Proposition 1. *Suppose $\mathcal{L}$ is sampled from a Random Graph model (such as graphon) that is independent of X_0. Denote $N = \max\{n, d\}$, then the following holds: For any $\epsilon > 0$, there exist universal constant $C > 0$ and constants $C_P, c_\alpha > 0$ depending only on α and f such that*

$$\mathbb{P}\left(|h_G - n^{-1}\operatorname{Tr} f(\mathcal{L})\mathcal{L}f(\mathcal{L})| \geq \frac{C_P(\log(N))^{1+\epsilon}}{\sqrt{N}}\right) \leq C \times N^{-c_\alpha(\log(N))^\epsilon}.$$

Proof. The proof this very similar to that of Theorem. We will again apply Hanson Wright Inequality. Suppose $n \geq d$, then we write

$$\begin{aligned} h_G &= n^{-1}\operatorname{Tr}\mathcal{L}XX^\top = (nd)^{-1}\operatorname{Tr}\mathcal{L}f(\mathcal{L})X_0X_0^\top f(\mathcal{L}) = (nd)^{-1}\operatorname{Tr}X_0^\top f(\mathcal{L})\mathcal{L}f(\mathcal{L})X_0 \\ &= \frac{1}{nd}\sum_{a=1}^{d}(X_0\mathbf{e}_a)^\top f(\mathcal{L})\mathcal{L}f(\mathcal{L})X_0\mathbf{e}_a = n^{-1}\operatorname{Tr} f(\mathcal{L})\mathcal{L}f(\mathcal{L}) + \Delta(n) \end{aligned}$$

where the error $\Delta(n)$ is defined as

$$|\Delta(n)| = \left|\frac{1}{d}\sum_{a=1}^{d}\underbrace{n^{-1}[(X_0\mathbf{e}_a)^\top f(\mathcal{L})\mathcal{L}f(\mathcal{L})X_0\mathbf{e}_a - \operatorname{Tr} f(\mathcal{L})\mathcal{L}f(\mathcal{L})]}_{E_a}\right| \leq \max_{1\leq a\leq d}|E_a|.$$

Applying

$$n^{-1}\|f(\mathcal{L})\mathcal{L}f(\mathcal{L})\|_F \leq n^{-1/2}\|f(\mathcal{L})\|_{\text{op}}^2\|\mathcal{L}\|_{\text{op}} \leq C_P \times n^{-1/2}.$$

We find

$$\mathbb{P}\left(|h_G - n^{-1}\operatorname{Tr} f(\mathcal{L})\mathcal{L}f(\mathcal{L})| \geq \frac{C_P(\log(n))^{1+\epsilon}}{\sqrt{n}}\right) \leq C \times n^{-c_\alpha(\log(n))^\epsilon}.$$

Similarly, if $d > n$, since $\mathcal{L}$ is p.s.d, we can write $\mathcal{L} = \mathcal{L}^{1/2}\mathcal{L}^{1/2}$, so we have

$$\begin{aligned} h_G &= \frac{1}{nd}\sum_{a=1}^{d}(X_0\mathbf{e}_a)^\top f(\mathcal{L})\mathcal{L}^{1/2}\mathcal{L}^{1/2}f(\mathcal{L})X_0\mathbf{e}_a = \frac{1}{nd}\sum_{a=1}^{d}\sum_{i=1}^{n}(X_0\mathbf{e}_a)^\top f(\mathcal{L})\mathcal{L}^{1/2}\mathbf{e}_i\mathbf{e}_i^\top\mathcal{L}^{1/2}f(\mathcal{L})X_0\mathbf{e}_a \\ &= \frac{1}{nd}\sum_{i=1}^{n}\sum_{a=1}^{d}(\mathbf{e}_i^\top\mathcal{L}^{1/2}f(\mathcal{L})X_0\mathbf{e}_a)^2 = \frac{1}{n}\sum_{i=1}^{n}\frac{\|f(\mathcal{L})\mathcal{L}^{1/2}\mathbf{e}_i\|_2^2}{d}\sum_{a=1}^{d}(\mathbf{g}_a)^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}\frac{\|f(\mathcal{L})\mathcal{L}^{1/2}\mathbf{e}_i\|_2^2}{d} \times \mathbf{g}^\top I_d\mathbf{g} = \frac{1}{n}\sum_{i=1}^{n}\|f(\mathcal{L})\mathcal{L}^{1/2}\mathbf{e}_i\|_2^2 + \Delta(d) \\ &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{e}_i^\top\mathcal{L}^{1/2}f(\mathcal{L})f(\mathcal{L})\mathcal{L}^{1/2}\mathbf{e}_i + \Delta(d) = n^{-1}\operatorname{Tr} f(\mathcal{L})\mathcal{L}f(\mathcal{L}) + \Delta(d) \end{aligned}$$

where we defined $\mathbf{g} \sim N(0, I_d)$ and the error term $\Delta(d)$ is defined to be

$$|\Delta(d)| = \left|\frac{1}{n}\sum_{i=1}^{n}\underbrace{d^{-1}\|f(\mathcal{L})\mathcal{L}^{1/2}\mathbf{e}_i\|_2^2[\mathbf{g}^\top I_d\mathbf{g} - \operatorname{Tr} I_d]}_{E_i}\right| \leq \max_{1\leq i\leq n}|E_i|.$$

Again, following similar steps as in the proof of Theorem , we deduce

$$\mathbb{P}\left(|h_G - n^{-1}\operatorname{Tr} f(\mathcal{L})\mathcal{L}f(\mathcal{L})| \geq \frac{C_P(\log(d))^{1+\epsilon}}{\sqrt{d}}\right) \leq C \times d^{-c_\alpha(\log(d))^\epsilon}.$$

Combing the cases $n \geq d$ and $n < d$ completes the proof. $\square$