1 APPENDIX

Lemma 1. *Let S_t be a fixed set of nodes, and v be a fixed node. Suppose that we generate an RS set R for v in a possible world X where X is the modified possible world constructed from possible world X sampled from G. Let ω_R be the probability that S_t covers R with weight ω_R, and ρ_2 be the probability that S_t, when used as a seed set for campaign t, achieves a score ω_R at v in a propagation process on G w.r.t. $g(\cdot)$. Then, $\rho_1 = \rho_2$.*

$$Y_i(S) = \begin{cases} S \cap R_i(v) = \omega_{R_i(v)} & \text{if } S \text{ covers } R_i(v) \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Lemma 2. *Given a random RS set $R_i(v)$ generated with importance sampling rooted at v, for any set $S \subseteq V$, we have, $\sigma(S) = E[Y_i(S)] \cdot |P|$.*

Proof. Let $\eta_X(SF \to v) = 1$ be an indicator for if SF can reach v in possible world X. Let (1) $\eta_X(SF \to v) = 1$ if S can reach v's neighbor in X along a path, (2) $\eta_X(SF \to v) = 2$ if S can reach v in X along a path and (3) $\eta_X(SF \to v) = 0$ otherwise. When the root v of R_i is selected uniformly at random from RX_F , we have □

$$\begin{aligned} \sigma(S) =& \sum_v [2Pr_X(\eta_X(SF \to v) \wedge Pr_X(\eta_X(S| = v) \\ &+ \sum_v [Pr_X(\eta_X(SF \to v) \wedge Pr_X(\eta_X(S| = v) \\ =& \sum_v Pr_X(\eta_X(SF \to v) \cdot [2Pr_X(\eta_X(S| = v) = 2) \\ &+ Pr_X(\eta_X(S| = v) = 1)] \\ =& |P|_F \cdot [2Pr_{X,v \in RX_F}(S \cap R_i(v) = 2) \\ &+ Pr_{X,v \in RX_F}(S \cap R_i(v) = 1)] \\ =& |P|_F \cdot E_{X,v \in RX_F}[Yi(S)] \end{aligned} \tag{2}$$

Lemma 6 states that we can estimate the expected scores of truth spread using random RS sets generated with IS. Let R be a collection of θ random RS sets generated with IS and let $\omega_R(S)$ be the total weight of RS sets in R covered by a node set S. Then, based on Lemmas 4, 5 and 6, we can prove:

Corollary 1. $E[\frac{\omega_R(S)}{\theta}] \cdot |P| = \sigma(S)$

Next, we analyze the random variables associated with random RS sets generated using IS. In particular, we show they have smaller variances than random RS sets generated by RS and, as a consequence, fewer samples are required by our reverse sampling framework. Define the random variable $Z_i(S) = \frac{Y_i(S) \cdot |P|}{n}$. Notice that the means of $Y_i(S)$ and $Z_i(S)$ are $E[Y_i(S)] = \sigma(S) \cdot |P|$and

$E[Zi(S)] = E[Yi(S)] \cdot \frac{|P|}{n} = \frac{\sigma(s)}{n}$ respectively. If we construct a set of random variables $Z_1(S), ..., Z_\theta(S)$, observe that $\frac{\theta}{n}\sum_{i=1}^{\theta} Z_i(S)$ is an empirical estimate of $\sigma(S)$.

Proposition 1. $Var[Z_i(S)] \leq 2 \cdot \frac{\theta(S)}{n}\frac{|P|}{n}$

Proof. Define the following random variables:

$$Y_i^{(2)}(S) = \begin{cases} 1 & \text{if } S \text{ covers } R_i(v) \text{ with weight } 1 \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

$$Y_i^{(2)}(S) = \begin{cases} 1 & \text{if } S \text{ covers } R_i(v) \text{ with weight } 2 \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Then, we can re-write the random variable $Z_i(S)$ as $Z_i(S) = \frac{|P|}{n} \cdot (Y_i^{(1)} + 2Y_i^{(2)})$. As a result, we have

$$E[Z_i(S)] = \frac{|P|}{n} \cdot (Y_i^{(1)} + 2Y_i^{(2)}) \tag{5}$$

Now, we can bound the variance of $Z_i(S)$.

$$\begin{aligned} Var[Z_i(S)] = & Var[\frac{|P|}{n} \cdot (Y_i^{(1)} + 2Y_i^{(2)})] \\ = & \frac{|P|^2}{n^2}(Var[Y_i^{(1)}] + 4Var[Y_i^{(2)}] + 4Cov[Y_i^{(1)}, Y_i^{(2)}]) \\ \leq & \frac{|P|^2}{n^2}(E[Y_i^{(1)}] + 4E[Y_i^{(2)}]) \leq 2 \cdot \frac{\theta(S)}{n}\frac{|P|}{n} \end{aligned} \tag{6}$$

Notice that the mutual exclusion of the events described by $Y_i^{(1)}$ and $Y_i^{(2)}$ ensures that $Cov[Y_i^{(1)}, Y_i^{(2)}] \leq 0$. □

Definition 1. *(Martingale). A sequence of random variables $Y_1, Y_2, Y_3, ...$ is a martingale if and only if $E[|Y_i|] \leq +\||P|ty$ and $E[Y_i|Y_1, Y_2, ..., Y_{i-1}] = Y_{i-1}$ for any i.*

It is straightforward to show that the random variables $Z_i(S)$ form a martingale. Thus, the Chernoff bounds for martingales given in Lemma 2 let us derive the following concentration bounds for the random variables $Z_i(S)$ associated with our RS sets generated with IS by plugging in the variance derived above. Note, we assume that $|P| \leq \frac{n}{2}$ as a necessary boundary condition for the derivation of our concentration bounds. Further, in our experiments the condition always held since an unrealistic number of mis|P|ormation seeds would be required to |P|luence over half the networks considered.

Lemma 3. *Given a fixed collection of θ RS sets R constructed with importance sampling and a seed set S, let $\Lambda(S) = \frac{|P|}{n}\omega R(S)$ be the normalized weighted coverage of S in R. For any $\lambda > 0$ we have,*

$$Pr[\Lambda(S) - \sigma(S) \cdot \frac{\theta}{n} \geq \lambda] \leq exp(\frac{-\lambda^2}{\frac{2}{3} + 4\sigma(S)\frac{\theta}{n}\frac{|P|}{n}}) \tag{7}$$

$$Pr[\Lambda(S) - \sigma(S) \cdot \frac{\theta}{n} \leq -\lambda] \leq exp(\frac{-\lambda^2}{4\sigma(S)\frac{\theta}{n}\frac{|P|}{n}}) \tag{8}$$

Proof. We begin by recalling the concentration bounds for martingales. Let $M1, M2, ...$ be a martingale such that $|M_1| \leq a, |M_j - M_{j-1}| \leq a \forall j \in [2, \theta]$ and
$Var[M_1] + \sum_{j=2}^{\theta} Var[M_j|M_1, M_2, ..., M_{j-1}] \leq b$
Where $Var[\cdot]$ denotes the variance of a random variable. Then, for any $\lambda > 0$,

$$Pr[M_\theta - E[M_\theta] \geq \lambda] \leq exp(\frac{-\lambda^2}{\frac{2}{3}a\lambda + 2b}) \tag{9}$$

and,

$$Pr[M_\theta - E[M_\theta] \leq -\lambda] \leq exp(\frac{-\lambda^2}{2b}) \tag{10}$$

Next, we apply the above concentration bounds on the martingale formed by the random variable $Z_i(S)$ associated with our RS sets generated with importance sampling. First, since each RS set R_i is generated randomly and independently of all the prior RS sets, we have

$$E[Z_i(S)|Z_1(S), Z_2(S), ..., Z_{i-1}(S)] = E[Z_i(S)] = \frac{\sigma(S)}{n}. \tag{11}$$

Let $p = \frac{\sigma(S)}{n}$ and $M_i = \sum_{j=1}^{i}(Z_j(S) - p)$. Then, we have $E[Mi] = 0$ and $E[M_i|M_1, M_2, ..., M_{i-1}] = Mi - 1$
Therefore, $M_1, M_2, ..., M_\theta$ is a martingale. We have $|M_1| \leq \frac{2|P|}{n} \leq 1$ assuming $|P| \leq \frac{n}{2}$ and$|M_j - M_{j-1}| \leq 1 \forall j \in [2, \theta]$. We also have

$$\begin{aligned} Var[M_1] + \sum_{j=2}^{\theta} Var[M_j|M_1, M_2, ..., M_{j-1}] \\ = \sum_{j=1}^{\theta} Var[Z_j(S)] \leq 2\theta \cdot \frac{\sigma(S)}{n}\frac{|P|}{n} \end{aligned} \tag{12}$$

□

Lemma 4. *Let $\delta_1 \in (0, 1), o\epsilon_1 > 0$ and*

$$\theta_1 = \frac{4|P|log(\frac{1}{\delta_1})}{\epsilon_1^2 OPT} \tag{13}$$

If $\theta \geq \theta_1$, then $n \cdot F_R(S^o) \geq (1 - \epsilon_1) \cdot OPT$ holds with at least $1 - \delta_1$ probability.

Proof. Let $p = E[F_R(S^o)]$. By Lemma 6 and Equation 7, $p = E[F_R(S^o)] = \sigma(S^o)n = OPTn$.
By Lemma 7,

$$\begin{aligned}
&Pr[n \cdot F_R(S^o) \leq (1-\epsilon_1) \cdot OPT] \\
&\quad = Pr[n \cdot F_R(S^o) \leq (1-\epsilon_1) \cdot np] \\
&\quad = Pr[\theta \cdot F_R(S^o) \leq (1-\epsilon_1) \cdot \theta p] \\
&\quad \leq Pr[\theta \cdot F_R(S^o) \leq (1-\epsilon_1) \cdot \theta p] \\
&\quad = Pr[\sum_i Z_i(S^o) - \theta p \leq \epsilon_1 \theta p] \\
&\quad = Pr[\sum_i Z_i(S^o) - \sigma(S^o)\frac{\theta}{n} \leq \epsilon_1 \sigma(S^o)\frac{\theta}{n}] \\
&\quad \leq exp(\frac{(-\epsilon_1^2 \sigma(S^o)^2 (\frac{\theta}{n})^2}{4\sigma(S^o)\frac{\theta}{n}\frac{|P|}{n}}) \\
&\quad = exp(\frac{-\epsilon_1^2 \sigma(S^o)\theta}{4|P|}) \\
&\quad \leq \delta_1
\end{aligned} \tag{14}$$

Thus, the lemma is proved. □

Suppose that $nF_R(S^o) \geq (1-\epsilon_1) \cdot OPT$ holds, by the properties of the greedy approach,

$$\begin{aligned}
n \cdot F_R(S^\star) &\geq (1-1/e)n \cdot F_R(S^o) \\
&\geq (1-1/e)(1-\epsilon_1) \cdot OPT
\end{aligned} \tag{15}$$

Intuitively, this indicates that the expected mitigation of $S^\star$ is likely to be large, since $n \cdot F_R(S^\star)$ is an indicator of $\sigma(S^\star)$. This is formalized in the following lemma.

Lemma 5. *Let $\delta_2 \in (0,1), \epsilon_1 < \epsilon$ and*

$$\theta_2 = \frac{2(1-1/e) \cdot nlog(\frac{\binom{n}{k}}{\delta_2})}{(\epsilon - (1-1/e\epsilon'))^2 \cdot OPT} \tag{16}$$

. If Equation 14 holds and $\theta \geq \theta_2$, then with at least $1-\delta_2$ probability, $\sigma(S^\star) \geq (1-1/e-\epsilon) \cdot OPT$.

Proof. Let S be an arbitrary size-k seed set. We say S is bad if $\sigma(S) < (1-1/e-\epsilon) \cdot OPT$. To prove the lemma, we show that each bad size-k seed set has at most $\frac{\delta_2}{\binom{n}{k}}$ probability to be returned by applying the greedy algorithm to a collection of θ_2 RS sets. This suffices to establish the lemma because (1) there exist only $\binom{n}{k}$ bad size-k seed sets, and (2) if each of them has at most $\frac{\delta_2}{\binom{n}{k}}$ probability to be returned, then by the union bound, there is at least $1-\delta_2$ probability that none of them is output by the greedy algorithm. Consider any bad size-k seed set S. Let $p = E[F_R(S)] = \sigma(S)/n$. We have

$$\begin{aligned}
&Pr[n \cdot F_R(S) - \sigma(S) \geq \epsilon_2] \\
&= Pr[n \cdot F_R(S) \geq \sigma(S) + \epsilon_2 \cdot OPT] \\
&\geq Pr[n \cdot F_R(S) \geq (1-1/e)OPT + \epsilon_2 \cdot OPT] \\
&= Pr[n \cdot F_R(S) \geq (1-1/e+\epsilon) + \epsilon_2) \cdot OPT]
\end{aligned} \tag{17}$$

We set ϵ_2 such that the multiplicative factor of OPT in Equation 16 is equal to the one in Equation 14.

$$\epsilon_2 = (\epsilon - (1-1/e\epsilon_1))^2 \tag{18}$$

Then, we can apply our Chernoff bounds by re-writing $Pr[n \cdot F_R(S) - \sigma(S) \geq \epsilon_2 \cdot OPT] = Pr[\theta \cdot F_R(S) - \sigma(S)\frac{\theta}{n} \geq 2\sigma(S) \cdot \frac{\theta}{n} \cdot OPT]$ and letting $\lambda = \sigma(S)2 \cdot \frac{\theta}{n} \cdot OPT$ to get

$$\begin{aligned}
&Pr[n \cdot F_R(S) - \sigma(S) \geq \epsilon_2] \\
&= Pr[\theta \cdot F_R(S) - \sigma(S) \cdot \frac{\theta}{n} \geq \epsilon_2 \cdot \frac{\theta}{n} \cdot OPT] \\
&\leq exp(\frac{-\epsilon_2^2 \cdot \frac{\theta^2}{n^2} \cdot OPT}{\frac{2}{3}\epsilon_2 \cdot \frac{\theta}{n} \cdot OPT + 4\theta\frac{\sigma(S)}{n} \cdot \frac{|P|}{n}}) \\
&\leq exp(\frac{-\epsilon_2^2 \cdot \theta^2 \cdot OPT}{\frac{2}{3}\epsilon_2 \cdot \theta n \cdot OPT + 4\theta \cdot (1-1/e-\epsilon) \cdot OPT \cdot |P|}) \\
&\leq exp(-\frac{(\epsilon - (1-1/e\epsilon_1))^2 \cdot \theta \cdot OPT}{2(1-1/e) \cdot n}) \\
&\leq exp(-\frac{(\epsilon - (1-1/e\epsilon_1))^2 \cdot \theta_2 \cdot OPT}{2(1-1/e) \cdot n}) \\
&\leq \frac{\delta_2}{\binom{n}{k}}
\end{aligned} \tag{19}$$

We can derive the third inequality through $\frac{2|P|}{n} \leq 1$ □

Theorem 1. *Given any ϵ_1 such that $\epsilon_1 < \epsilon$ and any $\delta_1, \delta_2 \in (0,1)$ with $\delta_1 + \delta_2 \leq \delta$, setting $\theta \geq argmax\{\theta_1, \theta_2\}$ ensures that the node selection phase of IMM returns a $(1-1/e-\epsilon)$-approximate solution with at least $1-\delta$ probability.*

Proof. By Lemma 9,$\sigma(S^\star) \geq (1-1/e-\epsilon) \cdot OPT$ holds with at least $1-\delta_2$ probability under the condition that Equation 14 holds. And by Lemma 8, Equation 14 holds with at least $1-\delta_1$ probability. By the union bound, $\sigma(S^\star) \geq (1-1/e-\epsilon) \cdot OPT$ holds with at least $1-\delta_1-\delta_2 \geq 1-\delta$ probability. Thus, the theorem is proved. □

Now, a natural question is, how should we select ϵ_1, δ_1, and δ_2 that adhere to the conditions of Theorem 2 in order to minimize θ? Assume that OPT is known. We are trying to minimize $\theta^\star = argmax\{\theta_1, \theta_2\}$ subject to $\delta_1 + \delta_2 \leq \delta$.

Following an approach similar to that of, we set $\delta_1 = \delta_2 = \frac{1}{2}$ and set $\theta_1 = \theta_2$ to derive an approximately minimal value for $\theta^\star$.

$$\theta^\star \leq \frac{8n(1-1/e)[ln\frac{2}{\delta} + ln\binom{n}{k}]}{OPT\epsilon^2} \tag{20}$$

Lemma 6. *Let R be a collection of random RS sets and $S^\star$ be a size-k seed set generated by applying the greedy algorithm on R. For fixed ϵ, and δ, and*

$$|R| \geq \frac{8n(1-1/e)[ln\frac{2}{\delta} + ln\binom{n}{k}]}{OPT\epsilon^2} \tag{21}$$

then $S^\star$ is a $(1-1/e-\epsilon)$-approximate solution with at least $1-\delta$ probability.

According to lemmma 10, we can acquire the N_{max} in algorithm 1:

$$N_{max} \geq \frac{8n(1-1/e)[ln\frac{2}{\delta} + ln\binom{n}{k}]}{LB\epsilon^2} \tag{22}$$

which is an upper bound on the number of RS sets needed to guarantee a $(1-1/e-\epsilon)$ approximation with at least $1-\delta$ probability when $LB \leq OPT$ is a lower bound of the optimal mitigation, we derive the LB from [**?**]. Then, we can derive the $\sigma^l(S^\star)$ and $\sigma^u(S^\star)$ through the similar means in [**?**].

Lemma 7. *For any $\delta \in$ (0,1),we have*

$$Pr[\sigma(S^\star) \geq ((\sqrt{\Lambda_2(S^\star) + \frac{25a}{36}} - \sqrt{a})^2 - \frac{a}{36})\frac{n}{\theta_2}] \geq 1-\delta \tag{23}$$

where $a = \ln(\frac{1}{\delta})$.And,

$$
\begin{aligned}
&\Pr[\sigma(S^*) < ((\sqrt{\varLambda^2(S^*) + \frac{25a}{36}} - \sqrt{a})^2 - \frac{a}{36})\frac{n}{\theta_2}] \\
\leq &\Pr[\sigma(S^*) < ((\sqrt{\varLambda^2(S^*) + \frac{25a}{36}} - \sqrt{a})^2 - \frac{a}{36})\frac{n}{\theta_2}] \\
= &\Pr[\sigma(S^*) \cdot \frac{\theta_2}{n} < (\sqrt{\varLambda^2(S^*) + \frac{25a}{36}} - \sqrt{a})^2 - \frac{a}{36}] \\
= &\Pr[b + \frac{a}{36} < (\sqrt{\varLambda^2(S^*) + \frac{25a}{36}} - \sqrt{a})^2] \\
= &\Pr[\sqrt{b + \frac{a}{36}} < \sqrt{\varLambda^2(S^*) + \frac{25a}{36}} - \sqrt{a}] \\
&+ \Pr[\sqrt{b + \frac{a}{36}} < \sqrt{a} - \sqrt{\varLambda^2(S^*) + \frac{25a}{36}}] \\
= &\Pr[\sqrt{b + \frac{a}{36}} < \sqrt{\varLambda^2(S^*) + \frac{25a}{36}} - \sqrt{a}] \\
= &\Pr[\sqrt{b + \frac{a}{36}} + \sqrt{a} < \sqrt{\varLambda^2(S^*) + \frac{25a}{36}}] \\
= &\Pr[b + \sqrt{4ab + \frac{a^2}{9}} + \frac{a}{3} < \varLambda^2(S^*)] \\
= &\Pr[\sqrt{4ab + \frac{a^2}{9}} + \frac{a}{3} < \varLambda^2(S^*) - b] \\
= &\Pr[\sqrt{4ab + \frac{a^2}{9}} + \frac{a}{3} < \varLambda^2(S^*) - \sigma(S^*) \cdot \frac{\theta_2}{n}] \\
\leq &\Pr[\sqrt{4ab + \frac{a^2}{9}} + \frac{a}{3} < \varLambda^2(S^*) - \sigma(S^*) \cdot \frac{\theta_2}{n}] \\
\leq &\exp(-\frac{(\sqrt{4ab + \frac{a^2}{9}} + \frac{a}{3})^2}{4\sigma(S^*) \cdot \frac{\theta_2}{n}\frac{|P|}{n} + \frac{2}{3}(\sqrt{4ab + \frac{a^2}{9}} + \frac{a}{3})}) \\
\leq &\exp(-\frac{(\sqrt{4ab + \frac{a^2}{9}} + \frac{a}{3})^2}{4\sigma(S^*) \cdot \frac{\theta_2}{n} + \frac{2}{3}(\sqrt{4ab + \frac{a^2}{9}} + \frac{a}{3})}) \\
= &\exp(-a) = \delta
\end{aligned}
$$

Lemma 8. *For any $\delta \in$ (0,1),we have*

$$
Pr[\sigma(S^o) \leq (\sqrt{\frac{\varLambda_1(S^\star)}{\frac{1}{e}} + a} + \sqrt{a})^2 \frac{n}{\theta_2}] \geq 1 - \delta \tag{24}
$$

where $a = \ln(\frac{1}{\delta})$,

$$
\begin{aligned}
&\Pr[\sigma(S^o) > (\sqrt{\frac{\Lambda_1(S^*)}{1-1/e} + a} + \sqrt{a})^2 \frac{n}{\theta_1}] \\
\leq &\Pr[\sigma(S^o) > (\sqrt{\Lambda_1(S^o) + a} + \sqrt{a})^2 \frac{n}{\theta_1}] \\
= &\Pr[\sigma(S^o) \cdot \frac{\theta_1}{n} > (\sqrt{\Lambda_1(S^o) + a} + \sqrt{a})^2] \\
= &\Pr[b > (\sqrt{\Lambda_1(S^o) + a} + \sqrt{a})^2] \\
= &\Pr[\sqrt{b} > \sqrt{\Lambda_1(S^o) + a} + \sqrt{a}] \\
= &\Pr[\sqrt{b} - \sqrt{a} > \sqrt{\Lambda_1(S^o) + a}] \\
= &\Pr[b + a - 2\sqrt{ab} > \Lambda_1(S^o) + a] \\
= &\Pr[-\sqrt{4ab} > \Lambda_1(S^o) - b] \\
= &\Pr[-\sqrt{4ab} > \Lambda_1(S^o) - \sigma(S^o) \cdot \frac{\theta_1}{n}] \\
\leq &\Pr[-\sqrt{4ab} > \Lambda_1(S^o) - \sigma(S^o) \cdot \frac{\theta_1}{n}] \\
\leq &\exp(\frac{-4a\sigma(S^o) \cdot \frac{\theta_1}{n}}{4\sigma(S^o) \cdot \frac{\theta_1}{n} \frac{|P|}{n}}) \\
= &\exp(\frac{-a}{\frac{|P|}{n}}) \\
\leq &\exp(-a) = \delta
\end{aligned}
$$

The reason NATS ensures a $(1 - 1/e - \epsilon)$ approximation with at least $1 - \delta$ probability is as follows. First, the algorithm has at most i_{max} iterations. In each of the first $i_{max} - 1$ iterations, Algorithm 1 generates a size-k seed set $S^\star$ and derives $\sigma_l(S^\star)$ and $\sigma_u(S^o)$ from $R2$ and $R1$, respectively, setting $\delta_1 = \delta_2 = \frac{\delta}{3}$. Then, it computes $\alpha \leftarrow \sigma_l(S^\star)/\sigma_u(S^o)$ as the approximation guarantee of $S^\star$. By Lemmas 11 and 12, α is correct with at least $1 - 2\delta/(3i_{max})$ probability. By the union bound, it has at most $\frac{2\delta}{3}$ probability to output an incorrect solution in the first $i_{max} - 1$ iterations. Meanwhile, in the last iteration, it returns a seed set $S^\star$ generated by applying the greedy algorithm on R_1, with $|R1| \geq N_{max}$. By Equation 19, this ensures that $S^\star$ is an $(1 - 1/e - \epsilon)$-approximation with at least $1 - \frac{\delta}{3}$ probability. Therefore, the probability that IMM returns an incorrect solution is at most $\frac{2\delta}{3} + \frac{\delta}{3} = \delta$ leading to the following result regarding Algorithm 1.

Theorem 2. *Algorithm 1 returns a $(1 - 1/e - \epsilon)$-approximate solution for $\underline{\mu}$ as well as $\overline{\mu}$ with at least $1 - \delta$ probability.*