We firstly prove bound on $\|V^* - \hat{V}\|$:

$$\|V^* - \hat{V}\| = \|T^* V^* - \hat{T}\hat{V}\| \leq \|T^* V^* - \hat{T} V^*\| + \|\hat{T} V^* - \hat{T}\hat{V}^*\|$$

$$\leq \gamma \|P V^* - \hat{P} V^*\| + \gamma \|V^* - \hat{V}^*\|$$

By solving this inequality with respect to $\|V^* - \hat{V}\|$ we deduce

$$\|V^* - \hat{V}\| \leq \gamma b \|(P - \hat{P}) V^*\|$$

By Hoeffding's inequality to bound $|(P - \hat{P})V^*(x)|$ for all $x \in \mathcal{X}$ in high probability:

$$\mathbb{P}\left(|((P - \hat{P})V^*)(x)| \geq \varepsilon\right) \leq 2 e^{\left(-\frac{N\varepsilon^2}{2b^2}\right)}$$

We deduce

$$\mathbb{P}\left(\|(P - \hat{P})V^*\| \geq \varepsilon\right) \leq 2|\mathcal{X}| e^{\left(-\frac{N\varepsilon^2}{2b^2}\right)}$$

We then define δ as

$$\delta \triangleq 2|\mathcal{X}| e^{\left(-\frac{N\varepsilon^2}{2b^2}\right)}$$

By plugging last two (in)equalities we deduce

$$\mathbb{P}\left[\|(P - \hat{P})V^*\| < b\sqrt{\ln\left(\frac{2|\mathcal{X}|}{\delta}\right)/N}\right] \geq 1 - \delta$$

(in our situation $b = 1$, as R is in $[0,1]$.)