A.3 Proof of Lemma 4.5

Proof. Let $C_{\ell+1}(u,v)$ denote the cost of one push operation through edge (u,v) from level ℓ to $\ell+1$. If $\frac{(1-\alpha)\hat{\pi}_\ell(v,t)}{d_{out}(u)} \geq \frac{\alpha\theta}{\lambda(u)}$, the push operation will be guaranteed once. Otherwise, the push happens with probability $\frac{\lambda(u)\cdot(1-\alpha)\hat{\pi}_\ell(v,t)}{\alpha\theta\cdot d_{out}(u)}$. Hence, we can derive the expectation of $C_{\ell+1}(u,v)$ conditioned on the value of all estimators $\{\hat{\pi}_\ell\}$ at ℓ-th level$\{\hat{\pi}_\ell\}$ that

$$\mathrm{E}\left[C_{\ell+1}(u,v) \mid \{\hat{\pi}_\ell\}\right] = \begin{cases} 1, & if \quad \frac{(1-\alpha)\hat{\pi}_\ell(v,t)}{d_{out}(u)} \geq \frac{\alpha\theta}{\lambda(u)}, \\ 1\cdot\frac{\lambda(u)\cdot(1-\alpha)\hat{\pi}_\ell(v,t)}{\alpha\theta\cdot d_{out}(u)}, & otherwise. \end{cases}$$

Note that if $\frac{(1-\alpha)\hat{\pi}_\ell(v,t)}{d_{out}(u)} \geq \frac{\alpha\theta}{\lambda(u)}$, the conditional expectation $\mathrm{E}\left[C_{\ell+1}(u,v) \mid \{\hat{\pi}_\ell\}\right]$ satisfies that $\mathrm{E}\left[C_{\ell+1}(u,v) \mid \{\hat{\pi}_\ell\}\right] = 1 \leq \frac{\lambda(u)\cdot(1-\alpha)\hat{\pi}_\ell(v,t)}{\alpha\theta\cdot d_{out}(u)}$. Thus, we can derive that $\mathrm{E}\left[C_{\ell+1}(u,v) \mid \{\hat{\pi}_\ell\}\right] \leq \frac{\lambda(u)\cdot(1-\alpha)\hat{\pi}_\ell(v,t)}{\alpha\theta\cdot d_{out}(u)}$ always holds. Applying the unbiasedness of $\hat{\pi}_\ell(v,t)$ according to Lemma 4.3, we have

$$\mathrm{E}\left[C_{\ell+1}(u,v)\right] = \mathrm{E}\left[\mathrm{E}\left[C_{\ell+1}(u,v) \mid \{\hat{\pi}_\ell\}\right]\right] \leq \frac{\lambda(u)\cdot(1-\alpha)\pi_\ell(v,t)}{\alpha\theta\cdot d_{out}(u)}.$$

Recall that C_{total} denotes the total cost in the whole process and $C_{total} = \sum_{i=1}^{L}\sum_{u\in V}\sum_{v\in N_{out}(u)} C_i(u,v)$. The expectation of C_{total} can be derived that

$$\mathrm{E}\left[C_{total}\right] = \sum_{i=1}^{L}\sum_{u\in V}\sum_{v\in N_{out}(u)} \mathrm{E}\left[C_i(u,v)\right]$$

$$\leq \sum_{i=1}^{\infty}\sum_{u\in V}\sum_{v\in N_{out}(u)} \frac{\lambda(u)\cdot(1-\alpha)\pi_{i-1}(v,t)}{\alpha\theta\cdot d_{out}(u)} = \frac{1}{\alpha\theta}\cdot\sum_{u\in V}\lambda(u)\cdot\sum_{i=1}^{\infty}\pi_i(u,t).$$

According to the property of ℓ-hop PPR that $\sum_{i=0}^{\infty}\pi_i(u,t) = \pi(u,t)$,

$$\mathrm{E}\left[C_{total}\right] \leq \frac{1}{\alpha\theta}\sum_{u\in V}\lambda(u)\cdot\pi(u,t),$$

which proves the lemma. □

A.4 Proof of Theorem 4.1

Proof. We first show that by truncating at the $L = \log_{1-\alpha}\theta$ hop, we only introduce an additive error of θ. More precisely, note that $\sum_{i=L+1}^{\infty}\alpha(1-\alpha)^i \leq (1-\alpha)^{(L+1)} \leq \theta$. By setting a θ that is significantly smaller than the relative error threshold δ or the additive error bound ε, we can accomodate the θ additive error without increasing the asymptotic query time.

According to Lemma 4.4, we have $\mathrm{Var}\left[\hat{\pi}_\ell(s,t)\right] \leq \theta\pi_\ell(s,t)$. By Chebyshev inequality, we have

$$\Pr\left[\left|\hat{\pi}_\ell(s,t)-\pi_\ell(s,t)\right| \geq \sqrt{3\theta\pi_\ell(s,t)}\right] \leq 1/3.$$

We claim that this variance implies an ε_r-relative error for all $\pi_\ell(s,t) \geq 3\theta/\varepsilon_r^2$. For a proof, note that $\theta \leq \varepsilon_r^2\pi_\ell(s,t)/3$ and consequently $\sqrt{3\theta\pi_\ell(s,t)} \leq \sqrt{\varepsilon_r^2\pi_\ell(s,t)^2} = \varepsilon_r\pi_\ell(s,t)$. It follows that $\Pr\left[\left|\hat{\pi}_\ell(s,t)-\pi_\ell(s,t)\right| \geq \varepsilon_r\pi_\ell(s,t)\right] \leq 1/3$ for all $\pi_\ell(s,t) \geq 3\theta/\varepsilon_r^2$. By setting $\theta = \frac{\varepsilon_r^2\delta}{3L}$, we obtain a constant relative error guarantee for all $\pi_\ell(s,t) \geq \delta/L$, and consequently a constant relative error for $\pi(s,t) \geq \delta$. To obtain a high probability result, we can apply the Median-of-Mean trick [12], which takes the median of $O(\log n)$ independent copies of $\hat{\pi}_\ell(s,t)$ as the final estimator to $\pi_\ell(s,t)$. This trick brought the failure probability from $1/3$ to $1/n^2$ by increasing the running time by a factor of $O(\log n)$. Applying the union bound to n source nodes $s \in V$ and $\ell = 0,\ldots,L$, the failure probability becomes $1/n$. Finally, by setting $\lambda(u) = 1$ in Lemma 4.5, we can rewrite the time cost as below.

$$\mathrm{E}\left[C_{total}\right] \leq \frac{1}{\alpha\theta}\sum_{u\in V}\lambda(u)\cdot\pi(u,t) = \frac{1}{\alpha\theta}\sum_{u\in V}\pi(u,t) = \frac{n\pi(t)}{\alpha\theta},$$

where $\pi(t)$ represents t's PageRank and $n\pi(t) = \sum_{u\in V}\pi_\ell(u,t)$ according to PPR's definition. By setting $\theta = \frac{\varepsilon_r^2\delta}{3L}$ and running $O(\log n)$ independent copies of Algorithm 2, the time complexity can be bounded by $O\left(\frac{n\pi(t)L\log n}{\alpha\varepsilon_r^2\delta}\right) = \tilde{O}\left(\frac{n\pi(t)}{\delta}\right)$. If we choose the target node t uniformly at random from set V, then $\mathrm{E}\left[\pi(t)\right] = \frac{1}{n}$, and the running time becomes $\tilde{O}\left(\frac{1}{\delta}\right)$.

□

A.5 Proof of Theorem 4.2

Proof. Applying Lemma 4.4, we have $\mathrm{Var}\left[\hat{\pi}_\ell(s,t)\right] \leq \alpha\theta^2$. Consequently, we have $\mathrm{Var}\left[\hat{\pi}(s,t)\right] = \mathrm{Var}\left[\sum_{\ell=0}^{L}\hat{\pi}_\ell(s,t)\right] \leq \alpha L\theta^2$. By Chebyshev's inequality, we have $\Pr\left[\left|\hat{\pi}(s,t)-\pi(s,t)\right| \geq \sqrt{3L\alpha}\theta\right] \leq 1/3$. By setting $\theta = \varepsilon/\sqrt{3L\alpha}$, it follows that $\hat{\pi}(s,t)$ is an ε additive error for all $\pi(s,t)$. Similar to the proof of Theorem 4.1, we can use the median of $O(\log n)$ independent copies of $\hat{\pi}(s,t)$ as the estimator to reduce the failure probability from $1/3$ to $1/n$ for all source nodes $s \in V$.

For the time cost, Lemma 4.5 implies that

$$\mathrm{E}\left[C_{total}\right] \leq \frac{1}{\alpha\theta}\sum_{u\in V}\lambda(u)\cdot\pi(u,t) = \frac{1}{\alpha\theta}\sum_{u\in V}\sqrt{d_{out}(u)}\cdot\pi(u,t).$$

Recall that we set $\theta = \varepsilon/\sqrt{3L\alpha}$ and run $O(\log n)$ independent copies of Algorithm 2, it follows that the running time can be bounded by $\tilde{O}\left(\frac{1}{\varepsilon}\sum_{u\in V}\sqrt{d_{out}(u)}\cdot\pi(u,t)\right)$. If t is chosen uniformly at random, we have $\sum_{t\in V}\pi(u,t) = 1$. Ignoring the $\tilde{O}$ notation, we have

$$\mathrm{E}\left[C_{total}\right] \leq \frac{1}{\varepsilon}\cdot\frac{1}{n}\cdot\sum_{t\in V}\sum_{u\in V}\sqrt{d_{out}(u)}\cdot\pi(u,t)$$

$$= \frac{1}{\varepsilon}\cdot\frac{1}{n}\cdot\sum_{u\in V}\sqrt{d_{out}(u)}\sum_{t\in V}\pi(u,t) = \frac{1}{\varepsilon}\cdot\frac{1}{n}\cdot\sum_{u\in V}\sqrt{d_{out}(u)}.$$

By the AM-GM inequality, we have $\frac{1}{n}\cdot\sum_{u\in V}\sqrt{d_{out}(u)} \leq \sqrt{\frac{\sum_{u\in V}d_{out}(u)}{n}} = \sqrt{\bar{d}}$, Hence, $\mathrm{E}\left[C_{total}\right] \leq \frac{1}{\varepsilon}\cdot\frac{1}{n}\cdot\sum_{u\in V}\sqrt{d_{out}(u)} \leq \frac{\sqrt{\bar{d}}}{\varepsilon}$, and the theorem follows. □