APPENDIX

Proof of Theorem 5.6. The primary time overhead of LSBM lies in (i) estimating the influence of seeds of misinformation; (ii) generating the CP sequences; (iii) executing the Max-Coverage algorithm and computing $\sigma^L(B_L)$ and $\sigma^U(B_L^o)$ in all iterations.

As shown in [45], the time complexity of influence estimation is $O(\frac{m\cdot\ln 1/\delta}{\beta^2})$. Then. we analyze the number of CP sequences generated by LSBM.

Let $\epsilon_1 = \epsilon$, $\tilde{\epsilon}_1 = \epsilon/e$, $\hat{\epsilon}_1 = \sqrt{\frac{2a_3\mathbb{E}[I_G(s)]}{D_s^L(B_L^o)\theta_1}}$, $\epsilon_2 = \sqrt{\frac{2a_3\mathbb{E}[I_G(s)]}{D_s^L(B_L)\theta_2}}$, $\tilde{\epsilon}_2 = (\sqrt{\frac{2a_3D_s^L(B_L)\theta_2}{\mathbb{E}[I_G(s)]}+\frac{a_3^2}{9}}+\frac{a_3}{3})\cdot\frac{\mathbb{E}[I_G(s)]}{D_s^L(B_L)\theta_2}$, $a_3 = c\ln(\frac{3i_{\max}}{\delta})$ for any $c \geq 1$. In addition, let $\theta_a = \frac{2\mathbb{E}[I_G(s)]\ln\frac{6}{\delta}}{(1-1/e-\epsilon)\epsilon_1^2D_s^L(B_L^o)}$, $\theta_b = \frac{(2+2\tilde{\epsilon}_1/3)\mathbb{E}[I_G(s)]\ln\frac{6\binom{n-|S|}{k}}{\delta}}{\tilde{\epsilon}_1^2D_s^L(B_L^o)}$, $\theta_c = \frac{27\mathbb{E}[I_G(s)]\ln\frac{3i_{\max}}{\delta}(1+\beta)^2}{(1-1/e-\epsilon)(\epsilon_1+\epsilon_1\beta-2\beta)^2D_s^L(B_L^o)}$, $\theta_d = \frac{5\ln\frac{3i_{\max}}{\delta}\mathbb{E}[I_G(s)]}{18(1-\epsilon_1)(1-1/e-\epsilon)D_s^L(B_L^o)}$ and $\theta' = \max\{\theta_a,\theta_b,\theta_c,\theta_d\}$. It is easy to verify that

$$\theta' = O\left(\frac{(k\ln(n-|S|)+\ln 1/\delta)\mathbb{E}[I_G(s)]}{(\epsilon+\epsilon\beta-2\beta)^2D_s^L(B_L^o)}\right). \tag{18}$$

When $\theta_1 = \theta_2 = c\theta'$, based on Eq. (8) and Eq. (9), we have:

$$\Pr[\frac{Cov_{\mathbb{C}_1^s}(B_L^o)}{\theta_1} < (1-\epsilon_1)\cdot D_s^L(B_L^o)] \leq \left(\frac{\delta}{6}\right)^c, \tag{19}$$

$$\Pr[\frac{Cov_{\mathbb{C}_2^s}(B_L)}{\theta_2} < (1-\epsilon_1)\cdot D_s^L(B_L)] \leq \left(\frac{\delta}{6}\right)^c, \tag{20}$$

$$\Pr[\frac{Cov_{\mathbb{C}_1^s}(B_L)}{\theta_1} > D_s^L(B_L)+\tilde{\epsilon}_1\cdot D_s^L(B_L^o)] \leq \left(\frac{\delta}{6\binom{n-|S|}{k}}\right)^c, \tag{21}$$

$$\Pr[\frac{Cov_{\mathbb{C}_1^s}(B_L^o)}{\theta_1} < (1-\hat{\epsilon}_1)\cdot D_s^L(B_L^o)] \leq \left(\frac{\delta}{3i_{\max}}\right)^c, \tag{22}$$

$$\Pr[\frac{Cov_{\mathbb{C}_2^s}(B_L)}{\theta_2} < (1-\epsilon_2)\cdot D_s^L(B_L)] \leq \left(\frac{\delta}{3i_{\max}}\right)^c, \tag{23}$$

$$\Pr[\frac{Cov_{\mathbb{C}_2^s}(B_L)}{\theta_2} > (1+\tilde{\epsilon}_2)\cdot D_s^L(B_L)] \leq \left(\frac{\delta}{3i_{\max}}\right)^c. \tag{24}$$

Specially, when $\theta_1 \geq c\theta_a$, Eq. (19) holds; when $\theta_1 \geq c\theta_b$, Eq. (21) holds. Eq. (22)-Eq. (24) are obtained based on the definition of $\hat{\epsilon}_1, \epsilon_2$ and $\tilde{\epsilon}_2$. When the event in Eq. (19) and Eq. (21) not happen, we get:

$$D_s^L(B_L) \geq (1-1/e-\epsilon)D_s^L(B_L^o). \tag{25}$$

Based on Eq. (25), when when $\theta_1 \geq c\theta_a$, Eq. (20) holds. Since B_L is not independent of $\mathbb{C}_1^s$ and there are at most $\binom{n-|S|}{k}$ blocker sets, based on the union bound, the probability that none of the events in Eq. (19)-Eq. (24) happens is at least:

$$1-\left((\frac{\delta}{6})^c\cdot 2+(\frac{\delta}{6\binom{n-|S|}{k}})^c\cdot\binom{n-|S|}{k}+i_{\max}\cdot(\frac{\delta}{3i_{\max}})^c\right) \geq 1-\delta^c.$$

And we have:

$$\hat{\epsilon}_1 \leq \sqrt{\frac{2(1-1/e-\epsilon)(\epsilon_1+\epsilon_1\beta-2\beta)^2}{27(1+\beta)^2}} \leq \frac{\epsilon_1+\epsilon_1\beta-2\beta}{3(1+\beta)},$$

$$\epsilon_2 \leq \sqrt{\frac{2(1-1/e-\epsilon)D_s^L(B_L^o)(\epsilon_1+\epsilon_1\beta-2\beta)^2}{27(1+\beta)^2D_s^L(B_L)}} \leq \frac{\epsilon_1+\epsilon_1\beta-2\beta}{3(1+\beta)},$$

$$\tilde{\epsilon}_2 \leq \sqrt{\frac{(\epsilon_1+\epsilon_1\beta-2\beta)^2(2+2\tilde{\epsilon}_2/3)}{27(1+\beta)^2}} \leq \frac{\epsilon_1+\epsilon_1\beta-2\beta}{3(1+\beta)}.$$

In addition, when the event in Eq. (22) not happen, we have:

$$\left(\sqrt{\frac{Cov_{\mathbb{C}_1^s}(B_L^o)\cdot(1+\beta)}{\hat{I}_G(s)}+\frac{a_3}{2}}+\sqrt{\frac{a_3}{2}}\right)^2\cdot\frac{1}{\theta_1} \geq \frac{D_s^L(B_L^o)}{\mathbb{E}[I_G(s)]}.$$

Thus, it holds that:

$$\begin{aligned} 1-\hat{\epsilon}_1 &= 1-\sqrt{\frac{2a_3\mathbb{E}[I_G(s)]}{D_s^L(B_L^o)\theta_1}} \\ &\leq 1-\frac{\sqrt{2a_3}}{\sqrt{Cov_{\mathbb{C}_1^s}(B_L^o)\cdot(1+\beta)/\hat{I}_G(s)+\frac{a_3}{2}}+\sqrt{\frac{a_3}{2}}} \\ &\leq \frac{Cov^u_{\mathbb{C}_1^s}(B_L^o)\cdot(1+\beta)/\hat{I}_G(s)}{(\sqrt{Cov^u_{\mathbb{C}_1^s}(B_L^o)\cdot(1+\beta)/\hat{I}_G(s)+\frac{a_3}{2}}+\sqrt{\frac{a_3}{2}})^2}. \end{aligned}$$

Since $a_2 = \ln(\frac{3i_{\max}}{\delta}) \leq a_3$, based on Line 19 of Algorithm 4, thus

$$\begin{aligned} \sigma^U(B_L^o) &\leq \left(\sqrt{\frac{Cov_{\mathbb{C}_1^s}(B_L^o)\cdot(1+\beta)}{\hat{I}_G(s)}+\frac{a_3}{2}}+\sqrt{\frac{a_3}{2}}\right)^2\cdot\frac{1}{\theta_1} \\ &\leq \frac{Cov^u_{\mathbb{C}_1^s}(B_L^o)\cdot(1+\beta)/\hat{I}_G(s)}{1-\hat{\epsilon}_1}\cdot\frac{1}{\theta_1}. \end{aligned} \tag{26}$$

When $\theta_2 \geq \theta_d$ and according to Eq. (20), we have:

$$\frac{Cov_{\mathbb{C}_2^s}(B_L)}{\mathbb{E}[I_G(s)]} \geq \frac{\theta_2\cdot(1-\epsilon_1)D_s^L(B_L)}{\mathbb{E}[I_G(s)]} \geq \frac{5a_1}{18}.$$

Thus, $f(x) = (\sqrt{x+\frac{2a_1}{9}}-\sqrt{\frac{a_1}{2}})^2-\frac{a_1}{18})$ monotonically increasing. In addition, when the event in Eq. (24) does not happen, we have:

$$\left((\sqrt{\frac{Cov_{\mathbb{C}_2^s}(B_L)\cdot(1-\beta)}{\hat{I}_G(s)}+\frac{2a_3}{9}}-\sqrt{\frac{a_3}{2}})^2-\frac{a_3}{18}\right)\cdot\frac{1}{\theta_2} \leq \frac{D_s^L(B_L)}{\mathbb{E}[I_G(s)]}.$$

Thus, it holds that:

$$\begin{aligned} &\frac{Cov_{\mathbb{C}_2^s}(B_L)\cdot(1-\beta)}{\hat{I}_G(s)}-\frac{\tilde{\epsilon}_2D_s^L(B_L)\cdot\theta_2}{\mathbb{E}[I_G(s)]} \\ =&\frac{Cov_{\mathbb{C}_2^s}(B_L)\cdot(1-\beta)}{\hat{I}_G(s)}-(\sqrt{2a_3\frac{D_s^L(B_L)\cdot\theta_2}{\mathbb{E}[I_G(s)]}+\frac{a_3^2}{9}}+\frac{a_3}{3}) \\ \leq&\frac{Cov_{\mathbb{C}_2^s}(B_L)\cdot(1-\beta)}{\hat{I}_G(s)}-(\sqrt{\frac{2Cov_{\mathbb{C}_2^s}(B_L)(1-\beta)a_3}{\hat{I}_G(s)}+\frac{4a_3^2}{9}}-\frac{2a_3}{3}) \\ =&\left(\sqrt{\frac{Cov_{\mathbb{C}_2^s}(B_L)\cdot(1-\beta)}{\hat{I}_G(s)}+\frac{2a_3}{9}}-\sqrt{\frac{a_3}{2}}\right)^2-\frac{a_3}{18}. \end{aligned}$$

Since $a_1 \leq a_3$, based on the Line 14 of Algorithm 4, thus

$$\sigma^L(B_L) \geq \frac{Cov_{\mathbb{C}_2^s}(B_L) \cdot (1-\beta)}{\hat{I}_G(s) \cdot \theta_2} - \frac{\tilde{\epsilon}_2 D_s^L(B_L)}{\mathbb{E}[I_G(s)]}. \tag{27}$$

Putting Eq. (26) and Eq. (27) together, when none of the events in Eq. (19)-Eq. (24) happens, we have:

$$\begin{aligned}
\frac{\sigma^L(B_L)}{\sigma^U(B_L^o)} &\geq \frac{\frac{Cov_{\mathbb{C}_2^s}(B_L)\cdot(1-\beta)}{\hat{I}_G(s)\cdot\theta_2} - \frac{\tilde{\epsilon}_2 D_s^L(B_L)}{\mathbb{E}[I_G(s)]}}{\frac{Cov^u_{\mathbb{C}_1^s}(B_L^o)\cdot(1+\beta)/\hat{I}_G(s)}{1-\hat{\epsilon}_1}\cdot\frac{1}{\theta_1}} \\
&\geq \frac{\theta_1\left(\frac{Cov_{\mathbb{C}_2^s}(B_L)\cdot(1-\beta)}{(1+\beta)\cdot\theta_2} - \tilde{\epsilon}_2 \cdot D_s^L(B_L)\right)(1-\hat{\epsilon}_1)}{Cov^u_{\mathbb{C}_1^s}(B_L^o)} \\
&\geq \frac{\theta_1\left(\frac{1-\beta}{1+\beta}\cdot(1-\epsilon_2) - \tilde{\epsilon}_2\right)\cdot D_s^L(B_L)(1-\hat{\epsilon}_1)}{Cov^u_{\mathbb{C}_1^s}(B_L^o)} \\
&\geq \frac{\theta_1\left(\frac{1-\beta}{1+\beta}\cdot(1-\epsilon_2) - \tilde{\epsilon}_2\right)\cdot D_s^L(B_L)(1-\hat{\epsilon}_1)}{Cov_{\mathbb{C}_1^s}(B_L)}(1-1/e) \\
&\geq \frac{\theta_1\left(1-\epsilon_2-\frac{2\beta}{1+\beta} - \tilde{\epsilon}_2 - \hat{\epsilon}_1\right)\cdot D_s^L(B_L)}{Cov_{\mathbb{C}_1^s}(B_L)}(1-1/e) \\
&\geq \frac{\theta_1\,(1-\epsilon_1)\cdot D_s^L(B_L)}{Cov_{\mathbb{C}_1^s}(B_L)}(1-1/e) \\
&\geq \frac{\theta_1\,(1-\epsilon_1)\cdot\left(Cov_{\mathbb{C}_1^s}(B_L)/\theta_1 - \tilde{\epsilon}_1\cdot D_s^L(B_L^o)\right)}{Cov_{\mathbb{C}_1^s}(B_L)}(1-1/e) \\
&\geq \frac{(1-\epsilon_1)\cdot\left(Cov_{\mathbb{C}_1^s}(B_L) - \tilde{\epsilon}_1\cdot\frac{Cov_{\mathbb{C}_1^s}(B_L^o)}{(1-\epsilon_1)}\right)}{Cov_{\mathbb{C}_1^s}(B_L)}(1-1/e) \\
&\geq (1-\epsilon_1)\left(1-\frac{\tilde{\epsilon}_1}{(1-\epsilon_1)(1-1/e)}\right)(1-1/e) \\
&= 1-1/e-\epsilon.
\end{aligned}$$

Therefore, when $\theta_1 = \theta_2 = c\theta'$ CP sequences are generated, LSBM does not stop only if at least one of the events in Eq. (19)-Eq. (24) happens. The probability is at most δ^c.

Let j be the first iteration in which the number of CP sequences generated by LSBM reaches θ'. From this iteration onward, the expected number of CP sequences further generated is at most

$$\begin{aligned}
2\cdot\sum_{z\geq j}\theta_0\cdot 2^z\cdot\delta^{2^{z-j}} &= 2\cdot 2^j\cdot\theta_0\sum_{z=0}2^z\cdot\delta^{2^z} \\
&\leq 4\theta'\sum_{z=0}2^{-2^z+z} \\
&\leq 4\theta'\sum_{z=0}2^{-z} \leq 8\theta'.
\end{aligned}$$

If the algorithm stops before this iteration, there are at most $2\theta'$ CP sequences generated. Therefore, the expected number of CP sequences generated is less than $10\theta'$, which is

$$O\left(\frac{(k\ln(n-|S|)+\ln 1/\delta)\mathbb{E}[I_G(s)]}{(\epsilon+\epsilon\beta-2\beta)^2 D_s^L(B_L^o)}\right). \tag{28}$$

We have shown that the expected time required to generate a CP sequence is $O(m\cdot\alpha(m,n))$. Based on Wald's equation [38], LSBM requires $O(\frac{(k\ln(n-|S|)+\ln 1/\delta)\mathbb{E}[I_G(s)]m\cdot\alpha(m,n)}{(\epsilon+\epsilon\beta-2\beta)^2 D_s^L(B_L^o)})$ in CP sequences generation. In addition, the total expected time used for executing the Max-Coverage and computing $\sigma^L(B_L)$ and $\sigma^U(B_L^o)$ in all the iterations is

$$\begin{aligned}
&O(k(n-|S|)\cdot i_{\max} + 2\mathbb{E}[|\mathbb{C}_1^s\cup\mathbb{C}_2^s|]\cdot\mathbb{E}[|C^s|]) \\
=&O\left(\frac{(k\ln(n-|S|)+\ln 1/\delta)\mathbb{E}[I_G(s)]}{(\epsilon+\epsilon\beta-2\beta)^2}\right).
\end{aligned}$$

In summary, LSBM runs in $O(\frac{(k\ln(n-|S|)+\ln 1/\delta)\mathbb{E}[I_G(s)]m\cdot\alpha(m,n)}{(\epsilon+\epsilon\beta-2\beta)^2 D_s^L(B_L^o)} + \frac{m\cdot\ln 1/\delta}{\beta^2})$ expected time. □