Submitted to the Annals of Statistics
arXiv: arXiv:1312.2645

SUPPLEMENT TO SUBSAMPLING BOOTSTRAP OF COUNT FEATURES OF NETWORKS

BY SHARMODEEP BHATTACHARYYA[*,†] AND PETER J. BICKEL[*]

University of California, Berkeley[*]
Oregon State University[†]

Appendix.

A1. *Variance of* $\hat{P}_{b1}(R)$. The variance of $\hat{P}_{b1}(R)$ is

$$\mathrm{Var}_b\left[\frac{1}{\binom{m}{p}|Iso(R)|}\sum_{S\subseteq K_m, S\cong R}\mathbf{1}(S\subseteq H)\Big|G\right] = \left(\frac{1}{\binom{m}{p}|Iso(R)|}\right)^2\mathrm{Var}_b\left[\sum_{S\subseteq K_m, S\cong R}\mathbf{1}(S\subseteq H)\Big|G\right]$$

$$\begin{aligned}\mathrm{Var}_b\left[\sum_{S\subseteq K_m, S\cong R}\mathbf{1}(S\subseteq H)\Big|G\right] &= \mathbb{E}_b\left[\left(\sum_{S\subseteq K_m, S\cong R}\mathbf{1}(S\subseteq H)\right)^2\Big|G\right] \\ &\quad -\left(\mathbb{E}_b\left[\sum_{S\subseteq K_m, S\cong R}\mathbf{1}(S\subseteq H)\Big|G\right]\right)^2\end{aligned}$$

$$\begin{aligned}\mathbb{E}_b\left[\left(\sum_{S\subseteq K_m, S\cong R}\mathbf{1}(S\subseteq H)\right)^2\Big|G\right] &= \mathbb{E}_b\left[\sum_{S\subseteq K_m, S\cong R}\mathbf{1}(S\subseteq H)\Big|G\right] \\ &\quad + \mathbb{E}_b\left[\sum_{\substack{S,T\subseteq K_m\\ S,T\cong R, S\neq T}}\mathbf{1}(S,T\subseteq H)\Big|G\right] \\ &= I + II \text{ (Suppose)}\end{aligned}$$

Thus,

$$I = \sum_{S\subseteq K_n, S\cong R}\frac{\binom{n-p}{m-p}}{\binom{n}{m}}\mathbf{1}(S\subseteq G)$$

$$II = \mathbb{E}_b \left[\sum_{\substack{S,T \subseteq K_m \\ S,T \cong R, S \neq T}} \mathbf{1}(S, T \subseteq H) \middle| G \right]$$

Now, a host of subgraphs can be formed by the intersection of two copies of R. The number of intersected vertices can range from 0 to $p-1$. Let us consider, that for number of vertices in intersection as k ($k = 1, \ldots, (p-1)$), the number of graph structures that can be formed is g_k and we represent that graph structure by W_{jk}, where, $j = 1, \ldots, g_k$. Thus,

$$II = \sum_{k=0}^{p-1} \sum_{j=1}^{g_k} \sum_{S \subseteq K_n, S \cong W_{jk}} \frac{\binom{n-(2p-k)}{m-(2p-k)}}{\binom{n}{m}} \mathbf{1}(S \subseteq G)$$

So,

$$\begin{aligned} \mathbb{E}_b \left[\left(\sum_{S \subseteq K_m, S \cong R} \mathbf{1}(S \subseteq H) \right)^2 \middle| G \right] &= \sum_{S \subseteq K_n, S \cong R} \frac{\binom{n-p}{m-p}}{\binom{n}{m}} \mathbf{1}(S \subseteq G) \\ &+ \sum_{k=0}^{p-1} \sum_{j=1}^{g_k} \\ &\quad \sum_{S \subseteq K_n, S \cong W_{jk}} \frac{\binom{n-(2p-k)}{m-(2p-k)}}{\binom{n}{m}} \mathbf{1}(S \subseteq G) \end{aligned}$$

$$\begin{aligned} \mathrm{Var}_b \left[\sum_{S \subseteq K_m, S \cong R} \mathbf{1}(S \subseteq H) \middle| G \right] &= \sum_{S \subseteq K_n, S \cong R} \frac{\binom{n-p}{m-p}}{\binom{n}{m}} \mathbf{1}(S \subseteq G) \\ &+ \sum_{k=0}^{p-1} \sum_{j=1}^{g_k} \\ &\quad \sum_{S \subseteq K_n, S \cong W_{jk}} \frac{\binom{n-(2p-k)}{m-(2p-k)}}{\binom{n}{m}} \mathbf{1}(S \subseteq G) \\ &- \left(\sum_{S \subseteq K_n, S \cong R} \frac{\binom{n-p}{m-p}}{\binom{n}{m}} \mathbf{1}(S \subseteq G) \right)^2 \end{aligned}$$

$$\begin{aligned}
&\mathrm{Var}_b\left[\frac{1}{\binom{m}{p}|Iso(R)|}\sum_{S\subseteq K_m,S\cong R}\mathbf{1}(S\subseteq H)\middle|G\right]\\
&\quad=\left(\frac{1}{\binom{m}{p}|Iso(R)|}\right)^2\left[\sum_{S\subseteq K_n,S\cong R}\frac{\binom{n-p}{m-p}}{\binom{n}{m}}\mathbf{1}(S\subseteq G)\right]\\
&\quad-\left(\frac{1}{\binom{m}{p}|Iso(R)|}\right)^2\left[\left(\sum_{S\subseteq K_n,S\cong R}\frac{\binom{n-p}{m-p}}{\binom{n}{m}}\mathbf{1}(S\subseteq G)\right)^2\right]\\
&\quad+\left(\frac{1}{\binom{m}{p}|Iso(R)|}\right)^2\left[\sum_{k=0}^{p-1}\sum_{j=1}^{g_k}\sum_{\substack{S\subseteq K_n\\ S\cong W_{jk}}}\frac{\binom{n-(2p-k)}{m-(2p-k)}}{\binom{n}{m}}\mathbf{1}(S\subseteq G)\right]
\end{aligned}$$

So,

$$\begin{aligned}
\mathrm{Var}_b\left[\hat{P}_{b1}(R)\right] &= \left(\frac{1}{\binom{m}{p}|Iso(R)|}\right)^2\left[\sum_{S\subseteq K_n,S\cong R}\frac{\binom{n-p}{m-p}}{\binom{n}{m}}\mathbf{1}(S\subseteq G)\right]\\
&- \left(\frac{1}{\binom{m}{p}|Iso(R)|}\right)^2\left[\left(\sum_{S\subseteq K_n,S\cong R}\frac{\binom{n-p}{m-p}}{\binom{n}{m}}\mathbf{1}(S\subseteq G)\right)^2\right]\\
&+ \left(\frac{1}{\binom{m}{p}|Iso(R)|}\right)^2\left[\sum_{k=0}^{p-1}\sum_{j=1}^{g_k}\sum_{\substack{S\subseteq K_n\\ S\cong W_{jk}}}\frac{\binom{n-(2p-k)}{m-(2p-k)}}{\binom{n}{m}}\mathbf{1}(S\subseteq G)\right]
\end{aligned}$$

A2. *Proof of Theorem 3.1.*

PROOF. (i) Now, let us try to try to find the expectation of $\hat{P}_{b1}(R)$

under the sampling distribution conditional on the given data G.

$$\begin{aligned}
&\mathbb{E}_b\left[\frac{1}{\binom{m}{p}|Iso(R)|}\sum_{S\subseteq K_m, S\cong R}\mathbf{1}(S\subseteq H)\Big|G\right]\\
&= \frac{1}{\binom{m}{p}|Iso(R)|}\mathbb{E}\left[\sum_{S\subseteq K_m, S\cong R}\mathbf{1}(S\subseteq H)\Big|G\right]\\
&= \frac{1}{\binom{m}{p}|Iso(R)|}\sum_{H\subseteq G, |H|=m}\frac{1}{\binom{n}{m}}\sum_{S\subseteq K_m, S\cong R}\mathbf{1}(S\subseteq H)\\
&= \frac{1}{\binom{m}{p}|Iso(R)|}\sum_{\substack{S\subseteq K_n\\ S\cong R}}\sum_{\substack{H\supseteq S, H\subseteq G\\ |H|=m}}\frac{1}{\binom{n}{m}}\mathbf{1}(S\subseteq G)\\
&= \frac{1}{\binom{m}{p}|Iso(R)|}\sum_{S\subseteq K_n, S\cong R}\frac{\binom{n-p}{m-p}}{\binom{n}{m}}\mathbf{1}(S\subseteq G)\\
&= \frac{1}{\binom{n}{p}|Iso(R)|}\sum_{S\subseteq K_n, S\cong R}\mathbf{1}(S\subseteq G)
\end{aligned}$$

So, we have,

$$\mathbb{E}_b[\bar{P}_{B1}(R)|G] = \mathbb{E}_b[\hat{P}_{b1}(R)|G] = \hat{P}(R)$$

(ii) Given G,

$$\begin{aligned}
\mathrm{Var}_b[\rho^{-e}\bar{P}_{B1}(R)|G] = \rho^{-2e}\frac{1}{B^2}\Bigg(&\sum_{b=1}^{B}\mathrm{Var}_b[\hat{P}_{b1}(R)]\\
&+\sum_{b,b'=1, b\neq b'}^{B}\mathrm{Cov}_b(\hat{P}_{b1}(R), \hat{P}_{b'1}(R))\Bigg)
\end{aligned}$$

Now, the formula for $\mathrm{Var}_b[\hat{P}_{b1}(R)]$ from A1 we get that

$$\begin{aligned}
\mathrm{Var}_b\left[\rho_n^{-e}\hat{P}_{b1}(R)\right] &= \left(\frac{\rho_n^{-e}}{\binom{m}{p}|Iso(R)|}\right)^2\left[\sum_{S\subseteq K_n, S\cong R}\frac{\binom{n-p}{m-p}}{\binom{n}{m}}\mathbf{1}(S\subseteq G)\right]\\
&- \left(\frac{\rho_n^{-e}}{\binom{m}{p}|Iso(R)|}\right)^2\left[\left(\sum_{S\subseteq K_n, S\cong R}\frac{\binom{n-p}{m-p}}{\binom{n}{m}}\mathbf{1}(S\subseteq G)\right)^2\right]\\
&+ \left(\frac{\rho_n^{-e}}{\binom{m}{p}|Iso(R)|}\right)^2\left[\sum_{k=0}^{p-1}\sum_{j=1}^{g_k}\sum_{\substack{S\subseteq K_n\\ S\cong W_{jk}}}\frac{\binom{n-(2p-k)}{m-(2p-k)}}{\binom{n}{m}}\mathbf{1}(S\subseteq G)\right]
\end{aligned}$$

So,

$$\begin{aligned}\mathrm{Var}_b\left[\rho_n^{-e}\hat{P}_{b1}(R)\right] &= O\left(\frac{1}{\rho_n^e\binom{m}{p}}\right) + O\left(\frac{\binom{m}{p_W}\rho_n^{e_W}}{\binom{m}{p}^2\rho_n^{2e}}\right)\\ &= O\left(\frac{1}{m^p\rho_n^e}\right) + O\left(\frac{1}{m}\right)\end{aligned}$$

where, $W = S \cup S'$, $p_W = |V(W)|$ and $e_W = |E(W)|$.

(iii) Here, we use properties of the underlying model. Let us condition on $\boldsymbol{\xi} = \{\xi_1, \ldots, \xi_n\}$ and the whole graph G separately. Now, conditioning on $\boldsymbol{\xi}$, we get the main term of $\hat{P}(R)$ to be,

(0.1)

$$\mathbb{E}(\rho^{-e}\hat{P}(R)|\boldsymbol{\xi}) = \frac{1}{\binom{n}{p}|Iso(R)|}\sum_{S\subseteq K_n, S\cong R}\left(\prod_{(i,j)\in E(S)} w(\xi_i,\xi_j)\right) + O(n^{-1}\lambda_n).$$

We shall use the same decomposition as used in [1] of $(\rho_n^{-e}\bar{P}_{B1}(R) - \tilde{P}(R))$ into

$$\begin{aligned}(\rho_n^{-e}\bar{P}_{B1}(R) - \tilde{P}(R)) &= \rho_n^{-e}\left(\bar{P}_{B1} - \mathbb{E}_b[\hat{P}_{b1}(R)|G]\right)\\ &\quad + \rho_n^{-e}\left(\hat{P}(R) - \mathbb{E}(\hat{P}(R)|\boldsymbol{\xi})\right)\\ &\quad + \mathbb{E}(\hat{P}(R)|\boldsymbol{\xi})\rho_n^{-e} - \tilde{P}(R)\end{aligned}$$

Let us define,

$$\begin{aligned}U_3 &= \mathbb{E}(\hat{P}(R)|\boldsymbol{\xi})\rho_n^{-e} - \tilde{P}(R)\\ U_2 &= \rho_n^{-e}\left(\hat{P}(R) - \mathbb{E}(\hat{P}(R)|\boldsymbol{\xi})\right)\\ U_1 &= \rho_n^{-e}\left(\bar{P}_{B1} - \mathbb{E}_b[\hat{P}_{b1}(R)|G]\right)\end{aligned}$$

Now, it is easy to see that

$$\begin{aligned}\mathrm{Var}(\rho^{-e}\bar{P}_{B1}(R)) &= \mathbb{E}(\mathrm{Var}(\rho^{-e}\bar{P}_{B1}(R)|G)) + \mathrm{Var}(\mathbb{E}(\rho^{-e}\bar{P}_{B1}(R)|G))\\ &= \mathbb{E}(\mathrm{Var}(\rho^{-e}\bar{P}_{B1}(R) - \hat{P}(R)|G) + \mathrm{Var}(\hat{P}(R))\\ &= \mathbb{E}(\mathrm{Var}(U_1|G)) + \mathbb{E}(\mathrm{Var}(\hat{P}(R)|\boldsymbol{\xi})) + \mathrm{Var}(\mathbb{E}(\hat{P}(R)|\boldsymbol{\xi}))\\ &= \mathbb{E}(\mathrm{Var}(U_1|G)) + \mathbb{E}(\mathrm{Var}(U_2|\boldsymbol{\xi})) + \mathrm{Var}(U_3)\end{aligned}$$

We shall try to see the behavior of $\mathrm{Var}(U_1|G) = \mathrm{Var}_b[\rho^{-e}\bar{P}_{B1}(R)|G]$. From (ii) we get that, $\mathrm{Var}_b\left[\rho_n^{-e}\hat{P}_{b1}(R)\right] = O\left(\frac{1}{m^p\rho_n^e} \vee \frac{1}{m}\right)$. Similarly, $\mathrm{Cov}_b[\rho_n^{-e}\hat{P}_{b1}(R), \rho_n^{-e}\hat{P}_{b'1}(R)] = O(\frac{1}{m})$ for acyclic and k-cycle R following similar steps as variance in Appendix A1. If we consider the uniform probability for bootstrap to be γ, then, $B = O(\gamma n^p)$. Note that, if $E(H_b) \cap E(H_{b'}) = \phi$, then, $\mathrm{Cov}_b(\hat{P}_{b1}(R), \hat{P}_{b'1}(R)) = 0$. The number of pairs such that $E(H_b) \cap E(H_{b'}) \neq \phi$ is $O(m^2\gamma^2 n^{2m-2})$. Also, the number of edges for the leading term in the covariance is equal to or more than $2e$. So,

$$\mathbb{E}(\mathrm{Var}_b[\rho^{-e}\bar{P}_{B1}(R)|G]) = O\left(\frac{1}{B(m^p\rho_n^e \wedge m)}\right) + O\left(\frac{m^2\gamma^2 n^{2m-2}}{m\gamma^2 n^{2m}}\right)$$
$$= O\left(\frac{1}{B(m^p\rho_n^e \wedge m)} + \frac{m}{n^2}\right) = O\left(\frac{1}{B(m^p\rho_n^e \wedge m)}\right) + o(n^{-1})$$

The second equality follows since we have $m/n \to 0$ as $n \to \infty$. So, since, $B(m^p\rho_n^e \wedge m) > O(n)$, we have, $\mathbb{E}(\mathrm{Var}(U_1|G)) = o(n^{-1})$.
Now, by proof of Theorem 1 in [1], we have,

$$\begin{aligned} \mathrm{Var}(U_2) &= o(n^{-1}) \\ \mathrm{Var}(U_3) &= o(n^{-1}) \end{aligned}$$

So, we get, $\mathrm{Var}(\rho^{-e}\bar{P}_{B1}(R)) = o(n^{-1})$. Since, we already know $\sqrt{n}$-consistency of $\left(\rho_n^{-e}\hat{P}(R) - \tilde{P}(R))\right)$, this proves the $\sqrt{n}$-consistency of $\rho_n^{-e}\bar{P}_{B1}(R)$ to $\rho_n^{-e}\hat{P}(R)$.

□

A3. *Proof of Theorem 3.2.* For variance calculation, we also need the joint inclusion probability of two items, $S, S' \in \mathcal{S}_p$, which are subgraphs of G induced by the set of vertices $\{w_1, \ldots, w_p\}$ and $\{w'_1, \ldots, w'_p\}$ respectively, where, we take that $w_{i+1} \succ w_i$ and $w'_{i+1} \succ w'_i$, $i = 1, \ldots, p-1$. So,

$$\pi_{SS'} \equiv \text{Inclusion Probability of } S \textit{ and } S' = \mathbb{P}[(w_1, \ldots, w_p) \text{ is selected and } (w'_1, \ldots, w'_p) \text{ is selected}]$$
$$= \prod_{d=1}^{p} (q_d)^{z_{1d}} \prod_{d=1}^{p} \left(q_d^2\right)^{z_{2d}}$$

where,

$$z_{1d} = \begin{cases} \mathbf{1}(w_d = w'_d), & \text{for } d = 1 \\ \mathbf{1}((w_d, w_{d-1}) = (w'_d, w'_{d-1})), & \text{for } d = 2, \ldots, p \end{cases}$$

$$z_{2d} = \begin{cases} \mathbf{1}(w_d \neq w'_d), & \text{for } d = 1 \\ \mathbf{1}((w_d, w_{d-1}) \neq (w'_d, w'_{d-1})), & \text{for } d = 2, \ldots, p \end{cases}$$

(i) We know that $\hat{P}_{b2}(R)$ is a Horvitz-Thompson estimator with inclusion probability of each population unit to be $\pi = \prod_{d=1}^{p} q_d$. So, according to the sampling theory [2], $\hat{P}_{b2}(R)$ is an unbiased estimator of $\hat{P}(R)$ given the network G, if $\mathbb{P}(\hat{P}_{b2}(R) = 0|\hat{P}(R)) \to 0$ as $n \to \infty$. Now, $\mathbb{P}(\hat{P}_{b2}(R) = 0|\hat{P}(R)) \leq (1 - q_d)^{\lambda_n}$ for all $d = 1, \ldots, p$. For all $d = 1, \ldots, p$, $(1 - q_d)^{\lambda_n} \to 0$ if $\lambda_n q_d \to \infty$. So, under the condition, $\lambda_n q_d \infty$ and $q_d \to 0$ as $n \to \infty$, we have, $\hat{P}_{b2}(R)$ is an asymptotically unbiased estimator of $\hat{P}(R)$.

(ii) The variance of $\hat{P}_{b2}(R)$ coming from the bootstrap sampling only is given by

$$\mathrm{Var}_b\left[\rho_n^{-e}\hat{P}_{b2}(R)\right] = \frac{1}{N^2}\left[\frac{1-\pi}{\pi}\sum_{S\in\mathcal{S}_p}\mathbf{1}(S \cong R) + \sum_{S,S'\in\mathcal{S}_p, S\neq S'}\frac{\pi_{SS'}-\pi^2}{\pi^2}\mathbf{1}(S \cong R,\ S' \cong R)\right]$$

where,

$$N = \rho_n^e\binom{n}{p}|Iso(R)|$$

From the formula of $\mathrm{Var}_b[\hat{P}_{b2}(R)|G]$, we see that, the covariance terms vanishes when $\pi_{SS'} = \pi^2$. Now, if $q_1 = 1$, then, $\pi_{SS'} = \pi^2$ if $E(S) \cap E(S') = \phi$. The number of pairs such that $E(S) \cap E(S') \neq \phi$ is $O(p^2 n^{2p-2})$.

Now, the condition of $q_1 = 1$ is a bit restrictive. In stead, if we have $q_1 \to 1$ as $n \to \infty$, then, the highest order term of covariance term comes from the case when $E(S) \cap E(S') \neq \phi$ but the root nodes are same that is $w_1 = w_1'$. So, for some constant $C > 0$,

$$\begin{aligned}
&\frac{1}{N^2}\sum_{S,S'\in\mathcal{S}_p, S\neq S'}\frac{\pi_{SS'}-\pi^2}{\pi^2}\mathbf{1}(S \cong R,\ S' \cong R)\\
&\leq \frac{C}{N^2}\sum_{S,S'\in\mathcal{S}_p, S\neq S'}\frac{q_1-q_1^2}{q_1^2}\mathbf{1}(S \cong R,\ S' \cong R)\\
&= O\left(\left(\frac{1}{q_1}-1\right)\frac{n^{2p-1}}{n^{2p}}\right)\\
&= O\left(\left(\frac{1}{q_1}-1\right)\frac{1}{n}\right)
\end{aligned}$$

Now, for the variance term to vanish we need the conditions $q_1 = 1$ or $q_1 \to 1$ and $q_d \to 0$ and $\lambda_n q_d \to \infty$ for $d = 2, \ldots, p$ as $n \to \infty$. Since, we know that $O(1) \leq \lambda_n O(n)$, we get $nq_d \to \infty$ for $d = 2, \ldots, p$ as $n \to \infty$. So, we have

$$\begin{aligned}\frac{1}{N^2}\frac{1-\pi}{\pi}\sum_{S\in\mathcal{S}_p}\mathbf{1}(S\cong R) &= \left(\frac{1}{\pi}-1\right)O\left(\frac{1}{n^p\rho^e}\right)\\ &= O\left(\frac{1}{n^p\rho_n^e\pi}\right)\\ &= O\left(\frac{1}{n\rho_n^{e-p+1}}\cdot\prod_{d=2}^{p}\frac{1}{\lambda_n q_d}\right)\end{aligned}$$

So,

$$\mathbb{E}\left(\mathrm{Var}_b\left[\rho_n^{-e}\hat{P}_{b2}(R)\right]\right) = O\left(\left(\frac{1}{q_1}-1\right)\frac{1}{n}\right) + O\left(\frac{1}{n\rho_n^{e-p+1}}\cdot\prod_{d=2}^{p}\frac{1}{\lambda_n q_d}\right)$$

(iii) We shall use the same decomposition as used in [1] of $(\rho_n^{-e}\bar{P}_{B2}(R) - \tilde{P}(R))$ into

$$\begin{aligned}(\rho_n^{-e}\bar{P}_{B2}(R) - \tilde{P}(R)) &= \rho_n^{-e}\left(\bar{P}_{B2} - \mathbb{E}_b[\hat{P}_{b2}(R)|G]\right)\\ &\quad +\rho_n^{-e}\left(\hat{P}(R) - \mathbb{E}(\hat{P}(R)|\boldsymbol{\xi})\right)\\ &\quad +\mathbb{E}(\hat{P}(R)|\boldsymbol{\xi})\rho_n^{-e} - \tilde{P}(R)\end{aligned}$$

Let us define,

$$\begin{aligned}U_3 &= \mathbb{E}(\hat{P}(R)|\boldsymbol{\xi})\rho_n^{-e} - \tilde{P}(R)\\ U_2 &= \rho_n^{-e}\left(\hat{P}(R) - \mathbb{E}(\hat{P}(R)|\boldsymbol{\xi})\right)\\ U_1 &= \rho_n^{-e}\left(\bar{P}_{B2} - \mathbb{E}_b[\hat{P}(R)|G]\right)\end{aligned}$$

Now, it is easy to see that

$$\begin{aligned}\mathrm{Var}(\rho^{-e}\bar{P}_{B2}(R)) &= \mathbb{E}(\mathrm{Var}(\rho^{-e}\bar{P}_{B2}(R)|G)) + \mathrm{Var}(\mathbb{E}(\rho^{-e}\bar{P}_{B2}(R)|G))\\ &= \mathbb{E}(\mathrm{Var}(\rho^{-e}\bar{P}_{B2}(R) - \hat{P}(R)|G) + \mathrm{Var}(\hat{P}(R))\\ &= \mathbb{E}(\mathrm{Var}(U_1|G)) + \mathbb{E}(\mathrm{Var}(\hat{P}(R)|\boldsymbol{\xi})) + \mathrm{Var}(\mathbb{E}(\hat{P}(R)|\boldsymbol{\xi}))\\ &= \mathbb{E}(\mathrm{Var}(U_1|G)) + \mathbb{E}(\mathrm{Var}(U_2|\boldsymbol{\xi})) + \mathrm{Var}(U_3)\end{aligned}$$

We shall try to see the behavior of $\mathrm{Var}_b[\rho_n^{-e}\hat{P}_{b2}(R)|G]$.

$$\mathbb{E}\left(\mathrm{Var}_b\left[\rho_n^{-e}\hat{P}_{b2}(R)\right]\right) = O\left(\left(\frac{1}{q_1}-1\right)\frac{1}{n}\right) + O\left(\frac{1}{n\rho_n^{e-p+1}}\cdot\prod_{d=2}^{p}\frac{1}{\lambda_n q_d}\right)$$

Now, since the bootstrap samples for *subgraph sampling* are selected independently, we have that,

$$\mathbb{E}\left(\mathrm{Var}_b\left[\rho_n^{-e}\bar{P}_{B2}(R)\right]\right) = O\left(\left(\frac{1}{q_1}-1\right)\frac{1}{nB}\right) + O\left(\frac{1}{Bn\rho_n^{e-p+1}}\cdot\prod_{d=2}^{p}\frac{1}{\lambda_n q_d}\right)$$

Now, under the condition $\frac{1}{B}\left(\frac{1}{q_1}-1\right)\to 0$, $q_d \to 0$ for all $d = 1,\ldots,p$ and $B\prod_{d=2}^{p} q_d \geq \frac{1}{n^{p-1}\rho_n^e}$, we have

$$\mathbb{E}(\mathrm{Var}(U_1|G)) = \mathbb{E}\left(\mathrm{Var}_b[\rho_n^{-e}\hat{P}_{b2}(R)|G]\right) = o(n^{-1})$$

Now, by proof of Theorem 1 in [1], we have,

$$\begin{aligned}\mathrm{Var}(U_2) &= o(n^{-1})\\ \mathrm{Var}(U_3) &= o(n^{-1})\end{aligned}$$

So, we get, $\mathrm{Var}(\rho^{-e}\bar{P}_{B2}(R)) = o(n^{-1})$. Since, we already know $\sqrt{n}$-consistency of $\left(\rho_n^{-e}\hat{P}(R)-\tilde{P}(R))\right)$, this proves the $\sqrt{n}$-consistency of $\rho_n^{-e}\bar{P}_{B2}(R)$ to $\rho_n^{-e}\hat{P}(R)$.

B1. *Proof of Proposition 6.*

$$\begin{aligned}\sigma^2(R;\rho) &= \mathrm{Var}\left[\rho^{-e}\hat{P}(R)\right]\\
&= \mathrm{Var}\left[\sum_{S\subseteq K_n, S\cong R}\frac{\mathbf{1}(S\subseteq G)}{\rho^e\binom{n}{p}|Iso(R)|}\right]\\
&= \frac{1}{\left(\rho^e\binom{n}{p}|Iso(R)|\right)^2}\mathbb{E}\left[\sum_{S\subseteq K_n, S\cong R}\mathbf{1}(S\subseteq G)\right]^2 - \left(\tilde{P}(R)\right)^2\\
&= \frac{1}{\left(\rho^e\binom{n}{p}|Iso(R)|\right)^2}\mathbb{E}\left[\sum_{\substack{S,T\subseteq K_n\\ S,T\cong R, S\cap T\neq\emptyset}}\mathbf{1}(S,T\subseteq G)\right] - \left(1-\frac{\binom{n-p}{p}}{\binom{n}{p}}\right)\left(\tilde{P}(R)\right)^2\end{aligned}$$

If R is a connected subgraph, then, we can write,

$$\begin{aligned}&\frac{1}{\left(\rho^e\binom{n}{p}|Iso(R)|\right)^2}\mathbb{E}\left[\sum_{\substack{S,T\subseteq K_n\\ S,T\cong R, S\cap T\neq\emptyset}}\mathbf{1}(S,T\subseteq G)\right]\\
&= \frac{1}{\left(\rho^e\binom{n}{p}|Iso(R)|\right)^2}\sum_{\substack{W:W=S\cup T\\ S,T\cong R, S\cap T\neq\emptyset}}\mathbb{E}\left[\sum_{W\subseteq K_n}\mathbf{1}(W\subseteq G)\right]\end{aligned}$$

$$
\begin{aligned}
\sigma(R_1, R_2; \rho) &= \mathrm{Cov}\left(\rho^{-e_1}\hat{P}(R_1), \rho^{-e_2}\hat{P}(R_2)\right)\\
&= \frac{1}{\left(\rho^{e_1+e_2}\binom{n}{p_1}\binom{n}{p_2}|Iso(R_1)||Iso(R_2)|\right)} \times \\
&\quad \mathbb{E}\left[\left(\sum_{\substack{S\subseteq K_n,\\ S\cong R_1}} \mathbf{1}(S\subseteq G)\right)\left(\sum_{\substack{S\subseteq K_n,\\ S\cong R_2}} \mathbf{1}(S\subseteq G)\right)\right] - \tilde{P}(R_1)\tilde{P}(R_2)\\
&= \frac{1}{\left(\rho^{e_1+e_2}\binom{n}{p_1}\binom{n}{p_2}|Iso(R_1)||Iso(R_2)|\right)}\mathbb{E}\left[\sum_{\substack{S,T\subseteq K_n\\ S\cong R_1, T\cong R_2, S\cap T\neq\emptyset}} \mathbf{1}(S,T\subseteq G)\right]\\
&\quad - \left(1 - \frac{\binom{n-p_1}{p_2}}{\binom{n}{p_2}}\right)\tilde{P}(R_1)\tilde{P}(R_2)
\end{aligned}
$$

If R is a connected subgraph, then, we can write,

$$
\begin{aligned}
&\frac{1}{\left(\rho^{e_1+e_2}\binom{n}{p_1}\binom{n}{p_2}|Iso(R_1)||Iso(R_2)|\right)}\mathbb{E}\left[\sum_{\substack{S,T\subseteq K_n\\ S\cong R_1, T\cong R_2, S\cap T\neq\emptyset}} \mathbf{1}(S,T\subseteq G)\right]\\
&= \frac{1}{\left(\rho^{e_1+e_2}\binom{n}{p_1}\binom{n}{p_2}|Iso(R_1)||Iso(R_2)|\right)} \sum_{\substack{W:W=S\cup T\\ S\cong R_1 T\cong R_2, S\cap T\neq\emptyset}} \mathbb{E}\left[\sum_{W\subseteq K_n} \mathbf{1}(W\subseteq G)\right]
\end{aligned}
$$

B2. *Proof of Lemma 7.* Let us define

$$
\tilde{\sigma}^2(R) = \frac{1/(1-x)}{\left(\rho_n^e\binom{n}{p}|Iso(R)|\right)^2} \sum_{\substack{W\subseteq K_n: W=S\cup T,\\ S,T\cong R, |S\cap T|=1,p}} \mathbf{1}(W\subseteq G) - \frac{x\rho_n^{-2e}\hat{P}(R)^2}{(1-x)}
$$

where, $x = \left(1 - \frac{((n-p)!)^2}{n!(n-2p)!}\right)$ and

$$
\begin{aligned}
\tilde{\sigma}(R_1, R_2) = &\frac{1/(1-y)}{\left(\rho_n^{e_1+e_2}\binom{n}{p_1}\binom{n}{p_2}|Iso(R_1)||Iso(R_2)|\right)}\\
&\sum_{\substack{W\subseteq K_n, W=S\cup T,\\ S\cong R_1, T\cong R_2, |S\cap T|=1}} \mathbf{1}(W\subseteq G) - \frac{y\rho_n^{-(e_1+e_2)}\hat{P}(R_1)\hat{P}(R_2)}{(1-y)}.
\end{aligned}
$$

where, $y = \left(1 - \frac{(n-p_1)!(n-p_2)!}{n!(n-p_1-p_2)!}\right)$.

Now,

$$\mathbb{E}\left[\tilde{\sigma}^2(R)\right] = \frac{1/(1-x)}{\left(\rho_n^e\binom{n}{p}|Iso(R)|\right)^2}\mathbb{E}\left[\sum_{\substack{W\subseteq K_n:W=S\cup T,\\ S,T\cong R,|S\cap T|=1,p}} \mathbf{1}(W\subseteq G)\right] - \frac{x\rho_n^{-2e}\mathbb{E}\left[\hat{P}(R)^2\right]}{(1-x)}$$

$$= \frac{1/(1-x)}{\left(\rho_n^e\binom{n}{p}|Iso(R)|\right)^2}\mathbb{E}\left[\sum_{\substack{W\subseteq K_n:W=S\cup T,\\ S,T\cong R,|S\cap T|=1,p}} \mathbf{1}(W\subseteq G)\right] - \frac{x\rho_n^{-2e}\left[\mathrm{Var}\left(\hat{P}(R)\right)+P(R)^2\right]}{(1-x)}$$

$$= \frac{1}{1-x}\frac{1}{\left(\rho_n^e\binom{n}{p}|Iso(R)|\right)^2}\mathbb{E}\left[\sum_{\substack{W\subseteq K_n:W=S\cup T,\\ S,T\cong R,|S\cap T|=1,p}} \mathbf{1}(W\subseteq G)\right] - \frac{x\tilde{P}(R)^2}{1-x} - \frac{x\rho_n^{-2e}\mathrm{Var}\left(\hat{P}(R)\right)}{(1-x)}$$

$$= \frac{1}{1-x}\frac{1}{\left(\rho^e\binom{n}{p}|Iso(R)|\right)^2}\sum_{\substack{W:W=S\cup T\\ S,T\cong R,S\cap T\neq\emptyset}}\mathbb{E}\left[\sum_{W\subseteq K_n}\mathbf{1}(W\subseteq G)\right] - \frac{x\tilde{P}(R)^2}{1-x}$$

$$-\frac{1}{(1-x)\left(\rho_n^e\binom{n}{p}|Iso(R)|\right)^2}\mathbb{E}\left[\sum_{\substack{W\subseteq K_n:W=S\cup T,\\ S,T\cong R,1<|S\cap T|<p}} \mathbf{1}(W\subseteq G)\right] - \frac{x\rho_n^{-2e}\mathrm{Var}\left(\hat{P}(R)\right)}{(1-x)}$$

$$= \frac{\mathrm{Var}\left(\rho_n^{-e}\hat{P}(R)\right)}{(1-x)} - \frac{x\rho_n^{-2e}\mathrm{Var}\left(\hat{P}(R)\right)}{(1-x)} - o\left(\mathrm{Var}\left(\rho_n^{-e}\hat{P}(R)\right)\right)$$

$$= \mathrm{Var}\left(\rho_n^{-e}\hat{P}(R)\right) - o\left(\mathrm{Var}\left(\rho_n^{-e}\hat{P}(R)\right)\right)$$

Similarly, we get,

$$\mathbb{E}\left[\tilde{\sigma}(R_1,R_2)\right] = \mathrm{Cov}\left(\rho_n^{-e_1}\hat{P}(R_1),\rho_n^{-e_2}\hat{P}(R_2)\right) - o\left(\mathrm{Cov}\left(\rho_n^{-e_1}\hat{P}(R_1),\rho_n^{-e_2}\hat{P}(R_2)\right)\right)$$

Now, from the Theorem 1(a) in [1], we know that as $\lambda_n\to\infty$, if $\hat{\rho}_n = \frac{\bar{D}}{n-1}$ as defined in (2.6),

$$\frac{\hat{\rho}_n}{\rho_n} \xrightarrow{P} 1$$

So, using the estimate $\hat{\rho}_n$, we get that,

$$\frac{\hat{\sigma}^2(R)}{\sigma^2(R;\rho)} \xrightarrow{P} 1, \quad \frac{\hat{\sigma}(R_1,R_2)}{\sigma(R_1,R_2;\rho)} \xrightarrow{P} 1$$

B3. *Proof of Lemma 8.* Given G,

$$\mathbb{E}\left[\hat{\sigma}^2_{Bi}(R)|G\right]$$
$$= \sum_{\substack{W=S\cup T, S,T\cong R,\\ |S\cap T|=1,p}} \frac{\left(\hat{\rho}_n^{e_W}\binom{n}{p_W}|Iso(R)|\right)}{(1-x)\left(\hat{\rho}_n^{e}\binom{n}{p}|Iso(R)|\right)^2}\mathbb{E}\left[\bar{P}_{Bi}(W)|G\right] - \frac{x\mathbb{E}\left[\hat{\rho}_n^{-2e}\bar{P}_{Bi}(R)^2|G\right]}{(1-x)}$$
$$= \sum_{\substack{W=S\cup T, S,T\cong R,\\ |S\cap T|=1,p}} \frac{\left(\hat{\rho}_n^{e_W}\binom{n}{p_W}|Iso(R)|\right)}{(1-x)\left(\hat{\rho}_n^{e}\binom{n}{p}|Iso(R)|\right)^2}\hat{P}(W) - \frac{x\hat{\rho}_n^{-2e}\hat{P}(R)^2}{(1-x)} - \frac{x\mathrm{Var}\left(\hat{T}_{Bi}(R)\right)}{1-x}$$
$$= \hat{\sigma}^2(R) - \frac{x\mathrm{Var}\left(\hat{T}_{Bi}(R)\right)}{1-x} = \hat{\sigma}^2(R) - o\left(\hat{\sigma}^2(R)\right)$$

where the last inequality follows since $x = O\left(\frac{1}{n}\right)$ and Theorem 3.1 and Theorem 3.2 for $i = 1, 2$.

Similarly, we get,

$$\mathbb{E}\left[\hat{\sigma}_{Bi}(R_1, R_2)|G\right] = \hat{\sigma}(R_1, R_2) - o\left(\hat{\sigma}^2(R)\right)$$

So using Lemma 7, we have that,

$$\frac{\hat{\sigma}^2_{Bi}(R)}{\sigma^2(R;\rho)} \xrightarrow{P} 1, \qquad \frac{\hat{\sigma}_{Bi}(R_1, R_2)}{\sigma(R_1, R_2;\rho)} \xrightarrow{P} 1 \quad \text{for } i = 1, 2$$

References.

[1] Bickel, P. J., Chen, A. and Levina, E. (2011). The method of moments and degree distributions for network models. *Ann. Statist.* **39** 2280–2301. . MR2906868

[2] Thompson, S. K. (2012). *Sampling*, third ed. *Wiley Series in Probability and Statistics*. John Wiley & Sons Inc., Hoboken, NJ. . MR2894042 (2012k:62024)

Department of Statistics
44 Kidder Hall
Corvallis, OR, 97331
E-mail: bhattash@science.oregonstate.edu

Department of Statistics
367 Evans Hall
Berkeley, CA, 94720
E-mail: bickel@stat.berkeley.edu