1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned}
\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\
&= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\
&= \Pr\{Err|W = 1\}.
\end{aligned}$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,

$$\Pr\{Err^c|W = 1\} \geq \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\}.$$

5. Consider

$$\begin{aligned}
&\Pr\{Err|W = 1\} \\
&\quad= 1 - \Pr\{Err^c|W = 1\} \\
&\quad\leq 1 - \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\} \\
&\quad= \Pr\{(E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c)^c|W = 1\} \\
&\quad= \Pr\{E_1^c \cup E_2 \cup E_3 \cup \cdots \cup E_M|W = 1\}.
\end{aligned}$$

6. By the union bound,

$$\Pr\{Err|W = 1\} \leq \Pr\{E_1^c|W = 1\} + \sum_{w=2}^{M} \Pr\{E_w|W = 1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W = 1\} = \Pr\{(\tilde{\mathbf{X}}(1), \mathbf{Y}) \notin T^n_{[XY]\delta}|W = 1\} < \nu.$$

8. Conditioning on $\{W = 1\}$, for $2 \leq w \leq M$, $(\tilde{\mathbf{X}}(w), \mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X', Y'), where $X' \sim X$ and $Y' \sim Y$.