Jackknife Estimate of Large Networks (Proof-2- kcycle and kstar)

Qiaohui Lin

1 Consistency of Jackknife Estimate on k-cycles Counts

Theorem 1.1. *G_n is a size n unweighted undirected graph with latent position $\xi_1, , \xi_n$ and graphon $w(\xi_i, \xi_j)$ and sparcity parameter ρ_n, $n\rho_n \approx \infty$. k-cycle is a k-node and k-edge closed cycle in G_n, with $k \geq 3, k^2 << n$. Let Z_n be the average number of k-cycles. Z_i is the average number of k-cycles in the graph when leaving out node i, $\bar{Z}_i$ is the mean of all Z_i for $1 \leq i \leq n$. Then we have, both $(n-1)E\sum_{i=1}^n (Z_i - \bar{Z}_i)^2$ and $(n-1)Var(Z_{n-1})$ stabilize at non-zero constants, and*

$$(n-1)E\sum_{i=1}^n (Z_i - \bar{Z}_i)^2 - (n-1)Var(Z_{n-1}) = \frac{k^2-1}{n^{2k-2}} E\left[Var\left(\sum_{t_1, t_2,, t_{k-1}}^{t_1 \neq,, \neq t_{k-1} \neq i} \prod_{\substack{g_1, g_2 \\ \in \{t_1, t_2,, t_{k-1}, i\}}} w(\xi_{g_1}, \xi_{g_2})) \Big| \xi_i \right) \right] + o(1)$$
$$\to 0 \quad as\ n \to \infty \tag{1}$$

Proof 1.1.

Denote X_n as the total number of k-cycles in graph G_n. By definition,

$$Z_n = \frac{X_n}{\binom{n}{k} \rho_n^k} \tag{2}$$

Denote X_i the number of k-cycles node i is involved in. $EX_i = EX_j$ for any node i, j when unconditioned on latent positions. The total number of k-cycles in this graph when leaving out node i, is X_n minus the number of k-cycles the node i is involved in. Thus,

$$Z_i = \frac{X_n - X_i}{\binom{n-1}{k} \rho_n^k} = \frac{X_n - X_i}{c_n} \tag{3}$$

Then we can write

$$E\sum_{i=1}^n (Z_i - \bar{Z}_i)^2 = \frac{1}{2n} \sum_{i \neq j} E(Z_i - Z_j)^2 = \frac{1}{2n} \sum_{i \neq j} E\left(\frac{X_i - X_j}{c_n} \right)^2 = (n-1)Var\left(\frac{X_i}{c_n} \right) - \frac{\sum_{i \neq j} cov(X_i, X_j)}{nc_n^2} \tag{4}$$

whereas the total number of k-cycles in a $(n-1)$ graph is $X_{n-1} = \sum_{i=1}^{n-1} X_i/k$ as each k-cycle is counted k times from each node.

$$Var(Z_{n-1}) = Var\left(\frac{\sum_{i=1}^{n-1} X_i/k}{\binom{n-1}{k}\rho_{n-1}^k}\right) = Var\left(\frac{\sum_{i=1}^{n-1} X_i/k}{c_{n-1}}\right) = \frac{n-1}{k^2}Var\left(\frac{X_i}{c_{n-1}}\right)+\frac{1}{k^2}\sum_{i,j,i\neq j} cov\left(\frac{X_i}{c_{n-1}}, \frac{X_j}{c_{n-1}}\right) \tag{5}$$

We assume $\rho_n \approx \rho_{n-1}$, $c_{n-1} \approx c_n$ Apply law of total variance gives us

$$\sum_{i=1}^{n-1} Var\left(\frac{X_i}{c_n}\right) = \sum_{i=1}^{n-1} E\left[Var\left(\frac{X_i}{c_n}\middle|\xi\right)\right] + \sum_{i=1}^{n-1} Var\left[E\left(\frac{X_i}{c_n}\middle|\xi\right)\right] \tag{6}$$

$$\sum_{i,j,i\neq j} cov\left(\frac{X_i}{c_n}, \frac{X_j}{c_n}\right) = \sum_{i,j,i\neq j} E\left[cov\left(\frac{X_i}{c_n}, \frac{X_j}{c_n}\middle|\xi\right)\right] + \sum_{i,j,i\neq j} cov\left(E\frac{X_i}{c_n}|\xi, E\frac{X_i}{c_n}\middle|\xi,\right) \tag{7}$$

Following Bickle et.al (2011), the expectation of variance(covariance) term is of smaller order compared to the variance(covariance) of expectation term, d is the number of edge X_i and X_j share and assume $n\rho_n \to \infty$,

$$\sum_{i,j,i\neq j} E\left[cov\left(\frac{X_i}{c_n}, \frac{X_j}{c_n}\middle|\xi\right)\right] \leq O(n^{-1}(n\rho_n)^{-d}) = o(n^{-1}); \quad \sum_{i=1}^{n-1} E\left[Var\left(\frac{X_i}{c_n}\middle|\xi\right)\right] \leq O(n^{-1}(n\rho_n)^{-k}) = o(n^{-1}) \tag{8}$$

For the variance(covariance) of expectation term, for any fixed i,

$$Var\left[E\frac{X_i}{c_n}\middle|\xi\right] = VarE\left(\frac{\frac{1}{(k-1)!}\sum_{t_1,,t_{k-1}}^{t_1\neq,,\neq t_{k-1}\neq i} \mathbf{1}(A_{g_1,g_2\in\{t_1,,t_{k-1},i\}}=1)}{\frac{(n-1)!}{k!(n-1-k)!}\rho_n^k}\middle|\xi\right)$$

$$= \frac{k^2}{n^{2k}}Var\left(\sum_{t_1,,t_{k-1}}^{t_1\neq,,\neq t_{k-1}\neq i} \prod_{\substack{g_1,g_2\\ \in\{t_1,t_2,,t_{k-1},i\}}} w(\xi_{g_1},\xi_{g_2}))\right)$$

$$= \frac{k^2}{n^{2k}}EVar\left(\sum_{t_1,,t_{k-1}}^{t_1\neq,,\neq t_{k-1}\neq i} \prod_{\substack{g_1,g_2\\ \in\{t_1,t_2,,t_{k-1},i\}}} w(\xi_{g_1},\xi_{g_2})\middle|\xi_i\right) + \frac{k^2}{n^{2k}}VarE\left(\sum_{t_1,,t_{k-1}}^{t_1\neq,,\neq t_{k-1}\neq i} \prod_{\substack{g_1,g_2\\ \in\{t_1,t_2,,t_{k-1},i\}}} w(\xi_{g_1},\xi_{g_2})\middle|\xi_i\right) \tag{9}$$

The second term can bring the expection into the sum, turing it into

$$\frac{k^2}{n^{2k}}VarE\left(\sum_{t_1,,t_{k-1}}^{t_1\neq,,\neq t_{k-1}\neq i} \prod_{\substack{g_1,g_2\\ \in\{t_1,t_2,,t_{k-1},i\}}} w(\xi_{g_1},\xi_{g_2})\middle|\xi_i\right) = \frac{k^2}{n^2}VarE\left(\prod_{\substack{g_1,g_2\\ \in\{1,2,,,,k\}}} w(\xi_{g_1},\xi_{g_2})\middle|\xi_1\right) \tag{10}$$

Similarly, for any fixed i and j, covariance of conditional expectation can be written as

$$
\begin{aligned}
& cov\left[E\frac{X_i}{c_n}\Big|\xi, E\frac{X_j}{c_n}\Big|\xi\right] \\
& = cov\left[E\frac{\frac{1}{(k-1)!}\sum_{t_1,,t_{k-1}}^{t_1\neq,,\neq t_{k-1}\neq i}\mathbf{1}(A_{g_1,g_2\in\{t_1,,t_{k-1},i\}}=1)}{\frac{(n-1)!}{k!(n-1-k)!}\rho_n^k}\Bigg|\xi, E\frac{\frac{1}{(k-1)!}\sum_{s_1,,s_{k-1}}^{s_1\neq,,\neq s_{k-1}\neq j}\mathbf{1}(A_{g_1,g_2\in\{s_1,,s_{k-1},i\}}=1)}{\frac{(n-1)!}{k!(n-1-k)!}\rho_n^k}\Bigg|\xi\right] \\
& = \frac{k^2}{n^{2k}}\sum_{t_1,,t_{k-1}}^{t_1\neq,,\neq t_{k-1}\neq i}\sum_{s_1,,s_{k-1}}^{s_1\neq,,\neq s_{k-1}\neq i} cov\left(\prod_{\substack{g_1,g_2\\ \in\{t_1,t_2,,t_{k-1},i\}}} w(\xi_{g_1},\xi_{g_2}), \prod_{\substack{q_1,q_2\\ \in\{s_1,s_2,,s_{k-1},j\}}} w(\xi_{q_1},\xi_{q_2})\right)
\end{aligned}
\tag{11}
$$

Here we split the total sum of covariance into a combinatoric combination based on the number of nodes that set $S_i=\{i,t_1,,,t_{k-1}\}$ and set $S_j=\{j,s_1,,s_{k-1}\}$ share, $|S_i\cap S_j|$. In $\sum_{i\neq j} cov\left[E\frac{X_i}{c_n}\Big|\xi, E\frac{X_j}{c_n}\Big|\xi\right]$, the sum of $|S_i\cap S_j|=a$ is of order $n(n-1)\binom{n-2}{2k-2-a}/n^{2k}=O(n^{-a})$. Thus, we only keep those $|S_i\cap S_j|=1$, i.e.,

$$
\sum_{|S_i\cap S_j|=1} cov\left(\prod_{\substack{g_1,g_2\\ \in\{t_1,t_2,,t_{k-1},i\}}} w(\xi_{g_1},\xi_{g_2}), \prod_{\substack{q_1,q_2\\ \in\{s_1,s_2,,s_{k-1},j\}}} w(\xi_{q_1},\xi_{q_2})\right) \tag{12}
$$

For any fixed i and j, $|S_i\cap S_j|=1$ means that there is one common node in $S_i=\{i,t_1,,,t_{k-1}\}$ and set $S_j=\{j,s_1,,s_{k-1}\}$ while $i\neq j$, which has n^2-1 cases. Thus, (11) can be continued as (WLOG write $i=s_1$)

$$
\begin{aligned}
(11) &= \frac{k^2}{n^{2k}}\left[(n-2)(n-3)..(n-(2k-2))*(k^2-1)cov\left(\prod_{\substack{g_1,g_2\\ \in\{t_1,t_2,,t_{k-1},i\}}} w(\xi_{g_1},\xi_{g_2}), \prod_{\substack{q_1,q_2\\ \in\{i,s_2,,s_{k-1},j\}}} w(\xi_{q_1},\xi_{q_2})\right)\right] \\
&= \frac{k^2}{n^3}*(k^2-1)*cov\left(\prod_{\substack{g_1,g_2\\ \in\{t_1,t_2,,t_{k-1},i\}}} w(\xi_{g_1},\xi_{g_2}), \prod_{\substack{q_1,q_2\\ \in\{i,s_2,,s_{k-1},j\}}} w(\xi_{q_1},\xi_{q_2})\right)
\end{aligned}
\tag{13}
$$

We further condition this covariance on ξ_i. The first equality below holds as conditioned on ξ_i, $\prod_{\substack{g_1,g_2\\ \in\{t_1,t_2,,t_{k-1},i\}}} w(\xi_{g_1},\xi_{g_2})$ and $\prod_{\substack{q_1,q_2\\ \in\{i,s_2,,s_{k-1},j\}}} w(\xi_{q_1},\xi_{q_2})$ are independent, the $E(cov(\cdot))$ part is thus 0.

$$
cov\left(\prod_{\substack{g_1,g_2\\ \in\{t_1,t_2,,t_{k-1},i\}}} w(\xi_{g_1},\xi_{g_2}), \prod_{\substack{q_1,q_2\\ \in\{i,s_2,,s_{k-1},j\}}} w(\xi_{q_1},\xi_{q_2})\right) = Var\left(E\prod_{\substack{g_1,g_2\\ \in\{t_1,t_2,,t_{k-1},j\}}} w(\xi_{g_1},\xi_{g_2})\Big|\xi_i\right)
$$

Thus, the scaled Jackknife estimate and true variance have the expressions below:

$$(n-1)E\sum_{i=1}^{n}(Z_i-\bar{Z}_i)^2 = \frac{k^2}{n^{2k-2}}EVar\left(\sum_{t_1,,t_{k-1}}^{t_1\neq,,\neq t_{k-1}\neq i}\prod_{\substack{g_1,g_2\\ \in\{t_1,t_2,,t_{k-1},i\}}} w(\xi_{g_1},\xi_{g_2})\middle|\xi_i\right) \tag{14}$$

$$+ k^2VarE\left(\prod_{\substack{g_1,g_2\\ \in\{t_1,t_2,,,,t_k,i\}}} w(\xi_{g_1},\xi_{g_2})\middle|\xi_i\right) + O(\frac{1}{n\rho_n}) \tag{15}$$

$$(n-1)Var(Z_{n-1}) \tag{16}$$

$$= \frac{1}{n^{2k-2}}EVar\left(\sum_{t_1,,t_{k-1}}^{t_1\neq,,\neq t_{k-1}\neq i}\prod_{\substack{g_1,g_2\\ \in\{t_1,t_2,,t_{k-1},i\}}} w(\xi_{g_1},\xi_{g_2})\middle|\xi_i\right) + VarE\left(\prod_{\substack{g_1,g_2\\ \in\{t_1,t_2,,,,t_k,i\}}} w(\xi_{g_1},\xi_{g_2})\middle|\xi_i\right) \tag{17}$$

$$+ (k^2-1)*cov\left(\prod_{\substack{g_1,g_2\\ \in\{t_1,t_2,,t_{k-1},i\}}} w(\xi_{g_1},\xi_{g_2}), \prod_{\substack{q_1,q_2\\ \in\{i,s_2,,s_{k-1},j\}}} w(\xi_{q_1},\xi_{q_2})\right)$$

$$= \frac{1}{n^{2k-2}}EVar\left(\sum_{t_1,,t_{k-1}}^{t_1\neq,,\neq t_{k-1}\neq i}\prod_{\substack{g_1,g_2\\ \in\{t_1,t_2,,t_{k-1},i\}}} w(\xi_{g_1},\xi_{g_2})\middle|\xi_i\right) + k^2VarE\left(\prod_{\substack{g_1,g_2\\ \in\{t_1,t_2,,,,t_k,i\}}} w(\xi_{g_1},\xi_{g_2})\middle|\xi_i\right)$$

The difference between the two is thus,

$$(n-1)E(Z_i-\bar{Z}_i)^2-(n-1)Var(Z_{n-1}) = \frac{k^2-1}{n^{2k-2}}E\left[Var\left(\sum_{t_1,t_2,,t_{k-1}}^{t_1\neq,,\neq t_{k-1}\neq i}\prod_{\substack{g_1,g_2\\ \in\{t_1,t_2,,t_{k-1},i\}}} w(\xi_{g_1},\xi_{g_2}))\middle|\xi_i\right)\right]+o(1) \tag{18}$$

where

$$\frac{1}{n^{2k-2}}E\left[Var\left(\sum_{t_1,t_2,,t_{k-1}}^{t_1\neq,,\neq t_{k-1}\neq i}\prod_{\substack{g_1,g_2\\ \in\{t_1,t_2,,t_{k-1},i\}}} w(\xi_{g_1},\xi_{g_2}))\middle|\xi_i\right)\right]$$

$$= E\left[Var\frac{1}{n^{k-1}}\left(\sum_{t_1,t_2,,t_{k-1}}^{t_1\neq,,\neq t_{k-1}\neq i}\prod_{\substack{g_1,g_2\\ \in\{t_1,t_2,,t_{k-1},i\}}} w(\xi_{g_1},\xi_{g_2}))\middle|\xi_i\right)\right] \to 0$$

which is asymptotically 0 as it is the expectation of variance of the U-statistic.

2 Consistency of Jackknife Estimate on k-star Counts

Theorem 2.1. *In the G_n defined as in Theorem 1, k-star is a shape that contains k edges linked to one node, with $k \geq 2, k^2 << n$. Let Z_n be the average number of k-stars. Z_i is the average number of k-stars in the graph when leaving out node i, $\bar{Z}_i$ is the mean of all Z_i for $1 \leq i \leq n$. Then we have, both $(n-1)E\sum_{i=1}^{n}(Z_i - \bar{Z}_i)^2$ and $(n-1)Var(Z_{n-1})$ stabilize at non-zero constants, and*

$$(n-1)E\sum_{i=1}^{n}(Z_i - \bar{Z}_i)^2 - (n-1)Var(Z_{n-1}) = \frac{k^2(k+1)^2}{n^{2k}} EVar[\sum_{t_1,,t_k}^{t_1 \neq t_2,, \neq i} w(\xi_{t_1}, \xi_i)w(\xi_{t_1}, \xi_{t2})..w(\xi_{t1}, \xi_{t_k})|\xi_i] + o(1)$$
$$\to 0 \quad as\ n \to \infty \tag{19}$$

Proof 2.1.

Denote Y_n as the total number of k-stars in graph G_n. By definition,

$$Z_n = \frac{Y_n}{\binom{n}{k+1}\rho_n^k} \tag{20}$$

Denote Y_i the number of k-star node i has as a root node, Y^i the number of k-star node i has as an end node. Thus the total number of k-stars in complete graph is $Y_n = \sum_i Y_i$; the total number of k-stars when leaving out node i, is $Y_n - Y_i - Y^i$. Thus,

$$Z_i = \frac{Y_n - Y_i - Y^i}{\binom{n-1}{k+1}\rho_n^k} = \frac{Y_n - Y_i - Y^i}{\gamma_n} \tag{21}$$

$$E\sum_{i=1}^{n}(Z_i - \bar{Z}_i)^2 = \frac{1}{2n}\sum_{i \neq j} E\left(\frac{(Y_i + Y^i) - (Y_j + Y^j)}{\gamma_n}\right)^2 = (n-1)Var\left(\frac{Y_i + Y^i}{\gamma_n}\right) - \frac{\sum_{i \neq j} cov\left(Y_i + Y^i, Y_j + Y^j\right)}{n\gamma_n^2} \tag{22}$$

while the true variance is

$$Var(Z_{n-1}) = Var\left(\frac{\sum_{i=1}^{n-1} Y_i}{\binom{n-1}{k+1}\rho_{n-1}^k}\right) = Var\left(\frac{\sum_{i=1}^{n-1} Y_i}{\gamma_{n-1}}\right) = \frac{n-1}{k^2}Var\left(\frac{X_i}{c_{n-1}}\right) + \frac{1}{k^2}\sum_{i,j,i \neq j} cov\left(\frac{Y_i}{\gamma_{n-1}}, \frac{Y_j}{\gamma_{n-1}}\right) \tag{23}$$

Following the same steps as in Proof 1, $\sum EVar(\cdot|\xi)$ and $\sum ECov(\cdot|\xi)$ are of smaller order $o(n^{-1})$, and $VarE(\cdot|\xi)$ and $CovE(\cdot|\xi)$ are calculated as below.

$$Var(E\frac{Y_i}{\gamma_n}\Big|\xi) = Var\left(E\frac{\frac{1}{k!}\sum_{t_1,,t_k \neq i} 1(A_{it_1} = A_{it_2} = .. = A_{it_k} = 1)}{\frac{(n-1)!}{(k+1)!(n-k-2)!}\rho_n^k}\Bigg|\xi\right) \tag{24}$$

$$= Var\left(\frac{k+1}{n^{k+1}}\sum_{t_1,,t_k\neq i} w(\xi_i,\xi_{t1})..w(\xi_i,\xi_{t_k})\right)$$

$$= \frac{(k+1)^2}{n^{2k+2}} E\left[Var\sum_{t_1,,t_k\neq i} w(\xi_i,\xi_{t1})..w(\xi_i,\xi_{t_k})\Big|\xi_i\right] + \frac{(k+1)^2}{n^2} Var\left[Ew(\xi_i,\xi_{t1})..w(\xi_i,\xi_{t_k})\Big|\xi_i\right]$$

The covariance of expectation include multiple cases of covariance of root stars, end stars, and root and end stars. First, the covariance of root stars with root i and root j is

$$Cov\left(E\frac{Y_i}{\gamma_n}\Big|\xi, E\frac{Y_j}{\gamma_n}\Big|\xi\right) \tag{25}$$

$$= Cov\left(E\frac{\frac{1}{k!}\sum_{t_1\neq..t_k\neq i} 1(A_{it_1}=A_{it_2}=..=A_{it_k}=1)}{\frac{(n-1)!}{(k+1)!(n-k-2)!}\rho_n^k}\Big|\xi, E\frac{\frac{1}{k!}\sum_{s_1\neq..s_k\neq i} 1(A_{is_1}=A_{is_2}=..=A_{is_k}=1)}{\frac{(n-1)!}{(k+1)!(n-k-2)!}\rho_n^k}\Big|\xi\right)$$

$$= \frac{(k+1)^2}{n^{2k+2}} \sum_{t_1\neq..t_k\neq i}\sum_{s_1\neq..s_k\neq i} cov(w(\xi_i,\xi_{t_1})..w(\xi_i,\xi_{t_k}), w(\xi_j,\xi_{s_1})..w(\xi_j,\xi_{s_k}))$$

Following the same stategy of calculating covariance of only keep $|S_i\cap S_j|=1$, only one overlap in the two sets $S_i=\{i,t_1,,,t_k\}$ and $S_j=\{j,s_1,,s_k\}$, which means there are $2(k+1)+1$ different nodes in the two sets, i.e. $2k-1$ nodes besides i adn j. Since j and i are both root nodes, there are k^2 cases where the two end nodes overlap (WLOG $t_1=s_1$), and $2k$ cases where one root node is overlapping with another's end node (WLOG $i=s_1$). Thus, (25) can be continued as

$$(25) = \frac{(k+1)^2}{n^{2k+2}} \sum_{|S_i\cap S_j|=1} cov(w(\xi_i,\xi_{t_1})..w(\xi_i,\xi_{t_k}), w(\xi_j,\xi_{s_1})..w(\xi_j,\xi_{s_k}))$$

$$= \frac{(k+1)^2}{n^{2k+2}} n^{2k-1}\left[k^2 cov\left(\prod_{g=t_1,,t_k} w(\xi_i,\xi_g), \prod_{q=t_1,s_2,,s_k} w(\xi_j,\xi_q)\right) + 2k cov\left(\prod_{g=t_1,,t_k} w(\xi_i,\xi_g), \prod_{u=i,s_2,,s_k} w(\xi_j,\xi_u)\right)\right] \tag{26}$$

Plug in (26) and (24) to (23) and scale by $n-1$, we can get the scaled true variance of average number of stars

$$(n-1)Var(Z_{n-1}) = \frac{(n-1)^2}{k^2} Var\left(\frac{X_i}{c_{n-1}}\right) + \frac{(n-1)}{k^2}\sum_{i,j,i\neq j} cov\left(\frac{Y_i}{\gamma_{n-1}}, \frac{Y_j}{\gamma_{n-1}}\right) \tag{27}$$

$$= o(1) + \frac{(k+1)^2}{n^{2k}} E\left[Var\sum_{t_1,,t_k\neq i}\prod_{g=t_1,,t_k} w(\xi_i,\xi_g))\Big|\xi_i\right] + (k+1)^2 Var\left[E\prod_{g=t_1,,t_k} w(\xi_i,\xi_g))\Big|\xi_i\right]$$

$$+ (k+1)^2k^2 cov\left(\prod_{g=t_1,,t_k} w(\xi_i,\xi_g), \prod_{q=t_1,s_2,,s_k} w(\xi_j,\xi_q)\right) + \boxed{2k(k+1)^2 cov\left(\prod_{g=t_1,,t_k} w(\xi_i,\xi_g), \prod_{u=i,s_2,,s_k} w(\xi_j,\xi_u)\right)}$$

To calculate the Jackknife estimate, we need a few more ingredients, Similar to (25), the covariance of root i and end j stars, end i and end j stars are of same order,

$$Cov\left(E\frac{Y_i}{\gamma_n}\bigg|\xi, E\frac{Y^j}{\gamma_n}\bigg|\xi\right) = O(n^{-3}); Cov\left(E\frac{Y^i}{\gamma_n}\bigg|\xi, E\frac{Y^j}{\gamma_n}\bigg|\xi\right) = O(n^{-3}) \tag{28}$$

We also need covariance between root i star and end i star, with which we only consider $S_i \cap S^i = \{i\}$, i.e.,$t_1 \neq t_2, , , \neq t_k \neq t'_1 \neq, , , t'_k$

$$Cov\left(E\frac{Y_i}{\gamma_n}\bigg|\xi, E\frac{Y^j}{\gamma_n}\bigg|\xi\right) \tag{29}$$

$$= Cov\left(E\frac{\frac{1}{k!}\sum_{t_1\neq..t_k\neq i} 1(A_{it_1} = A_{it_2} = .. = A_{it_k} = 1)}{\frac{(n-1)!}{(k+1)!(n-k-2)!}\rho_n^k}\Bigg|\xi, E\frac{\frac{1}{(k-1)!}\sum_{t'_1\neq..t'_k\neq i} 1(A_{t'_1 i} = A_{t'_1 t'_2} = .. = A_{t'_1 t'_k} = 1)}{\frac{(n-1)!}{(k+1)!(n-k-2)!}\rho_n^k}\Bigg|\xi\right)$$

$$= \frac{k^2(k+1)^2}{n^{2k+2}} \sum_{t_1\neq..t_k\neq i}\sum_{t'_1\neq..t'_k\neq i} cov(w(\xi_i,\xi_{t_1})..w(\xi_i,\xi_{t_k}), w(\xi_j,\xi_{s_1})..w(\xi_j,\xi_{s_k}))$$

$$= \frac{k^2(k+1)^2}{n^{2k+2}} * n^{2k} cov\left(\prod_{g=t_1,,t_k} w(\xi_i,\xi_g), \prod_{m=i,t'_2,,t'_k} w(\xi_{t'_1},\xi_m)\right)$$

Also the variance of end i star is needed

$$Var(E\frac{Y^i}{\gamma_n}\bigg|\xi) = Var\left(E\frac{\frac{1}{(k-1)!}\sum_{t'_1,,t'_k\neq i} 1(A_{it'_1} = A_{it'_2} = .. = A_{it'_k} = 1)}{\frac{(n-1)!}{(k+1)!(n-k-2)!}\rho_n^k}\Bigg|\xi\right) \tag{30}$$

$$= Var\left(\frac{k(k+1)}{n^{k+1}}\sum_{t'_1,,t'_k\neq i} w(\xi_{t_1},\xi_i)..w(\xi_{t_1},\xi_{t_k})\right)$$

$$= \frac{k^2(k+1)^2}{n^{2k+2}} E\left[Var\sum_{t'_1,,t'_k\neq i}\prod_{g=i,t'_2,,t'_k} w(\xi_{t'_1},\xi_g)\bigg|\xi_i\right] + \frac{k^2(k+1)^2}{n^2} Var\left[E\prod_{g=i,t'_2,,t'_k} w(\xi_{t'_1},\xi_g)\bigg|\xi_i\right]$$

Then in (22), the jackknife estimate, the second term, taking apart the covariance sum, is of order $(O(n^2)O(n^{-3})/n) = O(n^{-2})$. Scale jackknife by n, this term is $O(n^{-1})$, which is of smaller order than the first term. So here we only carry down the first term of (22) and through scale and transform, it turns into,

$$(n-1)E\sum_{i=1}^{n}(Z_i - \bar{Z}_i)^2 = (n-1)^2\left(Var\left(\frac{Y_i}{\gamma_n}\right) + Var\left(\frac{Y^i}{\gamma_n}\right) + 2cov\left(\frac{Y_i}{\gamma_n}, \frac{Y^i}{\gamma_n}\right))\right)$$

$$= o(1) + \frac{(k+1)^2}{n^{2k}} E\left[Var\sum_{t_1,,t_k\neq i}\prod_{g=t_1,,t_k} w(\xi_i,\xi_g))\bigg|\xi_i\right] + (k+1)^2 Var\left[E\prod_{g=t_1,,t_k} w(\xi_i,\xi_g))\bigg|\xi_i\right]$$

$$+\frac{k^2(k+1)^2}{n^{2k}}E\left[Var\sum_{t'_1,,t'_k\neq i}\prod_{g=i,t'_2,,t'_k}w(\xi_{t'_1},\xi_g)\middle|\xi_i\right]+\frac{k^2(k+1)^2}{n^2}Var\left[E\prod_{g=i,t'_2,,t'_k}w(\xi_{t'_1},\xi_g)\middle|\xi_i\right]$$

$$+\boxed{2k(k+1)^2cov\left(\prod_{g=t_1,,t_k}w(\xi_i,\xi_g),\prod_{u=i,s_2,,s_k}w(\xi_j,\xi_u)\right)} \tag{31}$$

Since on symmetry

$$cov\left(\prod_{g=t_1,,t_k}w(\xi_i,\xi_g),\prod_{q=t_1,s_2,,s_k}w(\xi_j,\xi_q)\right)=cov\left(\prod_{g=i,t_2,,,t_k}w(\xi_{t_1},\xi_g),\prod_{q=i,s_2,,s_k}w(\xi_j,\xi_q)\right) \tag{32}$$

condition on ξ_i, the two terms inside the covariance are independent

$$cov\left(\prod_{g=i,t_2,,,t_k}w(\xi_{t_1},\xi_g),\prod_{q=i,s_2,,s_k}w(\xi_j,\xi_q)\right)=Var\left[E\prod_{g=i,t'_2,,t'_k}w(\xi_{t'_1},\xi_g)\middle|\xi_i\right] \tag{33}$$

Then every term is the same in $(n-1)E\sum_{i=1}^n(Z_i-\bar{Z}_i)^2$ and $(n-1)Var(Z_{n-1})$ is the same besides one,

$$(n-1)[E\sum_{i=1}^n(Z_i-\bar{Z}_i)^2-Var(Z_{n-1})]=k^2(k+1)^2EVar\left(\frac{1}{n^k}\sum_{t_1,,,t_k\neq i}w(\xi_{t_1},\xi_i)w(\xi_{t_1},\xi_{t_2})..w(\xi_{t_1},\xi_{t_k})|\xi_i\right)+o(1) \tag{34}$$

where from symmetry,

$$\begin{aligned}\sum_{t_1,,,t_k\neq i}w(\xi_{t_1},\xi_i)w(\xi_{t_1},\xi_{t_2})..w(\xi_{t_1},\xi_{t_k})=\frac{1}{k}\Bigg[&\sum_{t_1,,,t_k\neq i}w(\xi_{t_1},\xi_i)w(\xi_{t_1},\xi_{t_2})..w(\xi_{t_1},\xi_{t_k})\\&+\sum_{t_1,,,t_k\neq i}w(\xi_{t_2},\xi_i)w(\xi_{t_2},\xi_{t_1})..w(\xi_{t_2},\xi_{t_k})\\&+..+\sum_{t_1,,,t_k\neq i}w(\xi_{t_k},\xi_i)w(\xi_{t_k},\xi_{t_1})..w(\xi_{t_k},\xi_{t_{k-1}})\Bigg]\end{aligned} \tag{35}$$

inside bracket is a U-statistic, thus,

$$Var(\sum_{t_1,,,t_k\neq i}w(\xi_{t_1},\xi_i)w(\xi_{t_1},\xi_{t_2})..w(\xi_{t_1},\xi_{t_k}|\xi_i)\to 0 \tag{36}$$

which means,$(n-1)[E\sum_{i=1}^n(Z_i-\bar{Z}_i)^2-Var(Z_{n-1})]\to 0$