For any policy $\vec{\mu} = \mu_1, \ldots, \mu_n$, for any state x and any stationary adversary policy ν, there exists a decomposition:

$$\vec{\mu} = \vec{\mu}_P \vec{\mu}_C \vec{\mu}_{P'}$$

such that

$$\mathbb{1}_x P_{\vec{\mu}_P,\nu} = \mathbb{1}_x P_{\vec{\mu}_P,\nu} P_{\vec{\mu}_C,\nu} = \mathbb{1}_y(\vec{\mu}, x, \nu)$$

If we consider the policy $\tilde{\mu}(\vec{\mu}, x, \nu) = \mu_P(\mu_C)^\infty$, then for any w,

$$\begin{aligned}
v_{\tilde{\mu}(\vec{\mu},x,\nu),\nu}(x) - T_{\vec{\mu}_P,\nu} T_{\vec{\mu}_C,\nu} w(x) &= \mathbb{1}_x [T_{\vec{\mu}_P,\nu} (T_{\vec{\mu}_C,\nu})^\infty w - T_{\vec{\mu}_P,\nu} T_{\vec{\mu}_C,\nu} w] \\
&= \sum_{k=1}^{\infty} \mathbb{1}_x T_{\vec{\mu}_P,\nu} (T_{\vec{\mu}_C,\nu})^{k+1} w - \mathbb{1}_x T_{\vec{\mu}_P,\nu} (T_{\vec{\mu}_C,\nu})^k w \\
&= \sum_{k=1}^{\infty} \mathbb{1}_x \Gamma_{\vec{\mu}_P,\nu} (\Gamma_{\vec{\mu}_C,\nu})^k (T_{\vec{\mu}_C,\nu} w - w) \\
&= \frac{\gamma^{i+c}}{1-\gamma^c} \mathbb{1}_{y(\vec{\mu},x,\nu)} (T_{\vec{\mu}_C,\nu} w - w).
\end{aligned}$$

Therefore by taking $w = T_{\vec{\mu}_{P'},\nu} v$, we get for any v,

$$v_{\tilde{\mu}(\vec{\mu},x,\nu),\nu}(x) - T_{\vec{\mu},\nu} v(x) = \frac{\gamma^{i+c}}{1-\gamma^c} \mathbb{1}_{y(\vec{\mu},x,\nu)} (T_{\vec{\mu}_C,\nu} T_{\vec{\mu}_{P'},\nu} v - T_{\vec{\mu}_{P'},\nu} v).$$

Symmetrically we have for any policy $\vec{\nu} = \nu_1 \ldots \nu_n$, any state x and any stationary adversary policy μ,

$$v_{\mu,\tilde{\nu}(\vec{\nu},x,\mu)}(x) - T_{\mu,\vec{\nu}} v(x) = \frac{\gamma^{j+d}}{1-\gamma^d} \mathbb{1}_{z(\vec{\nu},x,\mu)} (T_{\mu,\vec{\nu}_C} T_{\mu,\vec{\nu}_{P'}} v - T_{\mu,\vec{\nu}_{P'}} v).$$

Assume $v \geq Tv$ and

$$T_{\mu_1,\nu_1} \ldots T_{\mu_n,\nu_n} v = T_{\vec{\mu},\vec{\nu}} v = T^n v.$$

On the one hand, we have

$$\begin{aligned}
v_{\mu_*,\nu_*}(x) - T^n v(x) &\leq v_{\mu_*,\tilde{\nu}(\vec{\nu},x,\mu_*)}(x) - T^n v(x) \\
&= \frac{\gamma^{j+d}}{1-\gamma^d} \mathbb{1}_{z(\vec{\nu},x,\mu_*)} (T_{\mu_*,\vec{\nu}_C} T_{\mu_*,\vec{\nu}_{P'}} v - T_{\mu_*,\vec{\nu}_{P'}} v) \\
&\leq \frac{\gamma^{j+d}}{1-\gamma^d} \mathbb{1}_{z(\vec{\nu},x,\mu_*)} (T^{n-j-d} v - T^{n-j} v).
\end{aligned}$$

From z such that $T^{n-j-d}v - T^{n-j}v \leq \rho$, can we do $\vec{\nu}_C$ and then force any policy μ_* to revisit z ? If yes, then we have a bound on v_*.

On the other hand, we have for all $\bar{\nu}$;

$$\begin{aligned}
v_{\tilde{\mu}(\vec{\mu},x,\bar{\nu})}(x) - T^n v(x) &= \frac{\gamma^{i+c}}{1-\gamma^c} \mathbb{1}_{y(\vec{\mu},x,\bar{\nu})} (T_{\vec{\mu}_C,\bar{\nu}} T_{\vec{\mu}_{P'},\bar{\nu}} v - T_{\vec{\mu}_{P'},\bar{\nu}} v) \\
&\geq \frac{\gamma^{i+c}}{1-\gamma^c} \mathbb{1}_{y(\vec{\mu},x,\bar{\nu})} (T^{n-i-c} v - T^{n-i} v).
\end{aligned}$$

From y, such that $T^{n-j-d}v - T^{n-j}v > \rho$, can we do $\vec{\mu}_C$ and then force any policy $\bar{\nu}$ to revisit y ? If yes, then we make a significant improvement.