sigma} = \beta$, that indicates that G has been achieved. To more precisely quantify the difference between agent i sending the σ_G message at time K_G vs. never sending it, we define the following value:

$$
\Delta^T(t_0, i, \beta) \equiv E\left[\sum_{t=0}^{T-t_0} R^{t_0+t} \,\middle|\, \Sigma_i^{t_0} = \sigma_G, K_G = t_0, A_G = i, b_{i \bullet \Sigma}^{t_0} = \beta\right]
$$
$$
- E\left[\sum_{t=0}^{T-t_0} R^{t_0+t} \,\middle|\, \Sigma_i^{t_0} = \text{null}, K_G = t_0, A_G = i, b_{i \bullet \Sigma}^{t_0} = \beta\right] \tag{1}
$$

We assume that, for all times other than K_G, the agents follow some communication policy, $\boldsymbol{\pi}_{\alpha\Sigma}$, that never specifies σ_G. Thus, Δ^T measures the difference in expected reward that hinges on agent i's specific decision to send or not send σ_G at time t_0. Given this definition, it is locally optimal for agent i to send the special message, σ_G, at time t_0, if and only if $\Delta^T \geq 0$. We define the communication policy, $\boldsymbol{\pi}_{\alpha\Sigma+\sigma}$, as the communication policy following $\boldsymbol{\pi}_{\alpha\Sigma}$ for all agents at all times, except for agent i under belief state β, when agent i sends message σ. With this definition, $\boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}$, is the policy under which agent i communicates the achievement of G, and $\boldsymbol{\pi}_{\alpha\Sigma+\text{null}}$ is the policy under which it does not. Therefore, we can alternatively describe agent i's decision criterion as choosing $\boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}$ over $\boldsymbol{\pi}_{\alpha\Sigma+\text{null}}$ if and only if $\Delta^T \geq 0$.

We can use the COM-MTDP model to derive an operational expression of $\Delta^T \geq 0$. For simplicity, we define notational shorthand for various sequences and combinations of values. We define a partial sequence of random variables, $X^{<t}$, to be the sequence of random variables for all times before t: $X^0, X^1, \ldots, X^{t-1}$. We make similar definitions for the other relational operators. The expression, $(S)^T$, denotes the cross product over states of the world, $\prod_{t=0}^T S$, as distinguished from the time-indexed random variable, S^T, which denotes the value of the state at time T. The notation, $s^{\geq t_0}[t]$, specifies the element in slot t within the vector $s^{\geq t_0}$. We define the function, Υ, as shorthand within our probability expressions. It allows us to compactly represent a particular subsequence of world and agent belief states occurring, conditioned on the current situation, as follows:

$$
\Pr(\Upsilon(\langle t, t' \rangle, s, \boldsymbol{\beta}_{\bullet\Sigma})) \equiv \Pr(S^{\geq t, \leq t'} = s, \boldsymbol{b}_{\alpha\bullet\Sigma}^{\geq t, \leq t'} = \boldsymbol{\beta}_{\bullet\Sigma} \,\big|\, K_G = t_0, A_G = i, b_{i \bullet \Sigma}^{t_0} = \beta) \tag{2}
$$

Informally, $\Upsilon(\langle t, t' \rangle, s, \boldsymbol{\beta}_{\bullet\Sigma})$ represents the event that the world and belief states from time t through t' correspond to the specified sequences, s and $\boldsymbol{\beta}_{\bullet\Sigma}$, respectively, conditioned on agent i being the first to know of G's achievement at time t_0 with a belief state, β. We define the function, $\beta_{\Sigma\bullet}$, to map a pre-communication belief state into the post-communication belief state that arises from a communication policy:

$$
\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}, \boldsymbol{\pi}_{\alpha\Sigma}) \equiv \boldsymbol{SE}_{\alpha\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}, \boldsymbol{\pi}_{\alpha\Sigma}(\boldsymbol{\beta}_{\bullet\Sigma})) \tag{3}
$$

This definition of $\beta_{\Sigma\bullet}$ is a well-defined function because of the deterministic nature of the policy, $\boldsymbol{\pi}_{\alpha\Sigma}$, and state-estimator function, $\boldsymbol{SE}_{\alpha\Sigma\bullet}$.

Theorem 7 *If we assume that, upon achievement of G, no communication other than σ_G is possible, then the condition $\Delta^T(t_0, i, \beta) \geq 0$ holds if and only if:*

$$\sum_{s^{\le t_0} \in (S)^{t_0}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\le t_0} \in \boldsymbol{B}_\alpha^{t_0}} \Pr(\Upsilon(\langle 0, t_0 \rangle, s^{\le t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{\le t_0}))$$

$$\cdot \left(\sum_{s^{\ge t_0} \in (S)^{T-t_0+1}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\ge t_0} \in \boldsymbol{B}_\alpha^{T-t_0+1}} \Pr \left(\Upsilon(\langle t_0, T \rangle, s^{\ge t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{\ge t_0}) \,\middle|\, \Sigma_i^{t_0} = \sigma_G, \Upsilon(\langle 0, t_0 \rangle, s^{\le t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{\le t_0}) \right) \right.$$

$$\cdot \sum_{t=t_0}^{T} R_A \left(s^{\ge t_0}[t], \boldsymbol{\pi}_{\alpha A} \left(\beta_{\Sigma\bullet} \left(\boldsymbol{\beta}_{\bullet\Sigma}^{\ge t_0}[t], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G} \right) \right) \right)$$

$$- \sum_{s^{\ge t_0} \in (S)^{T-t_0+1}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\ge t_0} \in \boldsymbol{B}_\alpha^{T-t_0+1}} \Pr \left(\Upsilon(\langle t_0, T \rangle, s^{\ge t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{\ge t_0}) \,\middle|\, \Sigma_i^{t_0} = null, \Upsilon(\langle 0, t_0 \rangle, s^{\le t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{\le t_0}) \right)$$

$$\left. \cdot \sum_{t=t_0}^{T} R_A \left(s^{\ge t_0}[t], \boldsymbol{\pi}_{\alpha A} \left(\beta_{\Sigma\bullet} \left(\boldsymbol{\beta}_{\bullet\Sigma}^{\ge t_0}[t], \boldsymbol{\pi}_{\alpha\Sigma+null} \right) \right) \right) \right)$$

$$\ge - \sum_{s \in G} \sum_{\boldsymbol{\beta} \in \boldsymbol{B}_\alpha} \Pr \left(\Upsilon(\langle t_0, t_0 \rangle, s, \boldsymbol{\beta}) \right) R_\Sigma(s, \sigma_G) \tag{4}$$

Proof: We can rewrite the expectation as an explicit summation over the possible state and belief state sequences:

$$\Delta^T(t_0, i, \beta)$$

$$\equiv E \left[\sum_{t=0}^{T-t_0} R^{t_0+t} \,\middle|\, \Sigma_i^{t_0} = \sigma_G, K_G = t_0, A_G = i, b_{i\bullet\Sigma}^{t_0} = \beta \right]$$

$$- E \left[\sum_{t=0}^{T-t_0} R^{t_0+t} \,\middle|\, \Sigma_i^{t_0} = \text{null}, K_G = t_0, A_G = i, b_{i\bullet\Sigma}^{t_0} = \beta \right] \tag{5}$$

$$= \sum_{s^{\le T} \in (S)^T} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\le T} \in (\mathbf{B})^T} \sum_{\boldsymbol{\beta}_{\Sigma\bullet}^{\le T} \in (\mathbf{B})^T} \Pr \left(S^{\le T} = s^{\le T}, \mathbf{b}_{\bullet\Sigma}^{\le T} = \boldsymbol{\beta}_{\bullet\Sigma}^{\le T}, \mathbf{b}_{\Sigma\bullet}^{\le T} = \boldsymbol{\beta}_{\Sigma\bullet}^{\le T} \right.$$

$$\left. \middle|\, \Sigma_i^{t_0} = \sigma_G, K_G = t_0, A_G = i, b_{i\bullet\Sigma}^{t_0} = \beta \right)$$

$$\cdot \sum_{t=0}^{T} R(s^{\le T}[t], \boldsymbol{\pi}_{\alpha A}(\boldsymbol{\beta}_{\Sigma\bullet}^{\le T}[t]), \boldsymbol{\pi}_{\alpha\Sigma}(\boldsymbol{\beta}_{\bullet\Sigma}^{\le T}[t]))$$

$$- \sum_{s^{\le T} \in (S)^T} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\le T} \in (\mathbf{B})^T} \sum_{\boldsymbol{\beta}_{\Sigma\bullet}^{\le T} \in (\mathbf{B})^T} \Pr \left(S^{\le T} = s^{\le T}, \mathbf{b}_{\bullet\Sigma}^{\le T} = \boldsymbol{\beta}_{\bullet\Sigma}^{\le T}, \mathbf{b}_{\Sigma\bullet}^{\le T} = \boldsymbol{\beta}_{\Sigma\bullet}^{\le T} \right.$$

$$\left. \middle|\, \Sigma_i^{t_0} = \text{null}, K_G = t_0, A_G = i, b_{i\bullet\Sigma}^{t_0} = \beta \right)$$

$$\cdot \sum_{t=0}^{T} R(s^{\le T}[t], \boldsymbol{\pi}_{\alpha A}(\boldsymbol{\beta}_{\Sigma\bullet}^{\le T}[t]), \boldsymbol{\pi}_{\alpha\Sigma}(\boldsymbol{\beta}_{\bullet\Sigma}^{\le T}[t])) \tag{6}$$

We now define the communication policy, $\pi_{\Sigma+\sigma}$, as the communication policy following π_Σ for all agents at all times, except for agent i under belief state β, when agent i sends message σ. With this definition, $\pi_{\Sigma+\sigma_G}$, is the policy under which agent i communicates the achievement of G, and $\pi_{\Sigma+\text{null}}$ is the policy under which it does not. We can simplify this expression further by using our belief state mapping function, $\beta_{\Sigma\bullet}$, because the post-communication belief state, $\mathbf{b}_{\Sigma\bullet}$, is completely determined given the pre-communication belief state, $\mathbf{b}_{\bullet\Sigma}$, the communication policy, π_Σ, and the post-communication state estimator function:

$$
= \sum_{s^{\leq T}\in(S)^T} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq T}\in(\mathbf{B})^T} \Pr(S^{\leq T}=s^{\leq T}, \mathbf{b}_{\bullet\Sigma}^{\leq T}=\boldsymbol{\beta}_{\bullet\Sigma}^{\leq T}|\Sigma_i^{t_0}=\sigma_G, K_G=t_0, A_G=i, b_{i\bullet\Sigma}^{t_0}=\beta)
$$

$$
\cdot \sum_{t=0}^{T} R(s^{\leq T}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq T}[t], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G})), \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq T}[t]))
$$

$$
- \sum_{s^{\leq T}\in(S)^T} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq T}\in(\mathbf{B})^T} \Pr(S^{\leq T}=s^{\leq T}, \mathbf{b}_{\bullet\Sigma}^{\leq T}=\boldsymbol{\beta}_{\bullet\Sigma}^{\leq T}|\Sigma_i^{t_0}=\text{null}, K_G=t_0, A_G=i, b_{i\bullet\Sigma}^{t_0}=\beta)
$$

$$
\cdot \sum_{t=0}^{T} R(s^{\leq T}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq T}[t], \boldsymbol{\pi}_{\alpha\Sigma+\text{null}})), \boldsymbol{\pi}_{\alpha\Sigma+\text{null}}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq T}[t])) \tag{7}
$$

We can further simplify the exposition by using our Υ shorthand:

$$
= \sum_{s^{\leq T}\in(S)^T} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq T}\in(\mathbf{B})^T} \Pr(\Upsilon(\langle 0,T\rangle, s, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq T})|\Sigma_i^{t_0}=\sigma_G)
$$

$$
\cdot \sum_{t=0}^{T} R(s^{\leq T}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq T}[t], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G})), \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq T}[t]))
$$

$$
- \sum_{s^{\leq T}\in(S)^T} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq T}\in(\mathbf{B})^T} \Pr(\Upsilon(\langle 0,T\rangle, s, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq T})|\Sigma_i^{t_0}=\text{null})
$$

$$
\cdot \sum_{t=0}^{T} R(s^{\leq T}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq T}[t], \boldsymbol{\pi}_{\alpha\Sigma+\text{null}})), \boldsymbol{\pi}_{\alpha\Sigma+\text{null}}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq T}[t])) \tag{8}
$$

The innermost sum adds up the reward received at each point of time. We can move the innermost summations to the outside:

$$
= \sum_{t=0}^{T} \sum_{s^{\leq T}\in(S)^T} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq T}\in(\mathbf{B})^T} \Pr(\Upsilon(\langle 0,T\rangle, s, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq T})|\Sigma_i^{t_0}=\sigma_G)
$$

$$
\cdot R(s^{\leq T}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq T}[t], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G})), \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq T}[t]))
$$

$$
- \sum_{t=0}^{T} \sum_{s^{\leq T}\in(S)^T} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq T}\in(\mathbf{B})^T} \Pr(\Upsilon(\langle 0,T\rangle, s, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq T})|\Sigma_i^{t_0}=\text{null})
$$

$$
\cdot R(s^{\leq T}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq T}[t], \boldsymbol{\pi}_{\alpha\Sigma+\text{null}})), \boldsymbol{\pi}_{\alpha\Sigma+\text{null}}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq T}[t])) \tag{9}
$$

We can split the inner two summations at time t as follows:

$$= \sum_{t=0}^{T} \sum_{s^{\leq t} \in (S)^t} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t} \in (\mathbf{B})^t} \Pr(\Upsilon(\langle 0, t\rangle, s^{\leq t}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t})|\Sigma_i^{t_0} = \sigma_G)$$

$$\cdot \sum_{s^{>t} \in (S)^{T-t}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{>t} \in (\mathbf{B})^{T-t}} \Pr(\Upsilon(\langle t+1, T\rangle, s^{>t}, \boldsymbol{\beta}_{\bullet\Sigma}^{>t})|\Upsilon(\langle 0, t\rangle, s^{\leq t}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}), \Sigma_i^{t_0} = \sigma_G)$$

$$\cdot R(s^{\leq t}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}[t], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G})), \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}[t]))$$

$$- \sum_{t=0}^{T} \sum_{s^{\leq t} \in (S)^t} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t} \in (\mathbf{B})^t} \Pr(\Upsilon(\langle 0, t\rangle, s^{\leq t}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t})|\Sigma_i^{t_0} = \text{null})$$

$$\cdot \sum_{s^{>t} \in (S)^{T-t}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{>t} \in (\mathbf{B})^{T-t}} \Pr(\Upsilon(\langle t+1, T\rangle, s^{>t}, \boldsymbol{\beta}_{\bullet\Sigma}^{>t})|\Upsilon(\langle 0, t\rangle, s^{\leq t}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}), \Sigma_i^{t_0} = \text{null})$$

$$\cdot R(s^{\leq t}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}[t], \boldsymbol{\pi}_{\alpha\Sigma+\text{null}})), \boldsymbol{\pi}_{\alpha\Sigma+\text{null}}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}[t])) \qquad (10)$$

By our Markovian assumption, for any given point of time, t, the reward is independent of the subsequence of states and observations to follow time t. Thus, the inner two summations reduce to 1.0:

$$= \sum_{t=0}^{T} \sum_{s^{\leq t} \in (S)^t} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t} \in (\mathbf{B})^t} \Pr(\Upsilon(\langle 0, t\rangle, s^{\leq t}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t})|\Sigma_i^{t_0} = \sigma_G)$$

$$\cdot R(s^{\leq t}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}[t], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G})), \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}[t]))$$

$$- \sum_{t=0}^{T} \sum_{s^{\leq t} \in (S)^t} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t} \in (\mathbf{B})^t} \Pr(\Upsilon(\langle 0, t\rangle, s^{\leq t}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t})|\Sigma_i^{t_0} = \text{null})$$

$$\cdot R(s^{\leq t}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}[t], \boldsymbol{\pi}_{\alpha\Sigma+\text{null}})), \boldsymbol{\pi}_{\alpha\Sigma+\text{null}}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}[t])) \qquad (11)$$

We can split the summation at the time, t_0, when agent i knows that G has been achieved:

$$= \sum_{t=0}^{t_0-1} \sum_{s^{\leq t} \in (S)^t} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t} \in (\mathbf{B})^t} \Pr(\Upsilon(\langle 0, t \rangle, s^{\leq t}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}) | \Sigma_i^{t_0} = \sigma_G)$$

$$\cdot R(s^{\leq t}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}[t], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G})), \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}[t]))$$

$$+ \sum_{t=t_0}^{T} \sum_{s^{\leq t} \in (S)^t} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t} \in (\mathbf{B})^t} \Pr(\Upsilon(\langle 0, t \rangle, s^{\leq t}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}) | \Sigma_i^{t_0} = \sigma_G)$$

$$\cdot R(s^{\leq t}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}[t], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G})), \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}[t]))$$

$$- \sum_{t=0}^{t_0-1} \sum_{s^{\leq t} \in (S)^t} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t} \in (\mathbf{B})^t} \Pr(\Upsilon(\langle 0, t \rangle, s^{\leq t}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}) | \Sigma_i^{t_0} = \text{null})$$

$$\cdot R(s^{\leq t}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}[t], \boldsymbol{\pi}_{\alpha\Sigma+\text{null}})), \boldsymbol{\pi}_{\alpha\Sigma+\text{null}}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}[t]))$$

$$- \sum_{t=t_0}^{T} \sum_{s^{\leq t} \in (S)^t} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t} \in (\mathbf{B})^t} \Pr(\Upsilon(\langle 0, t \rangle, s^{\leq t}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}) | \Sigma_i^{t_0} = \text{null})$$

$$\cdot R(s^{\leq t}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}[t], \boldsymbol{\pi}_{\alpha\Sigma+\text{null}})), \boldsymbol{\pi}_{\alpha\Sigma+\text{null}}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}[t])) \tag{12}$$

The first and third terms cancel each other out, because the message sent at time t_0 has no effect on the rewards received before that time:

$$= \sum_{t=t_0}^{T} \sum_{s^{\leq t} \in (S)^t} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t} \in (\mathbf{B})^t} \Pr(\Upsilon(\langle 0, t \rangle, s^{\leq t}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}) | \Sigma_i^{t_0} = \sigma_G)$$

$$\cdot R(s^{\leq t}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}[t], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G})), \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}[t]))$$

$$- \sum_{t=t_0}^{T} \sum_{s^{\leq t} \in (S)^t} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t} \in (\mathbf{B})^t} \Pr(\Upsilon(\langle 0, t \rangle, s^{\leq t}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}) | \Sigma_i^{t_0} = \text{null})$$

$$\cdot R(s^{\leq t}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}[t], \boldsymbol{\pi}_{\alpha\Sigma+\text{null}})), \boldsymbol{\pi}_{\alpha\Sigma+\text{null}}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t}[t])) \tag{13}$$

We first separate out the case at time t_0:

$$= \sum_{s^{\le t_0} \in (S)^{t_0}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}{}^{\le t_0} \in (\mathbf{B})^{t_0}} \Pr(\Upsilon(\langle 0, t_0 \rangle, s^{\le t_0}, \boldsymbol{\beta}_{\bullet\Sigma}{}^{\le t_0}) | \Sigma_i^{t_0} = \sigma_G)$$

$$\cdot R(s^{\le t_0}[t_0], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}{}^{\le t_0}[t_0], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G})), \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}(\boldsymbol{\beta}_{\bullet\Sigma}{}^{\le t_0}[t_0]))$$

$$+ \sum_{t=t_0+1}^{T} \sum_{s^{\le t} \in (S)^{t}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}{}^{\le t} \in (\mathbf{B})^{t}} \Pr(\Upsilon(\langle 0, t \rangle, s^{\le t}, \boldsymbol{\beta}_{\bullet\Sigma}{}^{\le t}) | \Sigma_i^{t_0} = \sigma_G)$$

$$\cdot R(s^{\le t}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}{}^{\le t}[t], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G})), \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}(\boldsymbol{\beta}_{\bullet\Sigma}{}^{\le t}[t]))$$

$$- \sum_{s^{\le t_0} \in (S)^{t_0}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}{}^{\le t_0} \in (\mathbf{B})^{t_0}} \Pr(\Upsilon(\langle 0, t_0 \rangle, s^{\le t_0}, \boldsymbol{\beta}_{\bullet\Sigma}{}^{\le t_0}) | \Sigma_i^{t_0} = \mathbf{null})$$

$$\cdot R(s^{\le t_0}[t_0], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}{}^{\le t_0}[t_0], \boldsymbol{\pi}_{\alpha\Sigma+\mathbf{null}})), \boldsymbol{\pi}_{\alpha\Sigma+\mathbf{null}}(\boldsymbol{\beta}_{\bullet\Sigma}{}^{\le t_0}[t_0]))$$

$$- \sum_{t=t_0+1}^{T} \sum_{s^{\le t} \in (S)^{t}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}{}^{\le t} \in (\mathbf{B})^{t}} \Pr(\Upsilon(\langle 0, t \rangle, s^{\le t}, \boldsymbol{\beta}_{\bullet\Sigma}{}^{\le t}) | \Sigma_i^{t_0} = \mathbf{null})$$

$$\cdot R(s^{\le t}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}{}^{\le t}[t], \boldsymbol{\pi}_{\alpha\Sigma+\mathbf{null}})), \boldsymbol{\pi}_{\alpha\Sigma+\mathbf{null}}(\boldsymbol{\beta}_{\bullet\Sigma}{}^{\le t}[t])) \tag{14}$$

We can split the second and fourth terms at time t_0:

$$= \sum_{s^{\leq t_0} \in (S)^{t_0}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0} \in (\mathbf{B})^{t_0}} \Pr(\Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0}) | \Sigma_i^{t_0} = \sigma_G)$$

$$\cdot R(s^{\leq t_0}[t_0], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0}[t_0], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G})), \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0}[t_0]))$$

$$+ \sum_{t=t_0+1}^{T} \sum_{s^{\leq t_0} \in (S)^{t_0}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0} \in (\mathbf{B})^{t_0}} \Pr(\Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0}) | \Sigma_i^{t_0} = \sigma_G)$$

$$\cdot \sum_{s^{>t_0} \in (S)^{T-t_0}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{>t_0} \in (\mathbf{B})^{T-t_0}} \Pr(\Upsilon(\langle t_0+1, T \rangle, s^{>t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{>t_0})$$

$$| \Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0}), \Sigma_i^{t_0} = \sigma_G)$$

$$\cdot R(s^{>t_0}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{>t_0}[t], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G})), \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}(\boldsymbol{\beta}_{\bullet\Sigma}^{>t_0}[t]))$$

$$- \sum_{s^{\leq t_0} \in (S)^{t_0}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0} \in (\mathbf{B})^{t_0}} \Pr(\Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0}) | \Sigma_i^{t_0} = \text{null})$$

$$\cdot R(s^{\leq t_0}[t_0], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0}[t_0], \boldsymbol{\pi}_{\alpha\Sigma+\text{null}})), \boldsymbol{\pi}_{\alpha\Sigma+\text{null}}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0}[t_0]))$$

$$- \sum_{t=t_0+1}^{T} \sum_{s^{\leq t_0} \in (S)^{t_0}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0} \in (\mathbf{B})^{t_0}} \Pr(\Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0}) | \Sigma_i^{t_0} = \text{null})$$

$$\cdot \sum_{s^{>t_0} \in (S)^{T-t_0}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{>t_0} \in (\mathbf{B})^{T-t_0}} \Pr(\Upsilon(\langle t_0+1, T \rangle, s^{>t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{>t_0})$$

$$| \Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0}), \Sigma_i^{t_0} = \text{null})$$

$$\cdot R(s^{>t_0}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{>t_0}[t], \boldsymbol{\pi}_{\alpha\Sigma+\text{null}})), \boldsymbol{\pi}_{\alpha\Sigma+\text{null}}(\boldsymbol{\beta}_{\bullet\Sigma}^{>t_0}[t]))$$

$$(15)$$

In addition, the outermost probability distribution is the same for the two policies, since they do not deviate until G is known by some agent. Therefore, we can group the first and third terms and the second and fourth terms together as follows:

$$
= \sum_{s^{\leq t_0} \in (S)^{t_0}} \sum_{\boldsymbol{\beta_{\bullet\Sigma}}^{\leq t_0} \in (\mathbf{B})^{t_0}} \Pr(\Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{\leq t_0}) | \Sigma_i^{t_0} = \sigma_G)
$$

$$
\cdot \Big[R(s^{\leq t_0}[t_0], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta_{\bullet\Sigma}}^{\leq t_0}[t_0], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G})), \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}(\boldsymbol{\beta_{\bullet\Sigma}}^{\leq t_0}[t_0]))
$$

$$
- R(s^{\leq t_0}[t_0], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta_{\bullet\Sigma}}^{\leq t_0}[t_0], \boldsymbol{\pi}_{\alpha\Sigma+\mathrm{null}})), \boldsymbol{\pi}_{\alpha\Sigma+\mathrm{null}}(\boldsymbol{\beta_{\bullet\Sigma}}^{\leq t_0}[t_0])) \Big]
$$

$$
+ \sum_{t=t_0+1}^{T} \sum_{s^{\leq t_0} \in (S)^{t_0}} \sum_{\boldsymbol{\beta_{\bullet\Sigma}}^{\leq t_0} \in (\mathbf{B})^{t_0}} \Pr(\Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{\leq t_0}) | \Sigma_i^{t_0} = \sigma_G)
$$

$$
\cdot \Bigg[\sum_{s^{>t_0} \in (S)^{T-t_0}} \sum_{\boldsymbol{\beta_{\bullet\Sigma}}^{>t_0} \in (\mathbf{B})^{T-t_0}} \Pr(\Upsilon(\langle t_0+1, T \rangle, s^{>t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{>t_0})
$$

$$
|\Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{\leq t_0}), \Sigma_i^{t_0} = \sigma_G)
$$

$$
\cdot R(s^{>t_0}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta_{\bullet\Sigma}}^{>t_0}[t], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G})), \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}(\boldsymbol{\beta_{\bullet\Sigma}}^{>t_0}[t]))
$$

$$
- \sum_{s^{>t_0} \in (S)^{T-t_0}} \sum_{\boldsymbol{\beta_{\bullet\Sigma}}^{>t_0} \in (\mathbf{B})^{T-t_0}} \Pr(\Upsilon(\langle t_0+1, T \rangle, s^{>t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{>t_0})
$$

$$
|\Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{\leq t_0}), \Sigma_i^{t_0} = \mathrm{null})
$$

$$
\cdot R(s^{>t_0}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta_{\bullet\Sigma}}^{>t_0}[t], \boldsymbol{\pi}_{\alpha\Sigma+\mathrm{null}})), \boldsymbol{\pi}_{\alpha\Sigma+\mathrm{null}}(\boldsymbol{\beta_{\bullet\Sigma}}^{>t_0}[t])) \Bigg]
$$

$$
(16)
$$

The differences within the brackets measure the expected difference in reward over all possible future sequences of states and observations. The probability in the outer summation iterates over all possible past sequences. We now separate the domain-level and communication-level components of the reward function:

$$
= \sum_{s^{\leq t_0} \in (S)^{t_0}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0} \in (\mathbf{B})^{t_0}} \Pr(\Upsilon(\langle 0, t_0\rangle, s^{\leq t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0}) | \Sigma_i^{t_0} = \sigma_G)
$$

$$
\cdot \Big[R_A(s^{\leq t_0}[t_0], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0}[t_0], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}))) + R_\Sigma(s^{\leq t_0}[t_0], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0}[t_0]))
$$

$$
- R_A(s^{\leq t_0}[t_0], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0}[t_0], \boldsymbol{\pi}_{\alpha\Sigma+\mathbf{null}}))) - R_\Sigma(s^{\leq t_0}[t_0], \boldsymbol{\pi}_{\alpha\Sigma+\mathbf{null}}(\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0}[t_0])) \Big]
$$

$$
+ \sum_{t=t_0+1}^{T} \sum_{s^{\leq t_0} \in (S)^{t_0}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0} \in (\mathbf{B})^{t_0}} \Pr(\Upsilon(\langle 0, t_0\rangle, s^{\leq t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0}) | \Sigma_i^{t_0} = \sigma_G)
$$

$$
\cdot \Bigg[\sum_{s^{>t_0} \in (S)^{T-t_0}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{>t_0} \in (\mathbf{B})^{T-t_0}} \Pr(\Upsilon(\langle t_0+1, T\rangle, s^{>t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{>t_0})
$$

$$
| \Upsilon(\langle 0, t_0\rangle, s^{\leq t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0}), \Sigma_i^{t_0} = \sigma_G)
$$

$$
\cdot \Big[R_A(s^{>t_0}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{>t_0}[t], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}))) + R_\Sigma(s^{>t_0}[t], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}(\boldsymbol{\beta}_{\bullet\Sigma}^{>t_0}[t])) \Big]
$$

$$
- \sum_{s^{>t_0} \in (S)^{T-t_0}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{>t_0} \in (\mathbf{B})^{T-t_0}} \Pr(\Upsilon(\langle t_0+1, T\rangle, s^{>t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{>t_0})
$$

$$
| \Upsilon(\langle 0, t_0\rangle, s^{\leq t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{\leq t_0}), \Sigma_i^{t_0} = \mathbf{null})
$$

$$
\cdot \Big[R_A(s^{>t_0}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{>t_0}[t], \boldsymbol{\pi}_{\alpha\Sigma+\mathbf{null}}))) + R_\Sigma(s^{>t_0}[t], \boldsymbol{\pi}_{\alpha\Sigma+\mathbf{null}}(\boldsymbol{\beta}_{\bullet\Sigma}^{>t_0}[t])) \Big] \Bigg]
$$

$$
\tag{17}
$$

In the first term, we can isolate the difference between the communication-level reward values for te two policies. We define $\Sigma_{\neq i}^{t}$ to be the combined communication of all agents other than i, at time t. At time t_0, this variable is the same for both policies, since only the policies differ when communicating σ_G. Thus, we can reduce our expression to:

$$= \sum_{s^{\leq t_0} \in (S)^{t_0}} \sum_{\boldsymbol{\beta_{\bullet \Sigma}}^{\leq t_0} \in (\mathbf{B})^{t_0}} \Pr(\Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta_{\bullet \Sigma}}^{\leq t_0}) | \Sigma_i^{t_0} = \sigma_G)$$

$$\cdot \left[R_\Sigma(s^{\leq t_0}[t_0], \sigma_G \times \boldsymbol{\Sigma}_{\neq i}^t) - R_\Sigma(s^{\leq t_0}[t_0], \boldsymbol{\Sigma}_{\neq i}^t) \right]$$

$$+ \sum_{s^{\leq t_0} \in (S)^{t_0}} \sum_{\boldsymbol{\beta_{\bullet \Sigma}}^{\leq t_0} \in (\mathbf{B})^{t_0}} \Pr(\Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta_{\bullet \Sigma}}^{\leq t_0}) | \Sigma_i^{t_0} = \sigma_G)$$

$$\cdot \left[R_A(s^{\leq t_0}[t_0], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma \bullet}(\boldsymbol{\beta_{\bullet \Sigma}}^{\leq t_0}[t_0], \boldsymbol{\pi}_{\alpha \Sigma + \sigma_G}))) - R_A(s^{\leq t_0}[t_0], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma \bullet}(\boldsymbol{\beta_{\bullet \Sigma}}^{\leq t_0}[t_0], \boldsymbol{\pi}_{\alpha \Sigma + \text{null}}))) \right]$$

$$+ \sum_{t=t_0+1}^{T} \sum_{s^{\leq t_0} \in (S)^{t_0}} \sum_{\boldsymbol{\beta_{\bullet \Sigma}}^{\leq t_0} \in (\mathbf{B})^{t_0}} \Pr(\Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta_{\bullet \Sigma}}^{\leq t_0}) | \Sigma_i^{t_0} = \sigma_G)$$

$$\cdot \left[\sum_{s^{>t_0} \in (S)^{T-t_0}} \sum_{\boldsymbol{\beta_{\bullet \Sigma}}^{>t_0} \in (\mathbf{B})^{T-t_0}} \Pr(\Upsilon(\langle t_0 + 1, T \rangle, s^{>t_0}, \boldsymbol{\beta_{\bullet \Sigma}}^{>t_0}) \right.$$

$$| \Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta_{\bullet \Sigma}}^{\leq t_0}), \Sigma_i^{t_0} = \sigma_G)$$

$$\cdot R_A(s^{>t_0}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma \bullet}(\boldsymbol{\beta_{\bullet \Sigma}}^{>t_0}[t], \boldsymbol{\pi}_{\alpha \Sigma + \sigma_G}))) + R_\Sigma(s^{>t_0}[t], \boldsymbol{\pi}_{\alpha \Sigma + \sigma_G}(\boldsymbol{\beta_{\bullet \Sigma}}^{>t_0}[t]))$$

$$- \sum_{s^{>t_0} \in (S)^{T-t_0}} \sum_{\boldsymbol{\beta_{\bullet \Sigma}}^{>t_0} \in (\mathbf{B})^{T-t_0}} \Pr(\Upsilon(\langle t_0 + 1, T \rangle, s^{>t_0}, \boldsymbol{\beta_{\bullet \Sigma}}^{>t_0})$$

$$| \Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta_{\bullet \Sigma}}^{\leq t_0}), \Sigma_i^{t_0} = \text{null})$$

$$\left. \cdot \left[R_A(s^{>t_0}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma \bullet}(\boldsymbol{\beta_{\bullet \Sigma}}^{>t_0}[t], \boldsymbol{\pi}_{\alpha \Sigma + \text{null}}))) + R_\Sigma(s^{>t_0}[t], \boldsymbol{\pi}_{\alpha \Sigma + \text{null}}(\boldsymbol{\beta_{\bullet \Sigma}}^{>t_0}[t])) \right] \right]$$

$$\tag{18}$$

We can separate the third term into domain-level and communication-level sums. We can then merge the second term into the domain-level sum by rewriting this sum to include the t_0 case again:

$$= \sum_{s^{\leq t_0} \in (S)^{t_0}} \sum_{\boldsymbol{\beta_{\bullet\Sigma}}^{\leq t_0} \in (\mathbf{B})^{t_0}} \mathrm{Pr}(\Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{\leq t_0}) | \Sigma_i^{t_0} = \sigma_G)$$

$$\cdot \left[R_\Sigma(s^{\leq t_0}[t_0], \sigma_G \times \boldsymbol{\Sigma}_{\neq i}^t) - R_\Sigma(s^{\leq t_0}[t_0], \boldsymbol{\Sigma}_{\neq i}^t) \right]$$

$$+ \sum_{t=t_0}^{T} \sum_{s^{<t_0} \in (S)^{t_0-1}} \sum_{\boldsymbol{\beta_{\bullet\Sigma}}^{<t_0} \in (\mathbf{B})^{t_0-1}} \mathrm{Pr}(\Upsilon(\langle 0, t_0-1 \rangle, s^{<t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{<t_0}))$$

$$\cdot \left[\sum_{s^{\geq t_0} \in (S)^{T-t_0+1}} \sum_{\boldsymbol{\beta_{\bullet\Sigma}}^{\geq t_0} \in (\mathbf{B})^{T-t_0+1}} \mathrm{Pr}(\Upsilon(\langle t_0, T \rangle, s^{\geq t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{\geq t_0}) | \Upsilon(\langle 0, t_0-1 \rangle, s^{<t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{<t_0}), \Sigma_i^{t_0} = \sigma_G) \right.$$

$$\cdot R_A(s^{\geq t_0}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta_{\bullet\Sigma}}^{>t_0}[t], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G})))$$

$$- \sum_{s^{\geq t_0} \in (S)^{T-t_0+1}} \sum_{\boldsymbol{\beta_{\bullet\Sigma}}^{\geq t_0} \in (\mathbf{B})^{T-t_0+1}} \mathrm{Pr}(\Upsilon(\langle t_0, T \rangle, s^{\geq t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{\geq t_0}) | \Upsilon(\langle 0, t_0-1 \rangle, s^{<t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{<t_0}), \Sigma_i^{t_0} = \text{null})$$

$$\left. \cdot R_A(s^{\geq t_0}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta_{\bullet\Sigma}}^{>t_0}[t], \boldsymbol{\pi}_{\alpha\Sigma+\text{null}}))) \right]$$

$$+ \sum_{t=t_0+1}^{T} \sum_{s^{\leq t_0} \in (S)^{t_0}} \sum_{\boldsymbol{\beta_{\bullet\Sigma}}^{\leq t_0} \in (\mathbf{B})^{t_0}} \mathrm{Pr}(\Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{\leq t_0}) | \Sigma_i^{t_0} = \sigma_G)$$

$$\cdot \left[\sum_{s^{>t_0} \in (S)^{T-t_0}} \sum_{\boldsymbol{\beta_{\bullet\Sigma}}^{>t_0} \in (\mathbf{B})^{T-t_0}} \mathrm{Pr}(\Upsilon(\langle t_0+1, T \rangle, s^{>t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{>t_0}) | \Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{\leq t_0}), \Sigma_i^{t_0} = \sigma_G) \right.$$

$$\cdot R_\Sigma(s^{>t_0}[t], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}(\boldsymbol{\beta_{\bullet\Sigma}}^{>t_0}[t]))$$

$$- \sum_{s^{>t_0} \in (S)^{T-t_0}} \sum_{\boldsymbol{\beta_{\bullet\Sigma}}^{>t_0} \in (\mathbf{B})^{T-t_0}} \mathrm{Pr}(\Upsilon(\langle t_0+1, T \rangle, s^{>t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{>t_0}) | \Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{\leq t_0}), \Sigma_i^{t_0} = \text{null})$$

$$\left. \cdot R_\Sigma(s^{>t_0}[t], \boldsymbol{\pi}_{\alpha\Sigma+\text{null}}(\boldsymbol{\beta_{\bullet\Sigma}}^{>t_0}[t])) \right] \tag{19}$$

The Markovian property of the reward function allows us to further reduce the first term. In the second and third terms, we can move the outermost summation back inside:

$$= \sum_{s \in G} \sum_{\boldsymbol{\beta} \in \boldsymbol{B}} \Pr(\Upsilon(\langle t_0, t_0 \rangle, s, \boldsymbol{\beta})) \cdot \left[R_\Sigma(s, \sigma_G \times \boldsymbol{\Sigma}_{\neq i}^t) - R_\Sigma(s, \boldsymbol{\Sigma}_{\neq i}^t) \right]$$

$$+ \sum_{s^{<t_0} \in (S)^{t_0-1}} \sum_{\boldsymbol{\beta_{\bullet\Sigma}}^{<t_0} \in (\mathbf{B})^{t_0-1}} \Pr(\Upsilon(\langle 0, t_0-1 \rangle, s^{<t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{<t_0}))$$

$$\cdot \left[\sum_{s^{\geq t_0} \in (S)^{T-t_0+1}} \sum_{\boldsymbol{\beta_{\bullet\Sigma}}^{\geq t_0} \in (\mathbf{B})^{T-t_0+1}} \sum_{t=t_0}^{T} \Pr(\Upsilon(\langle t_0, T \rangle, s^{\geq t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{\geq t_0}) | \Upsilon(\langle 0, t_0-1 \rangle, s^{<t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{<t_0}), \Sigma_i^{t_0} = \sigma_G) \right.$$

$$\cdot R_A(s^{\geq t_0}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta_{\bullet\Sigma}}^{>t_0}[t], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G})))$$

$$- \sum_{s^{\geq t_0} \in (S)^{T-t_0+1}} \sum_{\boldsymbol{\beta_{\bullet\Sigma}}^{\geq t_0} \in (\mathbf{B})^{T-t_0+1}} \sum_{t=t_0}^{T} \Pr(\Upsilon(\langle t_0, T \rangle, s^{\geq t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{\geq t_0}) | \Upsilon(\langle 0, t_0-1 \rangle, s^{<t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{<t_0}), \Sigma_i^{t_0} = \text{null})$$

$$\left. \cdot R_A(s^{\geq t_0}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta_{\bullet\Sigma}}^{>t_0}[t], \boldsymbol{\pi}_{\alpha\Sigma+\text{null}}))) \right]$$

$$+ \sum_{s^{\leq t_0} \in (S)^{t_0}} \sum_{\boldsymbol{\beta_{\bullet\Sigma}}^{\leq t_0} \in (\mathbf{B})^{t_0}} \Pr(\Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{\leq t_0}) | \Sigma_i^{t_0} = \sigma_G)$$

$$\cdot \left[\sum_{s^{>t_0} \in (S)^{T-t_0}} \sum_{\boldsymbol{\beta_{\bullet\Sigma}}^{>t_0} \in (\mathbf{B})^{T-t_0}} \sum_{t=t_0+1}^{T} \Pr(\Upsilon(\langle t_0+1, T \rangle, s^{>t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{>t_0}) | \Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{\leq t_0}), \Sigma_i^{t_0} = \sigma_G) \right.$$

$$\cdot R_\Sigma(s^{>t_0}[t], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G}(\boldsymbol{\beta_{\bullet\Sigma}}^{>t_0}[t]))$$

$$- \sum_{s^{>t_0} \in (S)^{T-t_0}} \sum_{\boldsymbol{\beta_{\bullet\Sigma}}^{>t_0} \in (\mathbf{B})^{T-t_0}} \sum_{t=t_0+1}^{T} \Pr(\Upsilon(\langle t_0+1, T \rangle, s^{>t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{>t_0}) | \Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta_{\bullet\Sigma}}^{\leq t_0}), \Sigma_i^{t_0} = \text{null})$$

$$\left. \cdot R_\Sigma(s^{>t_0}[t], \boldsymbol{\pi}_{\alpha\Sigma+\text{null}}(\boldsymbol{\beta_{\bullet\Sigma}}^{>t_0}[t])) \right] \tag{20}$$

12

Given the sincerity of our agents, we can restrict the summation of the first term to consider only those states in which the agents have truly achieved the goal, G:

$$= \sum_{s \in G} \sum_{\boldsymbol{\beta} \in \boldsymbol{B}} \Pr(\Upsilon(\langle t_0, t_0 \rangle, s, \boldsymbol{\beta})) \cdot \left[R_\Sigma(s, \sigma_G \times \boldsymbol{\Sigma}_{\neq i}^t) - R_\Sigma(s, \boldsymbol{\Sigma}_{\neq i}^t) \right]$$

$$+ \sum_{s^{<t_0} \in (S)^{t_0-1}} \sum_{\boldsymbol{\beta_{\bullet \Sigma}}^{<t_0} \in (\mathbf{B})^{t_0-1}} \Pr(\Upsilon(\langle 0, t_0 - 1 \rangle, s^{<t_0}, \boldsymbol{\beta_{\bullet \Sigma}}^{<t_0}))$$

$$\cdot \left[\sum_{s^{\geq t_0} \in (S)^{T-t_0+1}} \sum_{\boldsymbol{\beta_{\bullet \Sigma}}^{\geq t_0} \in (\mathbf{B})^{T-t_0+1}} \sum_{t=t_0}^{T} \Pr(\Upsilon(\langle t_0, T \rangle, s^{\geq t_0}, \boldsymbol{\beta_{\bullet \Sigma}}^{\geq t_0}) | \Upsilon(\langle 0, t_0 - 1 \rangle, s^{<t_0}, \boldsymbol{\beta_{\bullet \Sigma}}^{<t_0}), \Sigma_i^{t_0} = \sigma_G) \right.$$

$$\cdot R_A(s^{\geq t_0}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma \bullet}(\boldsymbol{\beta_{\bullet \Sigma}}^{>t_0}[t], \boldsymbol{\pi}_{\alpha \Sigma + \sigma_G})))$$

$$- \sum_{s^{\geq t_0} \in (S)^{T-t_0+1}} \sum_{\boldsymbol{\beta_{\bullet \Sigma}}^{\geq t_0} \in (\mathbf{B})^{T-t_0+1}} \sum_{t=t_0}^{T} \Pr(\Upsilon(\langle t_0, T \rangle, s^{\geq t_0}, \boldsymbol{\beta_{\bullet \Sigma}}^{\geq t_0}) | \Upsilon(\langle 0, t_0 - 1 \rangle, s^{<t_0}, \boldsymbol{\beta_{\bullet \Sigma}}^{<t_0}), \Sigma_i^{t_0} = \text{null})$$

$$\left. \cdot R_A(s^{\geq t_0}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma \bullet}(\boldsymbol{\beta_{\bullet \Sigma}}^{>t_0}[t], \boldsymbol{\pi}_{\alpha \Sigma + \text{null}}))) \right]$$

$$+ \sum_{s^{\leq t_0} \in (S)^{t_0}} \sum_{\boldsymbol{\beta_{\bullet \Sigma}}^{\leq t_0} \in (\mathbf{B})^{t_0}} \Pr(\Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta_{\bullet \Sigma}}^{\leq t_0}) | \Sigma_i^{t_0} = \sigma_G)$$

$$\cdot \left[\sum_{s^{>t_0} \in (S)^{T-t_0}} \sum_{\boldsymbol{\beta_{\bullet \Sigma}}^{>t_0} \in (\mathbf{B})^{T-t_0}} \sum_{t=t_0+1}^{T} \Pr(\Upsilon(\langle t_0 + 1, T \rangle, s^{>t_0}, \boldsymbol{\beta_{\bullet \Sigma}}^{>t_0}) | \Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta_{\bullet \Sigma}}^{\leq t_0}), \Sigma_i^{t_0} = \sigma_G) \right.$$

$$\cdot R_\Sigma(s^{>t_0}[t], \boldsymbol{\pi}_{\alpha \Sigma + \sigma_G}(\boldsymbol{\beta_{\bullet \Sigma}}^{>t_0}[t]))$$

$$- \sum_{s^{>t_0} \in (S)^{T-t_0}} \sum_{\boldsymbol{\beta_{\bullet \Sigma}}^{>t_0} \in (\mathbf{B})^{T-t_0}} \sum_{t=t_0+1}^{T} \Pr(\Upsilon(\langle t_0 + 1, T \rangle, s^{>t_0}, \boldsymbol{\beta_{\bullet \Sigma}}^{>t_0}) | \Upsilon(\langle 0, t_0 \rangle, s^{\leq t_0}, \boldsymbol{\beta_{\bullet \Sigma}}^{\leq t_0}), \Sigma_i^{t_0} = \text{null})$$

$$\left. \cdot R_\Sigma(s^{>t_0}[t], \boldsymbol{\pi}_{\alpha \Sigma + \text{null}}(\boldsymbol{\beta_{\bullet \Sigma}}^{>t_0}[t])) \right] \tag{21}$$

We can further reduce the communication cost expression if we assume that the communication cost of each individual agent's message is independent of the messages communicated by the other agents. By our assumption that no other communication, other than σ_G, is possible upon the achievement of G, then there is no difference in communication cost after time t_0. Thus, the third term reduces to zero, leaving us with the following:

$$
= \sum_{s \in G} \sum_{\boldsymbol{\beta} \in \boldsymbol{B}} \Pr(\Upsilon(\langle t_0, t_0 \rangle, s, \boldsymbol{\beta})) \cdot R_\Sigma(s, \sigma_G)
$$

$$
+ \sum_{s^{<t_0} \in (S)^{t_0-1}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{<t_0} \in (\mathbf{B})^{t_0-1}} \Pr(\Upsilon(\langle 0, t_0 - 1 \rangle, s^{<t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{<t_0}))
$$

$$
\cdot \left[\sum_{s^{\geq t_0} \in (S)^{T-t_0+1}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\geq t_0} \in (\mathbf{B})^{T-t_0+1}} \sum_{t=t_0}^{T} \Pr(\Upsilon(\langle t_0, T \rangle, s^{\geq t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{\geq t_0}) | \Upsilon(\langle 0, t_0 - 1 \rangle, s^{<t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{<t_0}), \Sigma_i^{t_0} = \sigma_G) \right.
$$

$$
\cdot R_A(s^{\geq t_0}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{>t_0}[t], \boldsymbol{\pi}_{\alpha\Sigma+\sigma_G})))
$$

$$
- \sum_{s^{\geq t_0} \in (S)^{T-t_0+1}} \sum_{\boldsymbol{\beta}_{\bullet\Sigma}^{\geq t_0} \in (\mathbf{B})^{T-t_0+1}} \sum_{t=t_0}^{T} \Pr(\Upsilon(\langle t_0, T \rangle, s^{\geq t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{\geq t_0}) | \Upsilon(\langle 0, t_0 - 1 \rangle, s^{<t_0}, \boldsymbol{\beta}_{\bullet\Sigma}^{<t_0}), \Sigma_i^{t_0} = \text{null})
$$

$$
\left. \cdot R_A(s^{\geq t_0}[t], \boldsymbol{\pi}_{\alpha A}(\beta_{\Sigma\bullet}(\boldsymbol{\beta}_{\bullet\Sigma}^{>t_0}[t], \boldsymbol{\pi}_{\alpha\Sigma+\text{null}}))) \right]
$$

By substituting this expression into $\Delta_T \geq 0$ and subtracting the first term from both sides, we produce exactly the inequality from the statement of the theorem. $\square$