Supplement to “Directed Hybrid Random Networks Mixing Preferential Attachment with Uniform Attachment Mechanisms”

Tiandong Wang∗ Panpan Zhang†

January 18, 2022

This supplement provides the omitted technical details to the proofs of Theorem 1 and Proposition 1, in the main manuscript. The setting, notation, equation reference numbers are retained from the main paper.

1 Proof of Theorem 1

Here we only present the details for the in-degree sequence, and those for out-degrees will follow from a similar reasoning. The derivations for out-degree sequence are similar, so omitted.

For $k, \ell \geq 0$, we have

$$E\left[M_{k+\ell}^{\text{in}} - M_k^{\text{in}}\right] = \sum_{s=k+1}^{k+\ell} E\left[M_s^{\text{in}} - M_{s-1}^{\text{in}}\right]$$

$$= \sum_{s=k+1}^{k+\ell} E\left[\frac{(1-p)(\alpha + \beta)}{|V_{s+1}|} \prod_{\ell=1}^{s} \left(1 + \frac{p(\alpha + \beta)}{(1 + \delta_{\text{in}})} \right)^{(1-p)(\alpha + \beta)}\right]$$

$$\leq \sum_{s=k+1}^{k+\ell} \prod_{s=k+1}^{s} \left(1 + \frac{p(\alpha + \beta)}{(1 + \delta_{\text{in}})(\ell + 1)}\right) E\left[\frac{(1-p)(\alpha + \beta)}{|V_s|}\right],$$

where

$$\prod_{\ell=1}^{s} \left(1 + \frac{p(\alpha + \beta)}{(1 + \delta_{\text{in}})(\ell + 1)}\right) = \frac{\Gamma\left(s + 2 + \frac{p(\alpha + \beta)}{1 + \delta_{\text{in}}}\right)}{\Gamma(s + 2)\Gamma\left(2 + \frac{p(\alpha + \beta)}{1 + \delta_{\text{in}}}\right)} \sim s^{-p(\alpha + \beta)/(1 + \delta_{\text{in}})},$$

for $s$ large. For the expectation in the summand, we have

$$E\left[\frac{(1-p)(\alpha + \beta)}{|V(s)|}\right] = (1-p)(\alpha + \beta) \left(1 + \frac{1}{(1 - \beta)s} + \frac{1}{|V(s)|} - \frac{1}{(1 - \beta)s}\right).$$

Since for $n \geq 1$, $|V_n| - 1$ is a binomial random variable with success probability $1 - \beta$, then we apply the Chernoff bound to get

$$P\left(|V_n| - (1 - \beta)n \geq \sqrt{12(1 - \beta)n \log n}\right) \leq \frac{2}{n^4}. \quad (1)$$

∗Department of Statistics, Texas A&M University, 3143 TAMU, College Station, TX 77843, USA, twang@stat.tamu.edu

†Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania, 423 Guardian Drive, 501, Philadelphia, PA 19104, USA
Then by (1), we get
\[
\mathbb{E} \left[ \frac{1}{V(s)} - \frac{1}{(1-\beta) s} \right] = \mathbb{E} \left[ \frac{|V(s)| - (1-\beta) s}{(1-\beta) s |V(s)|} \right] \leq \frac{\sqrt{12(1-\beta) s \log s}}{(1-\beta) s - \sqrt{12(1-\beta) s \log s}} + \frac{(2-\beta) s}{(1-\beta) s \sqrt{12(1-\beta) s \log s}} \times \frac{2}{s^4}. \tag{2}
\]

Putting them together, we conclude that \( \mathbb{E} |M_{k+t}^{\text{in}} - M_k^{\text{in}}| \to 0 \) as \( k \to \infty \), suggesting that \( \{ \mathbb{E} |M_n^{\text{in}}| \}_{n \geq 1} \) is a cauchy sequence in \( L_1 \) space. Applying the martingale convergence theorem (Durrett, 2006, Theorem 4.2.11) gives that there exists some finite random variable \( L_i \) such that as \( n \to \infty \),
\[
M_n^{\text{in}} \xrightarrow{a.s.} L_i. \tag{3}
\]

Then it remains to show the almost sure convergence of
\[
X_n := \prod_{k=0}^{n-1} \left( 1 + \frac{(\alpha + \beta) p}{S_i + k + 1 + \delta_{\text{in}} |V_{S_i+k}|} \right).
\]

Consider \( \log X_n \), and rewrite it as
\[
\log X_n = \left[ \log \left( 1 + \frac{(\alpha + \beta) p}{S_i + k + 1 + \delta_{\text{in}} |V_{S_i+k}|} \right) - \sum_{k=0}^{n-1} \frac{(\alpha + \beta) p}{S_i + k + 1 + \delta_{\text{in}} |V_{S_i+k}|} \right] + \left[ \sum_{k=0}^{n-1} \frac{(\alpha + \beta) p}{S_i + k + 1 + \delta_{\text{in}} |V_{S_i+k}|} - \sum_{k=1}^{n} \frac{C_1}{S_i + k} - C_1 \log n \right] + C_1 \log n =: P_1(n) + P_2(n) + P_3(n) + C_1 \log n. \tag{4}
\]

For \( P_1(n) \), we first note that \( \log(1+x) - x \leq 0 \) for all \( x \geq 0 \). Then \( P_1(n+1) - P_1(n) \leq 0 \), i.e. \( P_1(n) \) is decreasing in \( n \), and it suffices to show
\[
P_1(\infty) := \sum_{k=0}^{\infty} \left( \log \left( 1 + \frac{(\alpha + \beta) p}{S_i + k + 1 + \delta_{\text{in}} |V_{S_i+k}|} \right) - \frac{(\alpha + \beta) p}{S_i + k + 1 + \delta_{\text{in}} |V_{S_i+k}|} \right)
\]
is finite almost surely. Note also that \( |\log(1+x) - x| \leq x^2/2 \), for all \( x \geq 0 \), then
\[
\mathbb{E} \left[ \sum_{k=0}^{\infty} \left( \log \left( 1 + \frac{(\alpha + \beta) p}{S_i + k + 1 + \delta_{\text{in}} |V_{S_i+k}|} \right) - \frac{(\alpha + \beta) p}{S_i + k + 1 + \delta_{\text{in}} |V_{S_i+k}|} \right) \right] \leq \sum_{k=0}^{\infty} \mathbb{E} \left[ \log \left( 1 + \frac{(\alpha + \beta) p}{S_i + k + 1 + \delta_{\text{in}} |V_{S_i+k}|} \right) - \frac{(\alpha + \beta) p}{S_i + k + 1 + \delta_{\text{in}} |V_{S_i+k}|} \right] \leq (\alpha + \beta)^2 p^2 \sum_{k=0}^{\infty} \mathbb{E} \left( \frac{1}{S_i + k + 1 + \delta_{\text{in}} |V_{S_i+k}|} \right)^2 \leq \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty.
\]

Hence, \( P_1(n) \xrightarrow{a.s.} P_1(\infty) \).
For $P_2(n)$, we apply (Athreya and Ney, 2004, Theorem 3.9.4) to conclude that there exists a finite r.v. $Z$ such that

$$\sum_{k=1}^{\infty} \left( \frac{(\alpha + \beta)p}{k + \delta_{in}|V(k-1)|} - \frac{C_1}{k} \right) \overset{a.s.}{\to} Z,$$

then

$$P_2(n) \overset{a.s.}{\to} Z - \sum_{k=1}^{S_i} \left( \frac{(\alpha + \beta)p}{k + \delta_{in}|V(k-1)|} - \frac{C_1}{k} \right) =: P_2(\infty).$$

Meanwhile, since $\sum_{k=1}^{n} 1/k - \log n \to \tilde{c}$ as $n \to \infty$, where $\tilde{c}$ is Euler’s constant, then for $i \geq 1$,

$$P_3(n) \overset{a.s.}{\to} C_1 \left( \tilde{c} + \log(S_i + 1) - \sum_{k=1}^{S_i} \frac{1}{k} \right) =: P_3(\infty).$$

Then we conclude from (4) that

$$\frac{X_n}{n^{\tilde{c} + 1}} \overset{a.s.}{\to} \exp \left( (P_1(\infty) + P_2(\infty) + P_3(\infty)) \right) < \infty. \quad (5)$$

The results in the theorem follow by combining (5) with (3), where we set

$$\zeta_i := L_i \exp \left( -(P_1(\infty) + P_2(\infty) + P_3(\infty)) \right).$$

□

2 Proof of Proposition 1

We explicitly demonstrate the derivations of Proposition 1 for the in-degree distribution, and the methodology is also applicable to out-degrees. Let $F_n$ denote the $\sigma$-field generated by the network evolution up to $n$ steps. Set $\tau$ to be an $(F_n)_{n\geq0}$-stopping time, then

$$F_\tau = \{ F : F \cap \{ \tau = n \} \in F_n \}.$$ 

For $i \geq 1$, let $S_i$ be the time when node $i$ is created, i.e.,

$$S_i = \inf \{ n \geq 0 : |V_n| = i \}.$$

Then $S_i$ is an $(F_n)_{n\geq0}$-stopping time. Also, for $n \geq k \geq 0$, we have

$$\{ S_i + k = n \} = \{ S_i = n - k \} \in F_{n-k} \subset F_n,$$

so $S_i + k, k \geq 0$, is a stopping time with respect to $(F_n)_{n\geq0}$.

Since the in-degree of $i$ is increased at most by 1 at each evolutionary step, then under $\mathbb{P}^{F_{S_i+n}}(\cdot) := \mathbb{P}(\cdot | F_{S_i+n})$, we have

$$\mathbb{E}^{F_{S_i+n}} \left[ \left( D_i^{in}(S_i + n + 1) + \delta_{in} \right)^2 \right]$$

$$= \left( D_i^{in}(S_i + n) + \delta_{in} \right)^2 + (\alpha + \beta) \left[ \frac{p(D_i^{in}(S_i + n) + \delta_{in})}{\sum_{k \in V_{S_i+n}} (D_k^{in}(S_i + n) + \delta_{in})} \right] \left[ \frac{1 - p}{|V_{S_i+n}|} \right]$$

$$+ 2(\alpha + \beta)(D_i^{in}(S_i + n) + \delta_{in}) \left[ \frac{p(D_i^{in}(S_i + n) + \delta_{in})}{\sum_{k \in V_{S_i+n}} (D_k^{in}(S_i + n) + \delta_{in})} \right] \left[ \frac{1 - p}{|V_{S_i+n}|} \right]$$

$$= \left( D_i^{in}(S_i + n) + \delta_{in} \right)^2 \left[ 1 + \frac{2(\alpha + \beta)p}{S_i + n + 1 + \delta_{in}|V_{S_i+n}|} \right]$$

$$+ (D_i^{in}(S_i + n) + \delta_{in}) \left[ \frac{(\alpha + \beta)p}{S_i + n + 1 + \delta_{in}|V_{S_i+n}|} \right]$$

$$+ (\alpha + \beta) \frac{1 - p}{|V_{S_i+n}|}. \quad (6)$$
With $C_1 = p(\alpha + \beta)/(1 + \delta_{in}(1 - \beta))$, we have
\[
E \left[ (D_{in}^n(S_i + n + 1) + \delta_{in})^2 \right] \\
= E \left[ (D_{in}^n(S_i + n) + \delta_{in})^2 \left( 1 + \frac{2(\alpha + \beta)p}{S_i + n + 1 + \delta_{in}|V_{S_i+n}|} \right) \right] \\
+ E \left[ (D_{in}^n(S_i + n) + \delta_{in}) \left( \frac{(\alpha + \beta)p}{S_i + n + 1 + \delta_{in}|V_{S_i+n}|} + \frac{2(\alpha + \beta)(1 - p)}{|V_{S_i+n}|} \right) \right] \\
+ (\alpha + \beta)E \left( \frac{1 - p}{|V_{S_i+n}|} \right) =: R_i^{(1)}(n) + R_i^{(2)}(n) + R_i^{(3)}(n). (7)
\]

First, we rewrite $R_i^{(1)}(n)$ to obtain
\[
R_i^{(1)}(n) = E \left[ (D_{in}^n(S_i + n) + \delta_{in})^2 \right] \left( 1 + \frac{2C_1}{n} \right) \\
+ E \left[ (D_{in}^n(S_i + n) + \delta_{in})^2 \left( \frac{2(\alpha + \beta)p}{S_i + n + 1 + \delta_{in}|V_{S_i+n}|} - \frac{2C_1}{n} \right) \right] \\
\leq E \left[ (D_{in}^n(S_i + n) + \delta_{in})^2 \right] \left( 1 + \frac{2C_1}{n} \right) \\
+ E \left[ (D_{in}^n(S_i + n) + \delta_{in})^2 \frac{2(\alpha + \beta)p\delta_{in}|V_n| - (1 - \beta)n}{(n + \delta_{in}|V_n|)(1 + \delta_{in}(1 - \beta))n} \right] \\
\leq E \left[ (D_{in}^n(S_i + n) + \delta_{in})^2 \right] \left( 1 + \frac{2C_1}{n} \right) \\
+ E \left[ (D_{in}^n(S_i + n) + \delta_{in})^2 \frac{2\delta_{in}|V_n| - (1 - \beta)n}{n^2} \right].
\]

Note also that $|V_n| - (1 - \beta)n \leq n + 1$, and $(D_{in}^n(S_i + n) + \delta_{in})^2 \leq (n + 1 + \delta_{in})^2$, then
\[
E \left[ (D_{in}^n(S_i + n) + \delta_{in})^2 \frac{2\delta_{in}|V_n| - (1 - \beta)n}{n^2} \right] \\
\leq E \left[ (D_{in}^n(S_i + n) + \delta_{in})^2 \right] \frac{2\delta_{in}\sqrt{T_2n \log n}}{n^2} + \frac{2\delta_{in}(n + 1 + \delta_{in})^2(n + 1)}{n^2} \times \frac{2}{n^4}.
\]

Therefore, we conclude that
\[
R_i^{(1)}(n) \leq E \left[ (D_{in}^n(S_i + n) + \delta_{in})^2 \right] \left( 1 + \frac{2C_1}{n} + \frac{2\delta_{in}\sqrt{T_2n \log n}}{n^2}\frac{2\delta_{in}(n + 1 + \delta_{in})^2(n + 1)}{n^2} \times \frac{2}{n^4} \right). (8)
\]

For $R_i^{(3)}(n)$, we apply the Cauchy-Schwartz inequality to obtain
\[
R_i^{(3)}(n) \leq \frac{1 - p}{n(1 - \beta)} \left( E \left[ ((1 - \beta)(S_i + n + 1) - |V_{S_i+n}|)^2 \right] \right)^{1/2} \left( E \left[ |V_n|^2 \right] \right)^{1/2}.
\]

By Theorem 3.9.4 in Athreya and Ney (2004), we have as $n \to \infty$,
\[
n^2E \left[ |V_n|^2 \right] \to (1 - \beta)^{-2}.
\]

Meanwhile,
\[
E \left[ ((1 - \beta)(S_i + n + 1) - |V_{S_i+n}|)^2 \right]
\]
\[ \text{Var}(|V_{S_i+n}|) + \text{Var}((1-\beta)(S_i + n + 1)) = \beta(1-\beta)\left(\mathbb{E}(S_i) + n\right) + 2(1-\beta)^2\text{Var}(S_i). \]

Hence, there exists some constant \( A_i > 0 \) such that
\[ R_i^{(3)}(n) \leq A_i n^{-3/2}. \tag{9} \]

For \( R_i^{(2)}(n) \), we now claim that
\[ \sup_{i \geq 1} \frac{\mathbb{E}(D_{i\text{in}}(n) + \delta_{i\text{in}})}{n^{C_1}} < \infty, \tag{10} \]
then for some constant \( K > 0 \), \( R_i^{(2)}(n) \) is bounded by
\[ Kn^{C_1} \mathbb{E}\left(\frac{(\alpha + \beta)p}{S_i + n + 1 + \delta_{i\text{in}}|V_{S_i+n}|} + \frac{2(\alpha + \beta)(1-p)}{|V_{S_i+n}|}\right), \]
and by (1) and (9) we see that
\[ \leq K n^{C_1} \left( \frac{C_1}{n} + \frac{2\delta_{i\text{in}}\sqrt{T2n \log n}}{n^2} + \frac{4\delta_{i\text{in}}(n+1+\delta_{i\text{in}})^2(n+1)}{n^6} + 2A_i n^{-3/2} \right). \tag{11} \]

Then combining (8), (9) and (11) gives that there exists some constant \( A_0 > 0 \) such that
\[ \mathbb{E}\left[(D_{i\text{in}}(S_i + n + 1) + \delta_{i\text{in}})^2\right] \leq \mathbb{E}\left[(D_{i\text{in}}(S_i + n) + \delta_{i\text{in}})^2\right] + 2\left(\alpha^2 + \beta^2\right)(S_i + n + 1 + \delta_{i\text{in}}|V_{S_i+n}|) + 2\delta_{i\text{in}}|V_{S_i+n}| + 2(1-p)\left|V_{S_i+n}\right|. \]

Note that
\[ \prod_{l=k+1}^{n} \left(1 + \frac{2C_1}{l} + \frac{2\delta_{i\text{in}}\sqrt{T2l \log l}}{l^2}\right) \leq \exp\left\{ \sum_{l=k+1}^{n} \left(\frac{2C_1}{l} + \frac{2\delta_{i\text{in}}\sqrt{T2l \log l}}{l^2}\right)\right\} \]
\[ \leq (n/k)^{2C_1} \exp\left\{ \sum_{l=k+1}^{\infty} \frac{2\delta_{i\text{in}}\sqrt{T2l \log l}}{l^2}\right\}. \tag{12} \]

Since \( \sum_{l=k+1}^{\infty} \frac{2\delta_{i\text{in}}\sqrt{T2l \log l}}{l^2} < \infty \), then we conclude that
\[ \sup_{i \geq 1} \frac{\mathbb{E}\left[(D_{i\text{in}}(n))^2\right]}{n^{2C_1}} < \infty. \]

Hence, we are left with verifying (10). Note that
\[ \mathbb{E}\left(D_{i\text{in}}(S_i + n + 1) + \delta_{i\text{in}}\right) = \mathbb{E}\left[(D_{i\text{in}}(S_i + n) + \delta_{i\text{in}})\left(1 + \frac{p(\alpha + \beta)}{S_i + n + 1 + \delta_{i\text{in}}|V_{S_i+n}|}\right)\right] + \mathbb{E}\left[1 - p\frac{\alpha + \beta}{|V_{S_i+n}|}\right]. \]
\[
\begin{align*}
&\leq \mathbb{E}\left[\left(D_i^{in}(S_i + n) + \delta_{in}\right)\left(1 + \frac{C_1}{n}\right)\right] \\
&\quad + \mathbb{E}\left[\left(D_i^{in}(S_i + n) + \delta_{in}\right)\left\{\frac{p(\alpha + \beta)}{n + 1 + \delta_{in}|V_n|} - \frac{C_1}{n}\right\}\right] + \mathbb{E}\left[1 - \frac{p}{|V_{S_i+n}|}\right],
\end{align*}
\]
and applying (1) and (9) gives
\[
\begin{align*}
&\leq \mathbb{E}\left(D_i^{in}(S_i + n) + \delta_{in}\right)\left(1 + \frac{C_1}{n} + \frac{\delta_{in}\sqrt{12n \log n}}{n^2}\right) \\
&\quad + \frac{2\delta_{in}(n + 1 + \delta_{in})(n + 1)}{n^6} + A_in^{-3/2}. 
\end{align*}
\]
Iterating the inequality in (13) backwards for \(n\) times gives
\[
\begin{align*}
\mathbb{E}\left[D_i^{in}(S_i + n + 1) + \delta_{in}\right] \\
&\leq \left(\delta_{in}\alpha + (1 + \delta_{in})\gamma\right)\prod_{r=1}^{n}\left(1 + \frac{C_1}{r} + \frac{1 + \sqrt{12r \log r}}{r^2}\right) \\
&\quad + \sum_{r=1}^{n}\left(\frac{2\delta_{in}(r + 1 + \delta_{in})(r + 1)}{r^6} + A_ir^{-3/2}\right)\prod_{s=r+1}^{n}\left(1 + \frac{C_1}{s} + \frac{1 + \sqrt{12s \log s}}{s^2}\right). 
\end{align*}
\]
By a similar argument as in (12), we see that the bound in (14) implies (10), which completes the proof of Proposition 1. \(\square\)

References

Athreya, K. and P. Ney (2004). Branching processes. Reprint of the 1972 original. Springer.

Durrett, R. T. (2006). Random Graph Dynamics. Cambridge, U.K.: Cambridge University Press.