Revision to the Proofs of: "Error Performance of Channel Coding in Random Access Communication"

Zheng Wang, *Member, IEEE*, Jie Luo, *Senior Member, IEEE*

The original proofs of Theorem 2 and Lemma 1 in the paper contain uncareful presentations that require further clarification. More specifically, the typicality thresholds used in the decoding algorithms should be functions of the codewords, since otherwise they may not be able to satisfy the formulas used in the proofs for their value determinations. The revised proofs are given below, and the results presented in Theorem 2 and Lemma 1 remain valid.

I. Proof of Theorem 2

Proof: We assume that the following decoding algorithm is used at the receiver.

Given the received channel symbols $\boldsymbol{y}$, the receiver outputs a message and rate vector pair $(\boldsymbol{w}, \boldsymbol{r})$, with $\boldsymbol{r} \in \mathcal{R}$, if for *all* user subsets $\mathcal{S} \subset \{1, \cdots, K\}$, the following two conditions are met.

$$\begin{aligned}
\text{C1R: } & -\frac{1}{N}\log Pr\{\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})}\} < -\frac{1}{N}\log Pr\{\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{w}},\tilde{\boldsymbol{r}})}\}, \\
& \text{for all } (\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{r}}) \text{ with } \tilde{\boldsymbol{r}} \in \mathcal{R}, (\tilde{\boldsymbol{w}}_{\mathcal{S}}, \tilde{\boldsymbol{r}}_{\mathcal{S}}) = (\boldsymbol{w}_{\mathcal{S}}, \boldsymbol{r}_{\mathcal{S}}), \\
& \text{and } (\tilde{w}_k, \tilde{r}_k) \neq (w_k, r_k), \forall k \notin \mathcal{S}, \\
\text{C2R: } & -\frac{1}{N}\log Pr\{\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})}\} < \tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}}, \boldsymbol{y}). \qquad (1)
\end{aligned}$$

Note that the typicality threshold $\tau_{(\boldsymbol{r},\mathcal{S})}(.)$ is a function of $(\boldsymbol{x}_{\mathcal{S}}, \boldsymbol{y})$. The threshold also depends on the rate vector $\boldsymbol{r}$ and the user subset $\mathcal{S}$.

Given a user subset $\mathcal{S} \subset \{1, \cdots, K\}$, we define the following three probability terms that will be extensively used in the probability bound derivation.

First, assume $(\boldsymbol{w}, \boldsymbol{r})$ is the transmitted message and rate pair with $\boldsymbol{r} \in \mathcal{R}$. We define $P_{m[\boldsymbol{r},\tilde{\boldsymbol{r}},\mathcal{S}]}$ as the probability that the receiver finds another message and rate pair $(\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{r}})$ with $\tilde{\boldsymbol{r}} \in \mathcal{R}$, $(\tilde{\boldsymbol{w}}_{\mathcal{S}}, \tilde{\boldsymbol{r}}_{\mathcal{S}}) = (\boldsymbol{w}_{\mathcal{S}}, \boldsymbol{r}_{\mathcal{S}})$, and $(\tilde{w}_k, \tilde{r}_k) \neq (w_k, r_k), \forall k \notin \mathcal{S}$, that has a likelihood value no worse than the transmitted codeword.

$$\begin{aligned}
P_{m[\boldsymbol{r},\tilde{\boldsymbol{r}},\mathcal{S}]} = & Pr\left\{P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})}) \leq P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{w}},\tilde{\boldsymbol{r}})})\right\}, \\
& (\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{r}}), \tilde{\boldsymbol{r}} \in \mathcal{R}, (\tilde{\boldsymbol{w}}_{\mathcal{S}}, \tilde{\boldsymbol{r}}_{\mathcal{S}}) = (\boldsymbol{w}_{\mathcal{S}}, \boldsymbol{r}_{\mathcal{S}}), \\
& (\tilde{w}_k, \tilde{r}_k) \neq (w_k, r_k), \forall k \notin \mathcal{S}. \qquad (2)
\end{aligned}$$

Second, assume $(\boldsymbol{w}, \boldsymbol{r})$ is the transmitted message and rate pair with $\boldsymbol{r} \in \mathcal{R}$. We define $P_{t[\boldsymbol{r},\mathcal{S}]}$ as the probability that the likelihood of the transmitted codeword is no larger than the predetermined threshold $\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}}, \boldsymbol{y})$.

$$P_{t[\boldsymbol{r},\mathcal{S}]} = Pr\left\{P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})}) \leq e^{-N\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})}\right\}, \qquad (3)$$

where the threshold $\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}}, \boldsymbol{y})$ will be optimized later[1].

Third, assume $(\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{r}})$ is the transmitted message and rate pair with $\tilde{\boldsymbol{r}} \notin \mathcal{R}$. We define $P_{i[\tilde{\boldsymbol{r}},\boldsymbol{r},\mathcal{S}]}$ as the probability that the receiver finds another message and rate pair $(\boldsymbol{w}, \boldsymbol{r})$ with $\boldsymbol{r} \in \mathcal{R}$, $(\boldsymbol{w}_{\mathcal{S}}, \boldsymbol{r}_{\mathcal{S}}) = (\tilde{\boldsymbol{w}}_{\mathcal{S}}, \tilde{\boldsymbol{r}}_{\mathcal{S}})$, and $(w_k, r_k) \neq (\tilde{w}_k, \tilde{r}_k), \forall k \notin \mathcal{S}$, that has a likelihood value above the required threshold $\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}}, \boldsymbol{y})$.

$$\begin{aligned}
P_{i[\tilde{\boldsymbol{r}},\boldsymbol{r},\mathcal{S}]} = & Pr\left\{P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})}) > e^{-N\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})}\right\}, \\
& (\boldsymbol{w}, \boldsymbol{r}), \boldsymbol{r} \in \mathcal{R}, (\boldsymbol{w}_{\mathcal{S}}, \boldsymbol{r}_{\mathcal{S}}) = (\tilde{\boldsymbol{w}}_{\mathcal{S}}, \tilde{\boldsymbol{r}}_{\mathcal{S}}), \\
& (w_k, r_k) \neq (\tilde{w}_k, \tilde{r}_k), \forall k \notin \mathcal{S}. \qquad (4)
\end{aligned}$$

With these probability definitions, we can upper bound the system error probability P_{es} by

$$\begin{aligned}
P_{es} \leq \max\Bigg\{ & \\
& \max_{\boldsymbol{r}\in\mathcal{R}} \sum_{\mathcal{S}\subset\{1,\cdots,K\}} \left[\sum_{\tilde{\boldsymbol{r}}\in\mathcal{R},\tilde{\boldsymbol{r}}_{\mathcal{S}}=\boldsymbol{r}_{\mathcal{S}}} P_{m[\boldsymbol{r},\tilde{\boldsymbol{r}},\mathcal{S}]} + P_{t[\boldsymbol{r},\mathcal{S}]}\right], \\
& \max_{\tilde{\boldsymbol{r}}\notin\mathcal{R}} \sum_{\mathcal{S}\subset\{1,\cdots,K\}} \sum_{\boldsymbol{r}\in\mathcal{R},\boldsymbol{r}_{\mathcal{S}}=\tilde{\boldsymbol{r}}_{\mathcal{S}}} P_{i[\tilde{\boldsymbol{r}},\boldsymbol{r},\mathcal{S}]}\Bigg\}. \qquad (5)
\end{aligned}$$

Next, we will upper bound each of the probability terms on the right hand side of (5).

Step 1: Upper-bounding $P_{m[\boldsymbol{r},\tilde{\boldsymbol{r}},\mathcal{S}]}$.

Assume $(\boldsymbol{w}, \boldsymbol{r})$ is the transmitted message and rate pair with $\boldsymbol{r} \in \mathcal{R}$. Given $\boldsymbol{r}, \tilde{\boldsymbol{r}} \in \mathcal{R}$, $P_{m[\boldsymbol{r},\tilde{\boldsymbol{r}},\mathcal{S}]}$ can be written as

$$P_{m[\boldsymbol{r},\tilde{\boldsymbol{r}},\mathcal{S}]} = E_{\boldsymbol{\theta}}\left[\sum_{\boldsymbol{y}} P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})\phi_{m[\boldsymbol{r},\tilde{\boldsymbol{r}},\mathcal{S}]}(\boldsymbol{y})\right], \qquad (6)$$

where $\phi_{m[\boldsymbol{r},\tilde{\boldsymbol{r}},\mathcal{S}]}(\boldsymbol{y}) = 1$ if $P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})}) \leq P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{w}},\tilde{\boldsymbol{r}})})$ for some $(\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{r}})$, with $(\tilde{\boldsymbol{w}}_{\mathcal{S}}, \tilde{\boldsymbol{r}}_{\mathcal{S}}) = (\boldsymbol{w}_{\mathcal{S}}, \boldsymbol{r}_{\mathcal{S}})$, and $(\tilde{w}_k, \tilde{r}_k) \neq (w_k, r_k), \forall k \notin \mathcal{S}$. $\phi_{m[\boldsymbol{r},\tilde{\boldsymbol{r}},\mathcal{S}]}(\boldsymbol{y}) = 0$ otherwise.

Zheng Wang is with Seagate Technology LLC, Longmont, CO 80503. E-mail: zen.wang@seagate.com. Jie Luo is with the Electrical and Computer Engineering Department, Colorado State University, Fort Collins, CO 80523. E-mail: rockey@engr.colostate.edu.

[1] As in the single-user case, the subscript $\boldsymbol{r}$ of $\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}}, \boldsymbol{y})$ represents the corresponding estimated rate of the receiver output. Note that we do not assume the receiver should know the transmitted rate.

For any $\rho > 0$ and $s > 0$, we can bound $\phi_{m[\boldsymbol{r},\tilde{\boldsymbol{r}},\mathcal{S}]}(\boldsymbol{y})$ by

$$\phi_{m[\boldsymbol{r},\tilde{\boldsymbol{r}},\mathcal{S}]}(\boldsymbol{y}) \leq \left[\frac{\sum_{\substack{\tilde{\boldsymbol{w}},(\tilde{\boldsymbol{w}}_{\mathcal{S}},\tilde{\boldsymbol{r}}_{\mathcal{S}})=(\boldsymbol{w}_{\mathcal{S}},\boldsymbol{r}_{\mathcal{S}}),\\ (\tilde{w}_k,\tilde{r}_k)\neq(w_k,r_k),\forall k\notin\mathcal{S}}} P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{w}},\tilde{\boldsymbol{r}})})^{\frac{s}{\rho}}}{P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{\frac{s}{\rho}}}\right]^{\rho}, \quad \rho > 0, s > 0. \tag{7}$$

Consequently, $P_{m[\boldsymbol{r},\tilde{\boldsymbol{r}},\mathcal{S}]}$ is upper bounded by

$$\begin{aligned}
P_{m[\boldsymbol{r},\tilde{\boldsymbol{r}},\mathcal{S}]} &\leq E_{\boldsymbol{\theta}}\Bigg[\sum_{\boldsymbol{y}} P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})}) \\
&\times \left[\frac{\sum_{\substack{\tilde{\boldsymbol{w}},(\tilde{\boldsymbol{w}}_{\mathcal{S}},\tilde{\boldsymbol{r}}_{\mathcal{S}})=(\boldsymbol{w}_{\mathcal{S}},\boldsymbol{r}_{\mathcal{S}}),\\ (\tilde{w}_k,\tilde{r}_k)\neq(w_k,r_k),\forall k\notin\mathcal{S}}} P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{w}},\tilde{\boldsymbol{r}})})^{\frac{s}{\rho}}}{P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{\frac{s}{\rho}}}\right]^{\rho}\Bigg] \\
&= \sum_{\boldsymbol{y}} E_{\boldsymbol{\theta}}\Bigg[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{1-s} \\
&\times \Bigg[\sum_{\substack{\tilde{\boldsymbol{w}},(\tilde{\boldsymbol{w}}_{\mathcal{S}},\tilde{\boldsymbol{r}}_{\mathcal{S}})=(\boldsymbol{w}_{\mathcal{S}},\boldsymbol{r}_{\mathcal{S}}),\\ (\tilde{w}_k,\tilde{r}_k)\neq(w_k,r_k),\forall k\notin\mathcal{S}}} P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{w}},\tilde{\boldsymbol{r}})})^{\frac{s}{\rho}}\Bigg]^{\rho}\Bigg] \\
&= \sum_{\boldsymbol{y}} E_{\boldsymbol{\theta}_{\mathcal{S}}}\Bigg[E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{1-s}\right] \\
&\times E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\Bigg[\Bigg[\sum_{\substack{\tilde{\boldsymbol{w}},(\tilde{\boldsymbol{w}}_{\mathcal{S}},\tilde{\boldsymbol{r}}_{\mathcal{S}})=(\boldsymbol{w}_{\mathcal{S}},\boldsymbol{r}_{\mathcal{S}}),\\ (\tilde{w}_k,\tilde{r}_k)\neq(w_k,r_k),\forall k\notin\mathcal{S}}} P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{w}},\tilde{\boldsymbol{r}})})^{\frac{s}{\rho}}\Bigg]^{\rho}\Bigg]\Bigg],
\end{aligned} \tag{8}$$

where in the last step, we can take the expectations operations over users not in $\mathcal{S}$ since codewords corresponding to $(\boldsymbol{w}_{\bar{\mathcal{S}}},\boldsymbol{r}_{\bar{\mathcal{S}}})$ and $(\tilde{\boldsymbol{w}}_{\bar{\mathcal{S}}},\tilde{\boldsymbol{r}}_{\bar{\mathcal{S}}})$ are generated independently.

Now assume $0 < \rho \leq 1$. Inequality (8) can be further bounded by

$$\begin{aligned}
P_{m[\boldsymbol{r},\tilde{\boldsymbol{r}},\mathcal{S}]} &\leq \sum_{\boldsymbol{y}} E_{\boldsymbol{\theta}_{\mathcal{S}}}\Bigg[E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{1-s}\right] \\
&\times E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\Bigg[\Bigg[\sum_{\tilde{\boldsymbol{w}},(\tilde{\boldsymbol{w}}_{\mathcal{S}},\tilde{\boldsymbol{r}}_{\mathcal{S}})=(\boldsymbol{w}_{\mathcal{S}},\boldsymbol{r}_{\mathcal{S}})} P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{w}},\tilde{\boldsymbol{r}})})^{\frac{s}{\rho}}\Bigg]^{\rho}\Bigg]\Bigg] \\
&\leq e^{N\rho\sum_{k\notin\mathcal{S}}\tilde{r}_k} \sum_{\boldsymbol{y}} E_{\boldsymbol{\theta}_{\mathcal{S}}}\left[E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{1-s}\right]\right. \\
&\times \left.\left[E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{w}},\tilde{\boldsymbol{r}})})^{\frac{s}{\rho}}\right]\right]^{\rho}\right].
\end{aligned} \tag{9}$$

Since (9) holds for all $0 < \rho \leq 1$, $s > 0$, and it is easy to verify that the bound becomes trivial for $s > 1$, we have

$$P_{m[\boldsymbol{r},\tilde{\boldsymbol{r}},\mathcal{S}]} \leq \exp\left\{-NE_m(\mathcal{S},\tilde{\boldsymbol{r}},\boldsymbol{P}_{\boldsymbol{X}|\boldsymbol{r}},\boldsymbol{P}_{\boldsymbol{X}|\tilde{\boldsymbol{r}}})\right\}, \tag{10}$$

where $E_m(\mathcal{S},\tilde{\boldsymbol{r}},\boldsymbol{P}_{\boldsymbol{X}|\boldsymbol{r}},\boldsymbol{P}_{\boldsymbol{X}|\tilde{\boldsymbol{r}}})$ is given by

$$\begin{aligned}
E_m(\mathcal{S},\tilde{\boldsymbol{r}},\boldsymbol{P}_{\boldsymbol{X}|\boldsymbol{r}},\boldsymbol{P}_{\boldsymbol{X}|\tilde{\boldsymbol{r}}}) &= \max_{0<\rho\leq 1} -\rho\sum_{k\notin\mathcal{S}}\tilde{r}_k \\
&+ \max_{0<s\leq 1} -\log\sum_{Y}\sum_{\boldsymbol{X}_{\mathcal{S}}}\prod_{k\in\mathcal{S}} P_{X|r_k}(X_k) \\
&\times \left(\sum_{\boldsymbol{X}_{\bar{\mathcal{S}}}}\prod_{k\notin\mathcal{S}} P_{X|r_k}(X_k)P(Y|\boldsymbol{X})^{1-s}\right) \\
&\times \left(\sum_{\boldsymbol{X}_{\bar{\mathcal{S}}}}\prod_{k\notin\mathcal{S}} P_{X|\tilde{r}_k}(X_k)P(Y|\boldsymbol{X})^{\frac{s}{\rho}}\right)^{\rho}.
\end{aligned} \tag{11}$$

Step 2: Upper-bounding $P_{t[\boldsymbol{r},\mathcal{S}]}$.

Assume $(\boldsymbol{w},\boldsymbol{r})$ is the transmitted message and rate pair with $\boldsymbol{r} \in \mathcal{R}$. Rewrite $P_{t[\boldsymbol{r},\mathcal{S}]}$ as

$$P_{t[\boldsymbol{r},\mathcal{S}]} = E_{\boldsymbol{\theta}}\left[\sum_{\boldsymbol{y}} P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})\phi_{t[\boldsymbol{r},\mathcal{S}]}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})\right], \tag{12}$$

where $\phi_{t[\boldsymbol{r},\mathcal{S}]}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y}) = 1$ if $P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})}) \leq e^{-N\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})}$, otherwise $\phi_{t[\boldsymbol{r},\mathcal{S}]}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y}) = 0$. Note that the value of $\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})$ will be specified later.

For any $s_1 > 0$, we can bound $\phi_{t[\boldsymbol{r},\mathcal{S}]}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})$ by

$$\phi_{t[\boldsymbol{r},\mathcal{S}]}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y}) \leq \frac{e^{-Ns_1\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})}}{P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{s_1}}, \quad s_1 > 0. \tag{13}$$

This yields

$$\begin{aligned}
P_{t[\boldsymbol{r},\mathcal{S}]} &\leq E_{\boldsymbol{\theta}}\left[\sum_{\boldsymbol{y}} P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{1-s_1}e^{-Ns_1\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})}\right] \\
&= \sum_{\boldsymbol{y}} E_{\boldsymbol{\theta}_{\mathcal{S}}}\left[E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{1-s_1}\right]e^{-Ns_1\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})}\right].
\end{aligned} \tag{14}$$

We will come back to this inequality later when we optimize $\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})$.

Step 3: Upper-bounding $P_{i[\tilde{\boldsymbol{r}},\boldsymbol{r},\mathcal{S}]}$.

Assume $(\tilde{\boldsymbol{w}},\tilde{\boldsymbol{r}})$ is the transmitted message and rate pair with $\tilde{\boldsymbol{r}} \notin \mathcal{R}$. Given $\boldsymbol{r} \in \mathcal{R}$, we first rewrite $P_{i[\tilde{\boldsymbol{r}},\boldsymbol{r},\mathcal{S}]}$ as

$$P_{i[\tilde{\boldsymbol{r}},\boldsymbol{r},\mathcal{S}]} = E_{\boldsymbol{\theta}}\left[\sum_{\boldsymbol{y}} P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{w}},\tilde{\boldsymbol{r}})})\phi_{i[\tilde{\boldsymbol{r}},\boldsymbol{r},\mathcal{S}]}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})\right], \tag{15}$$

where $\phi_{i[\tilde{\boldsymbol{r}},\boldsymbol{r},\mathcal{S}]}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y}) = 1$ if there exists $(\boldsymbol{w},\boldsymbol{r})$ with $\boldsymbol{r} \in \mathcal{R}$, $(\boldsymbol{w}_{\mathcal{S}},\boldsymbol{r}_{\mathcal{S}}) = (\tilde{\boldsymbol{w}}_{\mathcal{S}},\tilde{\boldsymbol{r}}_{\mathcal{S}})$, and $(w_k,r_k) \neq (\tilde{w}_k,\tilde{r}_k), \forall k \notin \mathcal{S}$ to satisfy $P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})}) > e^{-N\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})}$. Otherwise $\phi_{i[\tilde{\boldsymbol{r}},\boldsymbol{r},\mathcal{S}]}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y}) = 0$.

For any $s_2 > 0$ and $\tilde{\rho} > 0$, we can bound $\phi_{i[\tilde{\boldsymbol{r}},\boldsymbol{r},\mathcal{S}]}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})$ by

$$\phi_{i[\tilde{\boldsymbol{r}},\boldsymbol{r},\mathcal{S}]}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y}) \leq \left[\frac{\sum_{\substack{\boldsymbol{w},(\boldsymbol{w}_{\mathcal{S}},\boldsymbol{r}_{\mathcal{S}})=(\tilde{\boldsymbol{w}}_{\mathcal{S}},\tilde{\boldsymbol{r}}_{\mathcal{S}}),\\ (w_k,r_k)\neq(\tilde{w}_k,\tilde{r}_k),\forall k\notin\mathcal{S}}} P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{\frac{s_2}{\tilde{\rho}}}}{e^{-N\frac{s_2}{\tilde{\rho}}\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})}}\right]^{\tilde{\rho}}, \quad s_2 > 0, \tilde{\rho} > 0. \tag{16}$$

This gives,

$$P_{i[\tilde{\boldsymbol{r}},\boldsymbol{r},\mathcal{S}]} \leq \sum_{\boldsymbol{y}} E_{\boldsymbol{\theta}}\Bigg[P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{w}},\tilde{\boldsymbol{r}})}) \times \Bigg[\sum_{\substack{\boldsymbol{w},(\boldsymbol{w}_{\mathcal{S}},\boldsymbol{r}_{\mathcal{S}})=(\tilde{\boldsymbol{w}}_{\mathcal{S}},\tilde{\boldsymbol{r}}_{\mathcal{S}}),\\ (w_k,r_k)\neq(\tilde{w}_k,\tilde{r}_k),\forall k\notin\mathcal{S}}} P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{\frac{s_2}{\tilde{\rho}}}\Bigg]^{\tilde{\rho}} \times e^{Ns_2\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})}\Bigg]$$
$$\leq \sum_{\boldsymbol{y}} E_{\boldsymbol{\theta}_{\mathcal{S}}}\Bigg[E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{w}},\tilde{\boldsymbol{r}})})\right] \times E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\Bigg[\Bigg[\sum_{\boldsymbol{w},(\boldsymbol{w}_{\mathcal{S}},\boldsymbol{r}_{\mathcal{S}})=(\tilde{\boldsymbol{w}}_{\mathcal{S}},\tilde{\boldsymbol{r}}_{\mathcal{S}})} P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{\frac{s_2}{\tilde{\rho}}}\Bigg]^{\tilde{\rho}}\Bigg] \times e^{Ns_2\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})}\Bigg]. \quad (17)$$

Note that we can separate the expectation operators in the last step due to independence between the codewords of $(\boldsymbol{w}_{\bar{\mathcal{S}}},\boldsymbol{r}_{\bar{\mathcal{S}}})$ and $(\tilde{\boldsymbol{w}}_{\bar{\mathcal{S}}},\tilde{\boldsymbol{r}}_{\bar{\mathcal{S}}})$.

Assume $0<\tilde{\rho}\leq 1$. Inequality (17) leads to

$$P_{i[\tilde{\boldsymbol{r}},\boldsymbol{r},\mathcal{S}]} \leq \sum_{\boldsymbol{y}} E_{\boldsymbol{\theta}_{\mathcal{S}}}\Bigg[E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{w}},\tilde{\boldsymbol{r}})})\right] e^{N\tilde{\rho}\sum_{k\notin\mathcal{S}} r_k} \times \left\{E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{\frac{s_2}{\tilde{\rho}}}\right]\right\}^{\tilde{\rho}} e^{Ns_2\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})}\Bigg]$$
$$\leq \max_{\boldsymbol{r}'\notin\mathcal{R},\boldsymbol{r}'_{\mathcal{S}}=\boldsymbol{r}_{\mathcal{S}}} \sum_{\boldsymbol{y}} E_{\boldsymbol{\theta}_{\mathcal{S}}}\Bigg[E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w}',\boldsymbol{r}')})\right] \times \left\{E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{\frac{s_2}{\tilde{\rho}}}\right]\right\}^{\tilde{\rho}} e^{Ns_2\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})} \times e^{N\tilde{\rho}\sum_{k\notin\mathcal{S}} r_k}\Bigg]. \quad (18)$$

Note that the bound obtained in the last step is no longer a function of $\tilde{\boldsymbol{r}}_{\bar{\mathcal{S}}}$.

Step 4: Choosing $\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})$.

In this step, we determine the typicality threshold $\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})$ that optimizes the bounds in (14) and (18).

Define $\tilde{\boldsymbol{r}}^*\notin\mathcal{R}$ as

$$\tilde{\boldsymbol{r}}^* = \operatorname*{argmax}_{\boldsymbol{r}'\notin\mathcal{R},\boldsymbol{r}'_{\mathcal{S}}=\boldsymbol{r}_{\mathcal{S}}} \sum_{\boldsymbol{y}} E_{\boldsymbol{\theta}_{\mathcal{S}}}\Bigg[E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w}',\boldsymbol{r}')})\right] \times \left\{E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{\frac{s_2}{\tilde{\rho}}}\right]\right\}^{\tilde{\rho}} e^{Ns_2\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})} \times e^{N\tilde{\rho}\sum_{k\notin\mathcal{S}} r_k}\Bigg]. \quad (19)$$

Given $\boldsymbol{r}\in\mathcal{R}$, $\boldsymbol{y}$, and the auxiliary variables $s_1>0$, $s_2>0$, $0<\tilde{\rho}\leq 1$, we choose $\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})$ such that the following equality holds.

$$E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{1-s_1}\right] e^{-Ns_1\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})}$$
$$= E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{w}}^*,\tilde{\boldsymbol{r}}^*)})\right]\left\{E_{\boldsymbol{\theta}_{\mathcal{S}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{\frac{s_2}{\tilde{\rho}}}\right]\right\}^{\tilde{\rho}} \times e^{Ns_2\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})} e^{N\tilde{\rho}\sum_{k\notin\mathcal{S}} r_k}. \quad (20)$$

This is always possible if we do not enforce the natural constraint that $\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})\geq 0$.

Equation (20) implies

$$e^{-N\tau_{(\boldsymbol{r},\mathcal{S})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})} = \frac{\left\{E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{w}}^*,\tilde{\boldsymbol{r}}^*)})\right]\right\}^{\frac{1}{s_1+s_2}}}{\left\{E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{1-s_1}\right]\right\}^{\frac{1}{s_1+s_2}}} \times \left\{E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{\frac{s_2}{\tilde{\rho}}}\right]\right\}^{\frac{\tilde{\rho}}{s_1+s_2}} e^{N\frac{\tilde{\rho}}{s_1+s_2}\sum_{k\notin\mathcal{S}} r_k}. \quad (21)$$

Substitute (21) into (14), we get

$$P_{t[\boldsymbol{r},\mathcal{S}]} \leq \sum_{\boldsymbol{y}} E_{\boldsymbol{\theta}_{\mathcal{S}}}\Bigg[\left\{E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{1-s_1}\right]\right\}^{\frac{s_2}{s_1+s_2}} \times \left\{E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{w}}^*,\tilde{\boldsymbol{r}}^*)})\right]\right\}^{\frac{s_1}{s_1+s_2}} \times \left\{E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{\frac{s_2}{\tilde{\rho}}}\right]\right\}^{\frac{s_1\tilde{\rho}}{s_1+s_2}} e^{N\frac{s_1\tilde{\rho}}{s_1+s_2}\sum_{k\notin\mathcal{S}} r_k}\Bigg]. \quad (22)$$

Assume $s_2<\tilde{\rho}$. Let $s_1=1-\frac{s_2}{\tilde{\rho}}$. Inequality (22) becomes

$$P_{t[\boldsymbol{r},\mathcal{S}]} \leq \sum_{\boldsymbol{y}} E_{\boldsymbol{\theta}_{\mathcal{S}}}\Bigg[\left\{E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{\frac{s_2}{\tilde{\rho}}}\right]\right\}^{\frac{\tilde{\rho}^2}{\tilde{\rho}-(1-\tilde{\rho})s_2}} \times \left\{E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{w}},\tilde{\boldsymbol{r}}^*)})\right]\right\}^{\frac{\tilde{\rho}-s_2}{\tilde{\rho}-(1-\tilde{\rho})s_2}} \times e^{N\frac{\tilde{\rho}(\tilde{\rho}-s_2)}{\tilde{\rho}-(1-\tilde{\rho})s_2}\sum_{k\notin\mathcal{S}} r_k}\Bigg]. \quad (23)$$

Now do a variable change with $\rho=\frac{\tilde{\rho}(\tilde{\rho}-s_2)}{\tilde{\rho}-(1-\tilde{\rho})s_2}$ and $s=1-\frac{\tilde{\rho}-s_2}{\tilde{\rho}-(1-\tilde{\rho})s_2}$, and note that $s+\rho\leq 1$. Inequality (23) becomes

$$P_{t[\boldsymbol{r},\mathcal{S}]} \leq \sum_{\boldsymbol{y}} E_{\boldsymbol{\theta}_{\mathcal{S}}}\Bigg[\left\{E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{w},\boldsymbol{r})})^{\frac{s}{s+\rho}}\right]\right\}^{s+\rho} \times \left\{E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{w}}^*,\tilde{\boldsymbol{r}}^*)})\right]\right\}^{1-s} e^{N\rho\sum_{k\notin\mathcal{S}} r_k}\Bigg]$$
$$\leq \max_{\boldsymbol{r}'\notin\mathcal{R},\boldsymbol{r}'_{\mathcal{S}}=\boldsymbol{r}_{\mathcal{S}}} \Bigg\{\sum_{Y}\sum_{\boldsymbol{X}_{\mathcal{S}}}\prod_{k\in\mathcal{S}} P_{X|r_k}(X_k) \times \Bigg(\sum_{\boldsymbol{X}_{\bar{\mathcal{S}}}}\prod_{k\notin\mathcal{S}} P_{X|r_k}(X_k)P(Y|\boldsymbol{X})^{\frac{s}{s+\rho}}\Bigg)^{s+\rho} \times \Bigg(\sum_{\boldsymbol{X}_{\bar{\mathcal{S}}}}\prod_{k\notin\mathcal{S}} P_{X|r'_k}(X_k)P(Y|\boldsymbol{X})\Bigg)^{1-s}\Bigg\}^N \times e^{N\rho\sum_{k\notin\mathcal{S}} r_k}. \quad (24)$$

Following the same derivation, we can see that $P_{i[\tilde{\boldsymbol{r}},\boldsymbol{r},\mathcal{S}]}$ is also upper-bounded by the right hand side of (24). Because (24) holds for all $0<\rho\leq 1$ and $0<s\leq 1-\rho$, we have

$$P_{t[\boldsymbol{r},\mathcal{S}]}, P_{i[\tilde{\boldsymbol{r}},\boldsymbol{r},\mathcal{S}]} \leq \max_{\boldsymbol{r}'\notin\mathcal{R},\boldsymbol{r}'_{\mathcal{S}}=\boldsymbol{r}_{\mathcal{S}}} \exp\{-NE_i(\mathcal{S},\boldsymbol{r},\boldsymbol{P}_{\boldsymbol{X}|\boldsymbol{r}},\boldsymbol{P}_{\boldsymbol{X}|\boldsymbol{r}'})\}, \quad (25)$$

where

$$E_i(\mathcal{S},\boldsymbol{r},\boldsymbol{P}_{\boldsymbol{X}|\boldsymbol{r}},\boldsymbol{P}_{\boldsymbol{X}|\boldsymbol{r}'}) = \max_{0<\rho\le 1} -\rho\sum_{k\notin\mathcal{S}} r_k + \max_{0<s\le 1-\rho} -\log\sum_Y\sum_{\boldsymbol{X}_\mathcal{S}}\prod_{k\in\mathcal{S}} P_{X|r_k}(X_k) \times\left(\sum_{\boldsymbol{X}_{\bar{\mathcal{S}}}}\prod_{k\notin\mathcal{S}} P_{X|r_k}(X_k)P(Y|\boldsymbol{X})^{\frac{s}{s+\rho}}\right)^{s+\rho} \times\left(\sum_{\boldsymbol{X}_{\bar{\mathcal{S}}}}\prod_{k\notin\mathcal{S}} P_{X|r'_k}(X_k)P(Y|\boldsymbol{X})\right)^{1-s}. \quad (26)$$

Finally, substituting (10) and (25) into (5) gives the desired result.

■

II. Proof of Lemma 1

Proof: We assume a similar decoding algorithm as given in (1), with the second condition being revised to

$$\text{C2R:}\ -\frac{1}{N}\log Pr\{\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{W},\boldsymbol{r})}\} < \tau_{(\boldsymbol{r}_\mathcal{S},\boldsymbol{U}(\boldsymbol{r}_{\bar{\mathcal{S}}}))}(\boldsymbol{x}_\mathcal{S},\boldsymbol{y}). \quad (27)$$

In other words, we assume that the typicality threshold $\tau_{(\boldsymbol{r}_\mathcal{S},\boldsymbol{U}(\boldsymbol{r}_{\bar{\mathcal{S}}}))}(\boldsymbol{x}_\mathcal{S},\boldsymbol{y})$ should depend on the standard rates for users in $\mathcal{S}$ and the grid rates for users not in $\mathcal{S}$.

Given a user subset $\mathcal{S}\subset\{1,\cdots,K\}$, we define the following three probability terms.

First, assume $(\boldsymbol{W},\boldsymbol{r})$ is the transmitted message and rate pair with $\boldsymbol{r}\in\mathcal{R}$. We define $P_{m[\boldsymbol{r},\tilde{\boldsymbol{r}}^U,\mathcal{S}]}$ as the probability that the receiver finds another codeword and rate pair $(\tilde{\boldsymbol{W}},\tilde{\boldsymbol{r}})$ with $\tilde{\boldsymbol{r}}\in\mathcal{R}$, $\boldsymbol{U}(\tilde{\boldsymbol{r}})=\tilde{\boldsymbol{r}}^U$, $(\tilde{\boldsymbol{W}}_\mathcal{S},\tilde{\boldsymbol{r}}_\mathcal{S})=(\boldsymbol{W}_\mathcal{S},\boldsymbol{r}_\mathcal{S})$, and $(\tilde{W}_k,\tilde{r}_k)\neq(W_k,r_k),\forall k\notin\mathcal{S}$, that has a likelihood value no worse than the transmitted codeword. That is

$$P_{m[\boldsymbol{r},\tilde{\boldsymbol{r}}^U,\mathcal{S}]} = Pr\left\{P(\boldsymbol{y}|\boldsymbol{x}_{(W,r)})\le P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{W},\tilde{r})})\right\}, \quad (\tilde{\boldsymbol{W}},\tilde{\boldsymbol{r}}),\tilde{\boldsymbol{r}}\in\mathcal{R},\boldsymbol{U}(\tilde{\boldsymbol{r}})=\tilde{\boldsymbol{r}}^U,(\tilde{\boldsymbol{W}}_\mathcal{S},\tilde{\boldsymbol{r}}_\mathcal{S})=(\boldsymbol{W}_\mathcal{S},\boldsymbol{r}_\mathcal{S}), \quad (\tilde{W}_k,\tilde{r}_k)\neq(W_k,r_k),\forall k\notin\mathcal{S}. \quad (28)$$

Second, assume $(\boldsymbol{W},\boldsymbol{r})$ is the transmitted message and rate pair with $\boldsymbol{r}\in\mathcal{R}$. We define $P_{t[\boldsymbol{r},\mathcal{S}]}$ as in (3) except the typicality threshold is replaced by $\tau_{(\boldsymbol{r}_\mathcal{S},\boldsymbol{U}(\boldsymbol{r}_{\bar{\mathcal{S}}}))}(\boldsymbol{y})$.

Third, assume $(\tilde{\boldsymbol{W}},\tilde{\boldsymbol{r}})$ is the transmitted message and rate pair with $\tilde{\boldsymbol{r}}\notin\mathcal{R}$. We define $P_{i[\tilde{\boldsymbol{r}},\boldsymbol{r}^U,\mathcal{S}]}$ as the probability that the receiver finds another codeword and rate pair $(\boldsymbol{W},\boldsymbol{r})$ with $\boldsymbol{r}\in\mathcal{R}$, $\boldsymbol{U}(\boldsymbol{r})=\boldsymbol{r}^U$, $(\boldsymbol{W}_\mathcal{S},\boldsymbol{r}_\mathcal{S})=(\tilde{\boldsymbol{W}}_\mathcal{S},\tilde{\boldsymbol{r}}_\mathcal{S})$, and $(W_k,r_k)\neq(\tilde{W}_k,\tilde{r}_k),\forall k\notin\mathcal{S}$, that has a likelihood value above the required threshold $\tau_{(\tilde{\boldsymbol{r}}_\mathcal{S},\boldsymbol{r}^U_{\bar{\mathcal{S}}})}(\boldsymbol{x}_\mathcal{S},\boldsymbol{y})$. That is

$$P_{i[\tilde{\boldsymbol{r}},\boldsymbol{r}^U,\mathcal{S}]} = Pr\left\{P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{W},\boldsymbol{r})}) > e^{-N\tau_{(\tilde{\boldsymbol{r}}_\mathcal{S},\boldsymbol{r}^U_{\bar{\mathcal{S}}})}(\boldsymbol{x}_\mathcal{S},\boldsymbol{y})}\right\}, \quad (\boldsymbol{W},\boldsymbol{r}),\boldsymbol{r}\in\mathcal{R},\boldsymbol{U}(\boldsymbol{r})=\boldsymbol{r}^U,(\boldsymbol{W}_\mathcal{S},\boldsymbol{r}_\mathcal{S})=(\tilde{\boldsymbol{W}}_\mathcal{S},\tilde{\boldsymbol{r}}_\mathcal{S}), \quad (W_k,r_k)\neq(\tilde{W}_k,\tilde{r}_k),\forall k\notin\mathcal{S}. \quad (29)$$

With the probability definitions, we can upper bound the system error probability P_{es} by

$$P_{es}\le\max\Bigg\{ \max_{\boldsymbol{r}\in\mathcal{R}}\sum_{\mathcal{S}\subset\{1,\cdots,K\}}\left[\sum_{\tilde{\boldsymbol{r}}^U,\tilde{\boldsymbol{r}}^U_{\mathcal{S}}=\boldsymbol{U}(\boldsymbol{r}_\mathcal{S})} P_{m[\boldsymbol{r},\tilde{\boldsymbol{r}}^U,\mathcal{S}]}+P_{t[\boldsymbol{r},\mathcal{S}]}\right], \max_{\tilde{\boldsymbol{r}}\notin\mathcal{R}}\sum_{\mathcal{S}\subset\{1,\cdots,K\}}\sum_{\boldsymbol{r}^U,\boldsymbol{r}^U_\mathcal{S}=\boldsymbol{U}(\tilde{\boldsymbol{r}}_\mathcal{S})} P_{i[\tilde{\boldsymbol{r}},\boldsymbol{r}^U,\mathcal{S}]}\Bigg\}. \quad (30)$$

We will then follow similar steps as in the proof of Theorem 2 to upper bound each of the probability terms on the right hand side of (30).

To upper bound $P_{m[\boldsymbol{r},\tilde{\boldsymbol{r}}^U,\mathcal{S}]}$, we assume $0<\rho\le 1$, $0<s\le 1$, and get from (9) that

$$P_{m[\boldsymbol{r},\tilde{\boldsymbol{r}}^U,\mathcal{S}]}\le\sum_{\boldsymbol{y}}E_{\boldsymbol{\theta}_\mathcal{S}}\left[E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{W},\boldsymbol{r})})^{1-s}\right.\right. \times\left[\sum_{\substack{\tilde{\boldsymbol{W}},(\tilde{\boldsymbol{W}}_\mathcal{S},\tilde{\boldsymbol{r}}_\mathcal{S})=(\boldsymbol{W}_\mathcal{S},\boldsymbol{r}_\mathcal{S}),\\ \boldsymbol{U}(\tilde{\boldsymbol{r}})=\tilde{\boldsymbol{r}}^U}}E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{W}},\tilde{\boldsymbol{r}})})^{\frac{s}{\rho}}\right]^{\rho}\right]\Bigg]\Bigg]$$
$$\le e^{N\rho\sum_{k\notin\mathcal{S}}\tilde{r}^U_k}\sum_{\boldsymbol{y}}E_{\boldsymbol{\theta}_\mathcal{S}}\left[E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{W},\boldsymbol{r})})^{1-s}\right.\right. \times\left[\max_{\substack{\tilde{\boldsymbol{W}},(\tilde{\boldsymbol{W}}_\mathcal{S},\tilde{\boldsymbol{r}}_\mathcal{S})=(\boldsymbol{W}_\mathcal{S},\boldsymbol{r}_\mathcal{S}),\\ \boldsymbol{U}(\tilde{\boldsymbol{r}})=\tilde{\boldsymbol{r}}^U}}E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{W}},\tilde{\boldsymbol{r}})})^{\frac{s}{\rho}}\right]\right]^{\rho}\Bigg]\Bigg]$$
$$\le\exp\left\{-N\tilde{E}_m\left(\mathcal{S},\tilde{\boldsymbol{r}}^U,\boldsymbol{P}_{\boldsymbol{X}|\boldsymbol{r}},\boldsymbol{P}_{\substack{\boldsymbol{X}|\tilde{\boldsymbol{r}},\forall\tilde{\boldsymbol{r}}\in\mathcal{R},\\ \boldsymbol{U}(\tilde{\boldsymbol{r}})=\tilde{\boldsymbol{r}}^U,\\ \tilde{\boldsymbol{r}}_\mathcal{S}=\boldsymbol{r}_\mathcal{S}}}\right)\right\}, \quad (31)$$

where $\tilde{E}_m(\mathcal{S},\tilde{\boldsymbol{r}}^U,\boldsymbol{P}_{\boldsymbol{X}|\boldsymbol{r}},\boldsymbol{P}_{\boldsymbol{X}|\tilde{\boldsymbol{r}},\forall\tilde{\boldsymbol{r}}\in\mathcal{R},\boldsymbol{U}(\tilde{\boldsymbol{r}})=\tilde{\boldsymbol{r}}^U,\tilde{\boldsymbol{r}}_\mathcal{S}=\boldsymbol{r}_\mathcal{S}})$ is defined in the lemma.

To upper bound $P_{t[\boldsymbol{r},\mathcal{S}]}$, we get from (14) for $s_1>0$ that

$$P_{t[\boldsymbol{r},\mathcal{S}]}\le\sum_{\boldsymbol{y}}E_{\boldsymbol{\theta}_\mathcal{S}}\left[E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{W},\boldsymbol{r})})^{1-s_1}\right]\right. \left.\times e^{-Ns_1\tau_{(\boldsymbol{r}_\mathcal{S},\boldsymbol{U}(\boldsymbol{r}_{\bar{\mathcal{S}}}))}(\boldsymbol{x}_\mathcal{S},\boldsymbol{y})}\right]. \quad (32)$$

To upper bound $P_{i[\tilde{\boldsymbol{r}},\boldsymbol{r}^U,\mathcal{S}]}$, we get from (18) for $s_2>0$ and $0<\tilde{\rho}\le 1$ that

$$P_{i[\tilde{\boldsymbol{r}},\boldsymbol{r}^U,\mathcal{S}]}\le\sum_{\boldsymbol{y}}E_{\boldsymbol{\theta}_\mathcal{S}}\left[E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{W}},\tilde{\boldsymbol{r}})})\right]\right. \times\left\{\sum_{(\boldsymbol{W},\boldsymbol{r}),\boldsymbol{r}_\mathcal{S}=\tilde{\boldsymbol{r}}_\mathcal{S},\boldsymbol{U}(\boldsymbol{r}_{\bar{\mathcal{S}}})=\boldsymbol{r}^U_{\bar{\mathcal{S}}}}E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{W},\boldsymbol{r})})^{\frac{s_2}{\tilde{\rho}}}\right]\right\}^{\tilde{\rho}}\Bigg] \times e^{Ns_2\tau_{(\tilde{\boldsymbol{r}}_\mathcal{S},\boldsymbol{r}^U_{\bar{\mathcal{S}}})}(\boldsymbol{x}_\mathcal{S},\boldsymbol{y})}$$
$$\le\sum_{\boldsymbol{y}}E_{\boldsymbol{\theta}_\mathcal{S}}\left[E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}}\left[P(\boldsymbol{y}|\boldsymbol{x}_{(\tilde{\boldsymbol{W}},\tilde{\boldsymbol{r}})})\right]\right.$$

$$
\begin{aligned}
&\times \left\{ \max_{(\boldsymbol{W},\boldsymbol{r}),\boldsymbol{r}_{\mathcal{S}}=\tilde{\boldsymbol{r}}_{\mathcal{S}},\boldsymbol{U}(\boldsymbol{r}_{\bar{\mathcal{S}}})=\boldsymbol{r}^U_{\bar{\mathcal{S}}}} E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}} \left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{W},\boldsymbol{r})})^{\frac{s_2}{\tilde{\rho}}} \right] \right\}^{\tilde{\rho}} \Bigg] \\
&\times e^{N s_2 \tau_{(\tilde{\boldsymbol{r}}_{\mathcal{S}},\boldsymbol{r}^U_{\bar{\mathcal{S}}})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})} e^{N\tilde{\rho}\sum_{k\notin\mathcal{S}} r^U_k} \\
&\le \max_{\boldsymbol{r}'\notin\mathcal{R},\boldsymbol{r}'_{\mathcal{S}}=\tilde{\boldsymbol{r}}_{\mathcal{S}}} \sum_{\boldsymbol{y}} E_{\boldsymbol{\theta}_{\mathcal{S}}} \Big[E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}} \left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{W}',\boldsymbol{r}')}) \right] \\
&\times \left\{ \max_{(\boldsymbol{W},\boldsymbol{r}),\boldsymbol{r}_{\mathcal{S}}=\tilde{\boldsymbol{r}}_{\mathcal{S}},\boldsymbol{U}(\boldsymbol{r}_{\bar{\mathcal{S}}})=\boldsymbol{r}^U_{\bar{\mathcal{S}}}} E_{\boldsymbol{\theta}_{\bar{\mathcal{S}}}} \left[P(\boldsymbol{y}|\boldsymbol{x}_{(\boldsymbol{W},\boldsymbol{r})})^{\frac{s_2}{\tilde{\rho}}} \right] \right\}^{\tilde{\rho}} \Bigg] \\
&\times e^{N s_2 \tau_{(\tilde{\boldsymbol{r}}_{\mathcal{S}},\boldsymbol{r}^U_{\bar{\mathcal{S}}})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})} e^{N\tilde{\rho}\sum_{k\notin\mathcal{S}} r^U_k}. \qquad (33)
\end{aligned}
$$

Next, by following a derivation similar to Step 4 in the proof of Theorem 2, we can optimize (32) and (33) jointly over $\tau_{(\tilde{\boldsymbol{r}}_{\mathcal{S}},\boldsymbol{r}^U_{\bar{\mathcal{S}}})}(\boldsymbol{x}_{\mathcal{S}},\boldsymbol{y})$ to obtain the desired result.

■