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

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

Proof. The proof uses the Hanson–Wright inequality. Throughout we condition on $\mathcal{L}$ (independent of X_0), apply concentration, and then remove the conditioning at the end. Set

$$A := f(\mathcal{L})\,\mathcal{L}\,f(\mathcal{L}),$$

which is positive semidefinite because $\mathcal{L} \succeq 0$ and $f(\mathcal{L})$ is symmetric. We use the uniform bounds

$$\|\mathcal{L}\|_{\mathrm{op}} \leq 1, \qquad \|f(\mathcal{L})\|_{\mathrm{op}} \leq C_P \quad \Longrightarrow \quad \|A\|_{\mathrm{op}} \leq C_P, \quad \|A\|_F \leq \sqrt{n}\, C_P.$$

Case 1: $n \geq d$ (so $N = n$). We write

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

where

$$\Delta(n) := \frac{1}{d}\sum_{a=1}^{d} E_a, \qquad E_a := \frac{1}{n}\Big((X_0\mathbf{e}_a)^\top A(X_0\mathbf{e}_a) - \operatorname{Tr}(A)\Big).$$

Since $X_0\mathbf{e}_a \overset{\text{i.i.d.}}{\sim} N(0, I_n)$ and is independent of $\mathcal{L}$, the Hanson–Wright inequality for $\gamma \sim N(0, I_n)$ yields, for any $t > 0$,

$$\mathbb{P}(|E_a| \geq t \,|\, \mathcal{L}) = \mathbb{P}\Big(\big|\gamma^\top A\gamma - \operatorname{Tr}(A)\big| \geq nt \,\Big|\, \mathcal{L}\Big) \leq C\exp\left(-c\,\min\left\{\frac{n^2t^2}{\|A\|_F^2},\ \frac{nt}{\|A\|_{\mathrm{op}}}\right\}\right).$$

Choose $t = \dfrac{C_P(\log N)^{1+\epsilon}}{\sqrt{N}} = \dfrac{C_P(\log n)^{1+\epsilon}}{\sqrt{n}}$. Then

$$\frac{n^2t^2}{\|A\|_F^2} = (\log n)^{2+2\epsilon}, \qquad \frac{nt}{\|A\|_{\mathrm{op}}} \geq (\log n)^{1+\epsilon},$$

so

$$\mathbb{P}\left(|E_a| \geq \frac{C_P(\log n)^{1+\epsilon}}{\sqrt{n}} \,\middle|\, \mathcal{L}\right) \leq C\exp\big(-c(\log n)^{1+\epsilon}\big).$$

Since $|\Delta(n)| \leq \max_{1\leq a\leq d}|E_a|$, by a union bound over $a = 1, \ldots, d$ (with $d \leq n = N$),

$$\mathbb{P}\left(|\Delta(n)| \geq \frac{C_P(\log n)^{1+\epsilon}}{\sqrt{n}} \,\middle|\, \mathcal{L}\right) \leq C\, d\, \exp\big(-c(\log n)^{1+\epsilon}\big) = C\, n^{-c_\alpha(\log n)^\epsilon}.$$

Unconditioning preserves the bound:

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

Case 2: $d > n$ **(so** $N = d$**).** Since $\mathcal{L} \succeq 0$, write $\mathcal{L} = \mathcal{L}^{1/2}\mathcal{L}^{1/2}$. Then

$$\begin{aligned}
h_G &= \frac{1}{nd}\sum_{a=1}^{d}(X_0\mathbf{e}_a)^\top f(\mathcal{L})\mathcal{L}^{1/2}\mathcal{L}^{1/2}f(\mathcal{L})\,X_0\mathbf{e}_a \\
&= \frac{1}{nd}\sum_{a=1}^{d}\sum_{i=1}^{n}\left(\mathbf{e}_i^\top\mathcal{L}^{1/2}f(\mathcal{L})X_0\mathbf{e}_a\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\frac{\|f(\mathcal{L})\mathcal{L}^{1/2}\mathbf{e}_i\|_2^2}{d}\sum_{a=1}^{d}\gamma_a^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\|f(\mathcal{L})\mathcal{L}^{1/2}\mathbf{e}_i\|_2^2 + \Delta(d) = \frac{1}{n}\operatorname{Tr}\big(\mathcal{L}^{1/2}f(\mathcal{L})f(\mathcal{L})\mathcal{L}^{1/2}\big) + \Delta(d) \\
&= \frac{1}{n}\operatorname{Tr}(A) + \Delta(d),
\end{aligned}$$

where $\gamma \sim N(0, I_d)$ and

$$\Delta(d) := \frac{1}{n}\sum_{i=1}^{n}E_i, \qquad E_i := \frac{\|f(\mathcal{L})\mathcal{L}^{1/2}\mathbf{e}_i\|_2^2}{d}\left(\gamma^\top I_d\gamma - \operatorname{Tr}(I_d)\right).$$

Using $\|\mathcal{L}\|_{\mathrm{op}} \le 1$ and $\|f(\mathcal{L})\|_{\mathrm{op}} \le C_P$, we have $\|f(\mathcal{L})\mathcal{L}^{1/2}\mathbf{e}_i\|_2^2 \le C_P$, hence

$$|E_i| \le \frac{C_P}{d}\left|\gamma^\top I_d\gamma - \operatorname{Tr}(I_d)\right|.$$

Apply Hanson–Wright with $A = I_d$ ($\|I_d\|_F = \sqrt{d}$, $\|I_d\|_{\mathrm{op}} = 1$) and choose

$$t = \frac{C_P(\log N)^{1+\epsilon}}{\sqrt{N}} = \frac{C_P(\log d)^{1+\epsilon}}{\sqrt{d}}.$$

Then

$$\mathbb{P}(|E_i| \ge t \mid \mathcal{L}) \le \mathbb{P}\Big(\left|\gamma^\top I_d\gamma - \operatorname{Tr}(I_d)\right| \ge (\log d)^{1+\epsilon}\sqrt{d}\Big) \le C\exp\big(-c(\log d)^{1+\epsilon}\big).$$

A union bound over $i = 1, \dots, n$ (with $n < d = N$) gives

$$\mathbb{P}\left(|\Delta(d)| \ge \frac{C_P(\log d)^{1+\epsilon}}{\sqrt{d}}\,\middle|\,\mathcal{L}\right) \le C\,n\,\exp\big(-c(\log d)^{1+\epsilon}\big) = C\,d^{-c_\alpha(\log d)^\epsilon}.$$

Unconditioning yields

$$\mathbb{P}\left(\left|h_G - \frac{1}{n}\operatorname{Tr}(A)\right| \ge \frac{C_P(\log d)^{1+\epsilon}}{\sqrt{d}}\right) \le C\,d^{-c_\alpha(\log d)^\epsilon}.$$

Conclusion. In Case 1 we have $N = n$, in Case 2 we have $N = d$. Rewriting both bounds with $N = \max\{n, d\}$ gives the stated inequality:

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

The argument used only $\|\mathcal{L}\|_{\mathrm{op}} \le 1$ and the independence of $\mathcal{L}$ and X_0, so removing conditioning is valid. This completes the proof. $\square$