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 $2 \leq w \leq M$,

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

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,

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

13. Recall that ϵ is fixed. Since $\tau \to 0$ as $\delta \to 0$, we can choose δ to be sufficiently small so that

$$\frac{\epsilon}{4} - \tau > 0.$$

14. Then

$$2^{-n(\frac{\epsilon}{4}-\tau)} \to 0$$

as $n \to \infty$.

15. Let $\nu < \frac{\epsilon}{3}$ to obtain

$$\Pr\{Err\} < \frac{\epsilon}{2}$$

for sufficiently large n.