6. By the union bound,

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

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 $2 \le w \le M$, $(\tilde{\mathbf{X}}(w), \mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X', Y'), where $X' \sim X$ and $Y' \sim Y$.

9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \le w \le M$,

$$\begin{aligned} \Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\le 2^{-n(I(X;Y)-\tau)} \end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\begin{aligned} \Pr\{Err\} &< \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)} \\ &= \nu + 2^{-n(\frac{\epsilon}{4}-\tau)}. \end{aligned}$$