A Consistent Histogram Estimator for Exchangeable Graph Models (Supplementary Material)

Stanley H. Chan and Edoardo M. Airoldi

1 Proof of Lemma 1

We first prove the forward direction. Suppose that $\left|\frac{\sigma(i)}{n} - \frac{\sigma(j)}{n}\right| < \delta$ for some $\delta > 0$. Then,

$$
\begin{aligned}
\mathbb{P}\left(\left|U_{\sigma(i)} - U_{\sigma(j)}\right| > 3\delta\right) &\le \mathbb{P}\left(\left|U_{\sigma(i)} - \frac{\sigma(i)}{n}\right| + \left|U_{\sigma(j)} - \frac{\sigma(j)}{n}\right| + \left|\frac{\sigma(i)}{n} - \frac{\sigma(j)}{n}\right| > 3\delta\right) \\
&\le \mathbb{P}\left(\left|U_{\sigma(i)} - \frac{\sigma(i)}{n}\right| + \left|U_{\sigma(j)} - \frac{\sigma(j)}{n}\right| > 2\delta\right) \\
&\le \mathbb{P}\left(\left|U_{\sigma(i)} - \frac{\sigma(i)}{n}\right| > \delta\right) + \mathbb{P}\left(\left|U_{\sigma(j)} - \frac{\sigma(j)}{n}\right| > \delta\right) \\
&\overset{(a)}{\le} 2\exp\{-2n\delta^2\} + 2\exp\{-2n\delta^2\} \\
&= 4\exp\{-2n\delta^2\},
\end{aligned}
$$

where (a) is due to Dvoretzky. Consequently,

$$
\begin{aligned}
\mathbb{P}\left(\left|g(U_{\sigma(i)}) - g(U_{\sigma(j)})\right| > 3L_1\delta\right) &\overset{(b)}{\le} \mathbb{P}\left(\left|U_{\sigma(i)} - U_{\sigma(j)}\right| > 3\delta\right) \\
&\le 4\exp\{-2n\delta^2\},
\end{aligned}
$$

where (b) is due to Lipschitz. Therefore,

$$
\begin{aligned}
&\mathbb{P}\left(\left|d_{\sigma(i)} - d_{\sigma(j)}\right| > 6L_1\delta \;\middle|\; U_{\sigma(i)}, U_{\sigma(j)}\right) \\
&\le \mathbb{P}\left(\left|d_{\sigma(i)} - g(U_{\sigma(i)})\right| + \left|d_{\sigma(j)} - g(U_{\sigma(j)})\right| + \left|g(U_{\sigma(i)}) - g(U_{\sigma(j)})\right| > 6L_1\delta \;\middle|\; U_{\sigma(i)}, U_{\sigma(j)}\right) \\
&\overset{(c)}{\le}_p \mathbb{P}\left(\left|d_{\sigma(i)} - g(U_{\sigma(i)})\right| + \left|d_{\sigma(j)} - g(U_{\sigma(j)})\right| > 3L_1\delta \;\middle|\; U_{\sigma(i)}, U_{\sigma(j)}\right) \\
&\le 2\mathbb{P}\left(\left|d_{\sigma(i)} - g(U_{\sigma(i)})\right| > \frac{3}{2}L_1\delta \;\middle|\; U_{\sigma(i)}, U_{\sigma(j)}\right) \\
&\overset{(d)}{\le} 4\exp\left\{-2n^2\left(\frac{3}{2}L_1\delta\right)^2\right\} \\
&= 4\exp\left\{-\frac{9}{2}n^2L_1^2\delta^2\right\}.
\end{aligned}
$$

Here, (d) is due to Hoeffding. The inequality in (c) holds with probability at least $1-4\exp\{-2n\delta^2\}$. Letting two events

$$
\begin{aligned}
\mathcal{E}_1 &= \left\{\left|d_{\sigma(i)} - d_{\sigma(j)}\right| > 6L_1\delta \;\middle|\; U_{\sigma(i)}, U_{\sigma(j)}\right\} \\
\mathcal{E}_2 &= \left\{\left|g(U_{\sigma(i)}) - g(U_{\sigma(j)})\right| < 3L_1\delta\right\},
\end{aligned}
$$

and using the fact that

$$\begin{aligned}\mathbb{P}\left(\mathcal{E}_1\right) &= \mathbb{P}\left(\mathcal{E}_1 \cap \mathcal{E}_2\right) + \mathbb{P}\left(\mathcal{E}_1 \cap \mathcal{E}_2^c\right) \\ &\leq \mathbb{P}\left(\mathcal{E}_1 \cap \mathcal{E}_2\right) + \mathbb{P}\left(\mathcal{E}_2^c\right),\end{aligned}$$

then we have

$$\begin{aligned}\mathbb{P}\left(\left|d_{\sigma(i)} - d_{\sigma(j)}\right| > 6L_1\delta \;\middle|\; U_{\sigma(i)}, U_{\sigma(j)}\right) &\leq 4\exp\left\{-\frac{9}{2}n^2L_1^2\delta^2\right\} + 4\exp\{-2n\delta^2\} \\ &\leq 8\exp\{-2n\delta^2\},\end{aligned}$$

when $n > \frac{4}{9L_1^2}$. Putting $\delta = \frac{1}{6L_1}\sqrt{\frac{\log n}{n}}$, we have

$$\mathbb{P}\left(\left|d_{\sigma(i)} - d_{\sigma(j)}\right| > \sqrt{\frac{\log n}{n}} \;\middle|\; U_{\sigma(i)}, U_{\sigma(j)}\right) \leq 8e^{-\frac{1}{18L_1^2}\log n}.$$

We next prove the converse. First, by inverse Lipschitz we have

$$\begin{aligned}\left|\frac{\sigma(i)}{n} - \frac{\sigma(j)}{n}\right| &\leq \left|\frac{\sigma(i)}{n} - U_{\sigma(i)}\right| + \left|\frac{\sigma(j)}{n} - U_{\sigma(j)}\right| + \left|U_{\sigma(i)} - U_{\sigma(j)}\right| \\ &\leq \left|\frac{\sigma(i)}{n} - U_{\sigma(i)}\right| + \left|\frac{\sigma(j)}{n} - U_{\sigma(j)}\right| + \frac{1}{L_2}\left|g(U_{\sigma(i)}) - g(U_{\sigma(j)})\right|. \end{aligned} \tag{1}$$

By Dvoretzky, we have $\mathbb{P}\left(\left|\frac{\sigma(i)}{n} - U_{\sigma(i)}\right| > \eta\right) \leq 2\exp\{-2n\eta^2\}$ for any $\eta > 0$. Putting $\eta = \sqrt{\frac{1}{2n}\log\left(\frac{2}{\alpha}\right)}$, then $\mathbb{P}\left(\left|\frac{\sigma(i)}{n} - U_{\sigma(i)}\right| > \sqrt{\frac{1}{2n}\log\left(\frac{2}{\alpha}\right)}\right) \leq \alpha$. That is,

$$\left|\frac{\sigma(i)}{n} - U_{\sigma(i)}\right| \leq \sqrt{\frac{1}{2n}\log\left(\frac{2}{\alpha}\right)} \tag{2}$$

with probability at least $1-\alpha$.

Next, we note that

$$\left|g(U_{\sigma(i)}) - g(U_{\sigma(j)})\right| \leq \left|g(U_{\sigma(i)}) - d_{\sigma(i)}\right| + \left|g(U_{\sigma(j)}) - d_{\sigma(j)}\right| + \left|d_{\sigma(i)} - d_{\sigma(j)}\right|.$$

By Hoeffding, we know $\mathbb{P}\left(\left|g(U_{\sigma(i)}) - d_{\sigma(i)}\right| > \delta\right) \leq 2\exp\{-2n^2\delta\}$ for any $\delta > 0$. Putting $\delta = \sqrt{\frac{1}{2n^2}\log\left(\frac{2}{\alpha}\right)}$, then

$$\left|g(U_{\sigma(i)}) - d_{\sigma(i)}\right| \leq \sqrt{\frac{1}{2n^2}\log\left(\frac{2}{\alpha}\right)}, \tag{3}$$

with probability at least $1-\alpha$.

Substituting (2), (3) and that $\left|d_{\sigma(i)} - d_{\sigma(j)}\right| < \epsilon$ with probability at least $1-\alpha$ into (1), we have

$$\left|\frac{\sigma(i)}{n} - \frac{\sigma(j)}{n}\right| \leq 2\sqrt{\frac{1}{2n}\log\left(\frac{2}{\alpha}\right)} + \frac{2}{L_2}\sqrt{\frac{1}{2n^2}\log\left(\frac{2}{\alpha}\right)} + \frac{\epsilon}{L_2}, \tag{4}$$

which holds with probability at least $(1-\alpha)^5$.

Putting $\alpha = 8e^{-\frac{1}{18L_1^2}\log n}$, $\epsilon = \sqrt{\frac{\log n}{n}}$, and using the fact that

$$\log\left(\frac{2}{\alpha}\right) = \log\left(\frac{1}{4}\right) + \frac{\log n}{18L_1^2} \leq \frac{\log n}{18L_1^2},$$

we have

$$\left|\frac{\sigma(i)}{n} - \frac{\sigma(j)}{n}\right| \leq \sqrt{\frac{\log n}{n}}\left(\frac{1}{3L_1} + \frac{1}{3L_1L_2\sqrt{n}} + \frac{1}{L_2}\right),$$

with probability at least $(1-8e^{-\frac{1}{18L_1^2}\log n})^5 \approx 1-40e^{-\frac{1}{18L_1^2}\log n}$ for large n.

2 Proof of Lemma 2

For clarity and notational simplicity we prove a continuous version of the lemma. First, we define $w^{step} : [0,1]^2 \to [0,1]$ as the continuous version of $H^w \otimes \mathbf{1}_{h\times h}$. That is, we equally partition $[0,1]$ into k sub-intervals with width h/n. Then, for any (x,y) in the (i,j)th sub-interval $[i(h/n),\, (i+1)(h/n)] \times [j(h/n),\, (j+1)(h/n)]$, we let $w^{step}(x,y) = H^w_{i,j}$.

By assumption that w is smooth, there exists $\zeta_i \in \left[\frac{i-1}{k}, \frac{i}{k}\right]$ and $\xi_j \in \left[\frac{j-1}{k}, \frac{j}{k}\right]$ such that

$$w^{step}(u,v) = w(\zeta_i, \xi_j),$$

for $u \in \left[\frac{i-1}{k}, \frac{i}{k}\right]$, and $v \in \left[\frac{j-1}{k}, \frac{j}{k}\right]$. Therefore, the approximation error is bounded as

$$\begin{aligned}
\|w - w^{step}\|_2^2 &= \sum_{i=1}^{k}\sum_{j=1}^{k}\int_{\frac{i-1}{k}}^{\frac{i}{k}}\int_{\frac{j-1}{k}}^{\frac{j}{k}} \left(w(u,v) - w^{step}(u,v)\right)^2 dvdu \\
&= \sum_{i=1}^{k}\sum_{j=1}^{k}\int_{\frac{i-1}{k}}^{\frac{i}{k}}\int_{\frac{j-1}{k}}^{\frac{j}{k}} \left(w(u,v) - w^{step}(\zeta_i,\xi_j)\right)^2 dvdu \\
&\leq \left(\frac{1}{k^2}\right)^2 \sum_{i=1}^{k}\sum_{j=1}^{k}\int_{\frac{i-1}{k}}^{\frac{i}{k}}\int_{\frac{j-1}{k}}^{\frac{j}{k}} \sup_{\substack{u\,\in\,[\frac{i-1}{k},\,\frac{i}{k}] \\ v\,\in\,[\frac{j-1}{k},\,\frac{j}{k}]}} |\nabla w(u,v)|^2 \, dvdu \\
&\leq \frac{1}{k^2}\sup_{u\in[0,1],v\in[0,1]} |\nabla w(u,v)|^2 .
\end{aligned}$$

Therefore,

$$\|w - w^{step}\|_2 \leq \frac{1}{k}\sup_{u,v\in[0,1]} |\nabla w(u,v)| . \tag{5}$$

3 Proof of Lemma 3

First, by definition of $\widehat{H}$ and H, we have

$$\mathbb{E}[\|\widehat{H} - H\|_2^2] = \mathbb{E}\left[\sum_{i=1}^{k}\sum_{j=1}^{k}\left(\frac{1}{h^2}\sum_{i_1=1}^{h}\sum_{j_1=1}^{h}\left(\widehat{A}_{ih+i_1,jh+j_1} - A_{ih+i_1,jh+j_1}\right)\right)^2\right]. \tag{6}$$

To evaluate (6), it is clear that we have to estimate

$$\mathbb{E}\left[(\widehat{A}_{ij} - A_{ij})^2\right] \quad \text{and} \quad \mathbb{E}\left[\widehat{A}_{ij} - A_{ij}\right]$$

for all $i, j = 1, \ldots, k$. Let w_{ij} be the true graphon and

$$\widehat{w}_{ij} = w(U_{\widehat{\sigma}(i)}, U_{\widehat{\sigma}(j)})$$

be the empirical graphon ordered by $\widehat{\sigma}(1), \ldots, \widehat{\sigma}(n)$. Then it holds that

$$\mathbb{E}\left[(\widehat{A}_{ij} - A_{ij})^2\right] = \mathbb{E}[(\widehat{A}_{ij} - \widehat{w}_{ij})^2 + (A_{ij} - w_{ij})^2 + (w_{ij} - \widehat{w}_{ij})^2], \tag{7}$$

because $\mathbb{E}[\widehat{A}_{ij}] = \widehat{w}_{ij}$ and $\mathbb{E}[A_{ij}] = w_{ij}$.

To bound (7), we first show that

$$\mathbb{E}[(\widehat{A}_{ij} - \widehat{w}_{ij})^2] = \mathrm{Var}[\widehat{A}_{ij}] \le 1, \tag{8}$$

$$\mathbb{E}[(A_{ij} - w_{ij})^2] = \mathrm{Var}[A_{ij}] \le 1. \tag{9}$$

Next, we bound the term $(w_{ij} - \widehat{w}_{ij})^2$ as

$$\begin{aligned}(w_{ij} - \widehat{w}_{ij})^2 &\overset{(a)}{=} \left[w(U_{\sigma(i)}, U_{\sigma(j)}) - w(U_{\widehat{\sigma}(i)}, U_{\widehat{\sigma}(j)})\right]^2 \\ &\overset{(b)}{\le} \left[L\left(|U_{\sigma(i)} - U_{\widehat{\sigma}(i)}| + |U_{\sigma(j)} - U_{\widehat{\sigma}(j)}|\right)\right]^2 \\ &\overset{(c)}{\le} 4C^2L^2\frac{\log n}{n}, \end{aligned} \tag{10}$$

where $C = \frac{1}{3L_1} + \frac{1}{3L_1L_2} + \frac{1}{L_2}$. Here, in (a) we write $w_{ij} = w(U_{\sigma(i)}, U_{\sigma(j)})$. Since w is the true graphon, the permutation σ is the identity operator: $\sigma(i) = i$ for all i. The inequality in (b) holds because of the Lipschitz condition on w. The inequality in (c) is due to (4). Substituting (8), (9) and (10) into (7) yields

$$\mathbb{E}\left[(\widehat{A}_{ij} - A_{ij})^2\right] \le 2 + 4C^2L^2\frac{\log n}{n}. \tag{11}$$

Similarly, $\mathbb{E}\left[\widehat{A}_{ij} - A_{ij}\right]$ can be bounded as

$$\begin{aligned}\mathbb{E}\left[\widehat{A}_{ij} - A_{ij}\right] &\le \mathbb{E}\left[\widehat{A}_{ij} - \widehat{w}_{ij}\right] + \mathbb{E}\left[w_{ij} - A_{ij}\right] + |\widehat{w}_{ij} - w_{ij}| \\ &\le 2CL\sqrt{\frac{\log n}{n}}. \end{aligned} \tag{12}$$

Going back to (6), we can show that

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{i=1}^{k}\sum_{j=1}^{k}\left(\frac{1}{h^2}\sum_{i_1=1}^{h}\sum_{j_1=1}^{h}\left(\widehat{A}_{ih+i_1,jh+j_1}-A_{ih+i_1,jh+j_1}\right)\right)^2\right] \\
&\leq \sum_{i=1}^{k}\sum_{j=1}^{k}\frac{1}{h^4}\left(h^2\left(2+4C^2L^2\frac{\log n}{n}\right)+\frac{h^2(h^2-1)}{2}\left(2CL\sqrt{\frac{\log n}{n}}\right)^2\right) \\
&\leq \frac{k^2}{h^2}\left(2+4C^2L^2\frac{\log n}{n}\right)+k^2\left(2C^2L^2\frac{\log n}{n}\right).
\end{aligned}
$$

Substituting $n = kh$, we have

$$
\mathbb{E}[\|\widehat{H}-H\|_2^2] \leq \frac{k^4}{n^2}\left(2+4C^2L^2\frac{\log n}{n}\right)+k^2\left(4C^2L^2\frac{\log n}{n}\right). \tag{13}
$$

4 Proof of Lemma 4

By definitions of H_{ij} and H_{ij}^w, it holds that

$$
\begin{aligned}
\mathbb{E}[H_{ij}] = \mathbb{E}\left[\frac{1}{h^2}\sum_{i_1=1}^{h}\sum_{j_1=1}^{h}A_{ih+i_1,jh+j_1}\right] &= \frac{1}{h^2}\sum_{i_1=1}^{h}\sum_{j_1=1}^{h}\mathbb{E}\left[A_{ih+i_1,jh+j_1}\right] \\
&= \frac{1}{h^2}\sum_{i_1=1}^{h}\sum_{j_1=1}^{h}w_{ih+i_1,jh+j1} = H_{ij}^w.
\end{aligned}
$$

Consequently, we can show that

$$
\mathbb{E}\left[(H_{ij}-H_{ij}^w)^2\right] = \mathbb{E}\left[(H_{ij})^2\right]-(H_{ij}^w)^2,
$$

and hence

$$
\begin{aligned}
\mathbb{E}\left[H_{ij}^2\right] &= \frac{1}{h^4}\left(\sum_{i_1=1}^{h}\sum_{j_1=1}^{h}\sum_{i_2\neq i_1}\sum_{j_2\neq j_1}\mathbb{E}\left[A_{ih+i_1,jh+j_1}A_{ih+i2,jh+j_2}\right]+\sum_{i_1=1}^{h}\sum_{j1=1}^{h}\mathbb{E}\left[A_{ih+i_1,jh+j_1}^2\right]\right) \\
&= \frac{1}{h^4}\left(\sum_{i_1=1}^{h}\sum_{j_1=1}^{h}\sum_{i_2\neq i_1}\sum_{j_2\neq j_1}w_{ih+i_1,jh+j_1}w_{ih+i2,jh+j_2}+\sum_{i_1=1}^{h}\sum_{j1=1}^{h}w_{ih+i_1,jh+j_1}\right) \\
&= (H_{ij}^w)^2+\frac{1}{h^4}\sum_{i_1=1}^{h}\sum_{j_1=1}^{h}w_{ih+i_1,jh+j_1}(1-w_{ih+i_1,jh+j_1}) \\
&\leq (H_{ij}^w)^2+\frac{1}{h^2}.
\end{aligned}
$$

Therefore,

$$
\mathbb{E}[\|H-H^w\|_2^2] = \sum_{i=1}^{k}\sum_{j=1}^{k}\mathbb{E}[(H_{ij}-H_{ij}^w)^2] \leq \frac{k^2}{h^2} = \frac{k^4}{n^2}.
$$

5 Proof of Theorem 3

By the definition of MSE, we have

$$
\begin{aligned}
\text{MSE} &\stackrel{\text{def}}{=} \frac{1}{n^2}\mathbb{E}[\|\widehat{w}^{est} - w\|_2^2] \\
&= \frac{1}{n^2}\Big(\mathbb{E}[h^2\|\widehat{w}^{tv} - H^w\|_2^2] + \mathbb{E}[\|H^w \otimes \mathbf{1}_{h\times h} - w\|_2^2] + 2\mathbb{E}[(\widehat{w}^{tv} - H^w)^T(H^w \otimes \mathbf{1}_{h\times h} - w)]\Big),
\end{aligned} \tag{14}
$$

The first term above can be bounded by Lemma 5:

$$
\|\widehat{w}^{tv} - H^w\|_2^2 \le \varepsilon^2,
$$

because by assumption $\|\nabla H^w - (\nabla H^w)_s\|_1 = 0$. Now, ε can further be bounded by Lemma 3 and Lemma 4:

$$
\begin{aligned}
\varepsilon^2 &\stackrel{\text{def}}{=} \mathbb{E}[\|\eta + \rho\|_2^2] \\
&\stackrel{(a)}{=} \mathbb{E}[\|\widehat{H} - H\|_2^2] + \mathbb{E}[\|H - H^w\|_{\ell_2}^2], \\
&\le \frac{k^4}{n^2}\left(2 + 4C^2L^2\frac{\log n}{n}\right) + k^2\left(4C^2L^2\frac{\log n}{n}\right) + \frac{k^4}{n^2},
\end{aligned} \tag{15}
$$

where in (a) we used the fact that $\mathbb{E}[H_{ij}] = H_{ij}^w$ so that $\mathbb{E}[\rho] = 0$. Therefore,

$$
\begin{aligned}
&\frac{1}{n^2}\mathbb{E}\left[h^2\|\widehat{w}^{tv} - H^w\|_2^2\right] \\
&= \frac{k^2}{n^2}\left(2 + 4C^2L^2\frac{\log n}{n}\right) + \left(4C^2L^2\frac{\log n}{n}\right) + \frac{k^2}{n^2} \to 0
\end{aligned}
$$

as $n \to \infty$ and $k/n \to 0$.

The second term in (14) can be bounded by Lemma 2, which gives

$$
\frac{1}{n^2}\|H^w \otimes \mathbf{1}_{h\times h} - w\|_2^2 \le \frac{C'}{k^2n^2} \to 0 \tag{16}
$$

as $n \to \infty$, where $C' = \sup|\nabla w|^2$.

Substituting (15) and (16) into (14) completes the proof.