7 SUPPLEMENT TO "ESTIMATING GRAPHLETS VIA LIFTING"

7.1 Proof of Prop. 3.1.

Proof. Let ϕ_i be as defined in (12), (13). For both estimators, because of the form of (5) and (6), if a single term ϕ_i is unbiased then $\hat{N}_m$ is as well. Let us begin with $\hat{N}_{O,m}$, by considering a draw from the lifting process, $A = [v_1, \ldots, v_k]$ which induces the k-subgraph, $G|A$. By the definition of $\tilde{\pi}$,

$$\mathbb{E}\left(\phi_{O,1}\right) = \sum_{A \in V_G^k} \tilde{\pi}(A) \left(\frac{\mathbb{1}(T(A) \sim H_m)}{\mathrm{co}(T(A))\tilde{\pi}(A)} \right)$$

$$= \sum_{T \in \mathcal{V}_k} \sum_{A \in V_G^k : T(A) = T} \frac{\mathbb{1}(T \sim H_m)}{\mathrm{co}(T)} = \sum_{T \in \mathcal{V}_k} \mathbb{1}(T \sim H_m) = N_m.$$

Hence, the $\hat{N}_{O,m}$ is unbiased. Consider the shotgun estimator, $\hat{N}_{S,m}$,

$$\mathbb{E}\left(\phi_{S,1}\right) = \sum_{B \in V_G^{k-1}} \tilde{\pi}(B) \sum_{u \in \mathcal{N}_v(B)} \left(\frac{\mathbb{1}(G|B \cup \{u\} \sim H_m)}{\mathrm{co}(H_m)\tilde{\pi}(B)} \right)$$

$$= \sum_{T \in \mathcal{V}_k} \sum_{B \in V_G^{k-1}} \mathbb{1}(G|B \cup \{u\} = T, u \in \mathcal{N}_v(B)) \frac{\mathbb{1}(T \sim H_m)}{\mathrm{co}(T)}$$

$$= \sum_{T \in \mathcal{V}_k} \mathbb{1}(T \sim H_m) = N_m.$$

Hence, the shotgun estimator is unbiased as well. □

7.2 Proof of Theorem 4.1.

We can bound the variance in (11) by the second moment, which is bounded by,

$$\mathbb{E}\phi_1^2 \leq \mathbb{E}\phi_1 \max \phi_1 = N_m(G) \max \phi_1.$$

Seeking to control the the maximum of ϕ_1, we see that,

$$\max_T \frac{1}{\pi_U(T)} \leq \max_A \frac{1}{|\mathrm{co}(T)|\tilde{\pi}(A)} \leq \max \frac{\prod_{r=1}^{k-1}(d_1 + \ldots + d_r)}{|\mathrm{co}(H_m)|\pi_1(d_1)},$$

$$\max_B \frac{|\mathcal{N}_v(B)|}{|\mathrm{co}(H_m)|\tilde{\pi}(B)} \leq \max \frac{\prod_{r=1}^{k-1}(d_1 + \ldots + d_r)}{|\mathrm{co}(H_m)|\pi_1(d_1)}.$$

Thus, we can construct a bound on $V_m^{\perp\!\!\perp}(\phi_1)$.

7.3 Proof of Theorem 4.4

Let ϕ_i be as defined in (12), (13) or (14). Given two starting vertices v_i and v_j of the lifting process, notice that random variables $\phi_i|v_i$ and $\phi_j|v_j$ are independent. Therefore

$$\mathbb{E}\left(\phi_i \phi_{i+1})\right) = \mathbb{E}_{\pi_1(v_i)\times\pi_1(v_{i+1})} \mathbb{E}\left(\phi_i \phi_{i+1} | v_i, v_{i+1}\right) =$$
$$\mathbb{E}_{\pi_1(v_i)\times\pi_1(v_{i+1})} \left(\mathbb{E}\left(\phi_i|v_i\right) \mathbb{E}\left(\phi_{i+1}|v_{i+1}\right)\right).$$

Using the equation above, we can bound the covariance of ϕ_i and ϕ_{i+1} with basic inequalities:

$$|\mathrm{Cov}\left(\phi_i, \phi_{i+1}\right)| \leq$$
$$\sum_{x_1, x_2 \in V_G} \mathbb{E}(\phi_i | v_i = x_1) \mathbb{E}\left(\phi_{i+1} | v_{i+1} = x_2\right)$$
$$|\mathbb{P}(v_i = x_1, v_{i+1} = x_2) - \pi_1(x_1)\pi_1(x_2)| \leq$$
$$\max_{x_2 \in V_G} \mathbb{E}(\phi_{i+1}|v_{i+1} = x_2) \sum_{x_1} \mathbb{E}\left(\phi_i | v_i = x_1\right)$$
$$\max_{x_1} \sum_{x_2} |\mathbb{P}(v_i = x_1, v_{i+1} = x_2) - \pi(x_1)\pi(x_2)| =$$
$$2\gamma_{G_V}(h) \max_{x_2} \mathbb{E}\left(\phi_{i+1}|v_{i+1} = x_2\right) \sum_{x_1} \mathbb{E}\left(\phi_i | v_i = x_1\right),$$

where $\gamma_{G_V}(h)$ is the mixing coefficient from (18) for the random walk on vertices. Next, estimate factors from the RHS as follows:

$$\sum_x \mathbb{E}\left(\phi | v = x\right) \leq \max_x \frac{1}{\pi(x)} \sum_x \mathbb{E}\left(\phi | v = x\right) \pi(x) \leq$$
$$2|E_G| N_m(G). \quad (21)$$

For $\max_x \mathbb{E}\left(\phi | v = x\right)$, consider the expressions for ϕ from (12), (13) or (14).

Using notation $D = \prod_{r=2}^{k-1}(\Delta_1 + \ldots + \Delta_r)$, for the Ordered Lift estimator,

$$\max_x \mathbb{E}\left(\phi_O | v = x\right) \leq \max_x \sum_A \frac{\mathbb{P}(A|v = x)}{\tilde{\pi}(A)} \leq$$
$$\max_x \frac{|\{A \mid A[1] = x\}|}{\pi(x)} \leq 2|E_G|D.$$

For the Shotgun Lift estimator,

$$\max_x \mathbb{E}\left(\phi_S | v = x\right) \leq \max_x \sum_B |\mathcal{N}_v(B)| \frac{\mathbb{P}(B|v = x)}{\tilde{\pi}(B)} \leq$$
$$\max_x \frac{|\mathcal{N}_v(B)||\{B \mid B[1] = x\}|}{\pi(x)} \leq 2|E_G|D,$$

For the Unordered Lift estimator,

$$\max_x \mathbb{E}\left(\phi_U | v = x\right) \leq \max_x \sum_T \frac{\mathbb{P}(T|v = x)}{\pi_U(T)} \leq$$
$$\max_x \frac{|\{T \mid x \in V_T\}|}{\pi(x)} \leq 2|E_G|D.$$

Combining the results, we get the desired bound.