Proof. If $t = 0$, the RHS is 1 and the result trivial. Henceforth we assume $t > 0$. The proof splits into cases $f(\pi^{i-1}) = 1$ and $f(\pi^{i-1}) > 1$, which we consider in turn. Let $[x]$ denote the set $\{1, 2, \dots, x\}$.

If $f(\pi^{i-1}) = 1$, $s = 1$ is switchable in π^{i-1}. From Lemma 10, $\pi^{i-1}(1) = 0$. Thus, let $\pi^{i-1} = 0^{s'}x$ for some $1 \le s' \le n$, $x \in \{0,1\}^{n-s'}$, and x starts with 1 or x is empty. Applying Lemma 10 for states $1, 2, \dots, s'+1$, we get that states $1, 2, \dots, s'$ are switchable in π^{i-1} and $s'+1$ is not switchable in π^{i-1}, if $s'+1 \in [n]$. If $f(\pi^i) \ge t+1$, the states $1, 2, \dots, t$ are not switchable in π^i. Applying Lemma 10 for states $1, 2, \dots, t$, we get that $\pi^i = 10^{t-1}y$ where $y \in \{0,1\}^{n-t}$. If $s' = n$, $t \le n = s'$. t cannot be greater than s' if $s'+1 \in [n]$ as that will imply $\pi^i(s'+1) = 0 \ne \pi^{i-1}(s'+1)$, despite $s'+1$ not being switchable in π^{i-1}. Hence, if $t > s'$, $\mathbb{P}\{f(\pi^i) \ge t+1\} = 0 \le \frac{1}{2^t}$. Otherwise $t \le s'$. Therefore, states $1, 2, \dots, t$ are switchable in π^{i-1}. To get to π^i from π^{i-1}, the state 1 must be switched and the states $2, 3, \dots, t$ must not be switched. As each state is switched with probability $\frac{1}{2}$ by RPI1, the probability of this event happening is exactly $\frac{1}{2^t}$.

If $f(\pi^{i-1}) = s > 1$, s is switchable in π^{i-1} and $1, 2, \dots, s-1$ are not switchable in π^{i-1}. Applying Lemma 10 for states $1, 2, \dots, s$, we get $\pi^{i-1} = 10^{s-2}10^{s'}x$ for some $0 \le s' \le n-s$, $x \in \{0,1\}^{n-s-s'}$, and x starts with 1 or x is empty. Applying Lemma 10 for states $s+1, s+2, \dots, s+s'$, we get that states $s+1, s+2, \dots, s+s'$ are also switchable in π^{i-1} and $s+s'+1$ is not switchable in π^{i-1}, if $s+s'+1 \in [n]$. Note that since $i-1 < m$, $\pi^{i-1} \ne \pi^*$ and hence $s \le n$. If $f(\pi^i) \ge s+t$, the states $1, 2, \dots, s+t-1$ are not switchable in π^i. Applying Lemma 10 for states $1, 2, \dots, s+t-1$, we get that $\pi^i = 10^{s+t-2}y$ where $y \in \{0,1\}^{n-s-t+1}$. If $s+s' = n$, $s+t-1 \le n = s+s'$. $s+t-1$ cannot be greater than $s+s'$ if $s+s'+1 \in [n]$ as that will imply $\pi^i(s+s'+1) = 0 \ne \pi^{i-1}(s+s'+1)$, despite $s+s'+1$ not being switchable in π^{i-1}. Hence, if $s+t-1 > s+s'$, $\mathbb{P}\{f(\pi^i) \ge s+t\} = 0 \le \frac{1}{2^t}$. Otherwise $s+t-1 \le s+s'$. Therefore, states $s, s+1, \dots, s+t-1$ are switchable in π^{i-1}. To get to π^i from π^{i-1}, the state s must be switched and the states $s+1, s+2, \dots, s+t-1$ must not be switched. As each state is switched with probability $\frac{1}{2}$ by RPI1, the probability of this event happening is exactly $\frac{1}{2^t}$. □

Definition 14. *We define $L : \Pi \to \mathbb{R}_{\ge 0}$, where $L(\pi)$ is the expected number of policies evaluated by RPI1 starting from π.*

Note that even if we start from $\pi^0 = \pi^*$, we need to evaluate π^0 to know that it is optimal. Hence $L(\pi^*) = 1$.

Definition 15. *We define $N : [n+1] \to \mathbb{R}_{\ge 0}$, where*

$$N(s) = \min_{\pi \in \Pi, f(\pi) = s} L(\pi).$$

It directly follows from the definition that $N(f(\pi)) \le L(\pi)$ for any $\pi \in \Pi$.

Theorem 16. *For $s \in [n+1]$, $N(s) \ge n + 2 - s$.*

Proof. If $s = n+1$, $f(\pi) = n+1$ is true only for $\pi = \pi^*$. Hence $N(n+1) = L(\pi^*) = 1 \ge n + 2 - (n+1)$.

Now, let $s \in [n]$. Let π be a policy such that $N(s) = L(\pi)$. Hence $f(\pi) = s$. Since $f(\pi^*) = n+1$, π is not optimal. Let π' be obtained from π by an RPI1 update.

First we upper-bound the expectation of $f(\pi')$. Since $f(\pi')$ is a non-negatively valued random variable, we can use the following expression for its expectation.

$$\begin{aligned}
\mathbb{E}[f(\pi')] &= \sum_{n+1 \ge s' \ge 1} \mathbb{P}\{f(\pi') \ge s'\} \\
&= \sum_{s \ge s' \ge 1} 1 + \sum_{n+1 \ge s' > s} \mathbb{P}\{f(\pi') \ge s'\} \\
&\le s + \sum_{n+1 \ge s' > s} \frac{1}{2^{s'-s}} \\
&\le s + \sum_{k=1}^{\infty} \frac{1}{2^k} \\
&= s + 1.
\end{aligned}$$

Now, assuming inductively that $N(s') \ge n + 2 - s'$ for $s < s' \le n+1$, we can lower-bound $N(s) = L(\pi)$ as

$$\begin{aligned}
N(s) = \quad & 1 + \sum_{\pi'' \in \Pi} L(\pi'')\mathbb{P}\{\pi' = \pi''\} \\
\ge \quad & 1 + \sum_{\pi'' \in \Pi} N(f(\pi''))\mathbb{P}\{\pi' = \pi''\} \\
= \quad & 1 + \sum_{n+1 \ge s' \ge 1} \left[\sum_{\pi'' \in \Pi, f(\pi'') = s'} N(s')\mathbb{P}\{\pi' = \pi''\} \right] \\
= \quad & 1 + \sum_{n+1 \ge s' \ge 1} N(s') \left[\sum_{\pi'' \in \Pi, f(\pi'') = s'} \mathbb{P}\{\pi' = \pi''\} \right] \\
= \quad & 1 + \sum_{n+1 \ge s' \ge 1} N(s')\mathbb{P}\{f(\pi') = s'\} \\
= \quad & 1 + \sum_{n+1 \ge s' \ge s} N(s')\mathbb{P}\{f(\pi') = s'\},
\end{aligned}$$

since $\mathbb{P}\{f(\pi') < s = f(\pi)\} = 0$. We rearrange terms in