Theorem 13.1. *Assume $\mathcal{F}$ is a VC-subgraph class and $VC(\mathcal{F}) = V$. Suppose $-1 \le f(x) \le 1$ for all $f \in \mathcal{F}$ and $x \in \mathcal{X}$. Let $x_1, \ldots, x_n \in \mathcal{X}$ and define $d(f,g) = \frac{1}{n}\sum_{i=1}^n |f(x_i) - g(x_i)|$. Then*

$$\mathcal{D}(\mathcal{F}, \varepsilon, d) \le \left(\frac{8e}{\varepsilon}\log\frac{7}{\varepsilon}\right)^V .$$

(which is $\le \left(\frac{K}{\varepsilon}\right)^{V+\delta}$ for some δ.)

Proof. Let $m = \mathcal{D}(\mathcal{F}, \varepsilon, d)$ and $f_1, \ldots, f_m$ be ε-separated, i.e.

$$\frac{1}{n}\sum_{i=1}^n |f_r(x_i) - f_\ell(x_i)| > \varepsilon.$$

Let $(z_1, t_1), \ldots, (z_k, t_k)$ be constructed in the following way: z_i is chosen uniformly from $x_1, \ldots, x_n$ and t_i is uniform on $[-1, 1]$.

Consider f_r and f_ℓ from the ε-packing. Let C_{f_r} and C_{f_ℓ} be subgraphs of f_r and f_ℓ. Then

$$\begin{aligned}
&\mathbb{P}\left(C_{f_r} \text{ and } C_{f_\ell} \text{ pick out different subsets of } (z_1, t_1), \ldots, (z_k, t_k)\right)\\
&= \mathbb{P}\left(\text{At least one point } (z_i, t_i) \text{ is picked by } C_{f_r} \text{ or } C_{f_\ell} \text{ but not picked by the other}\right)\\
&= 1 - \mathbb{P}\left(\text{All points } (z_i, t_i) \text{ are picked either by both or by none}\right)\\
&= 1 - \mathbb{P}\left((z_i, t_i) \text{ is picked either by both or by none}\right)^k
\end{aligned}$$

□

Since z_i is drawn uniformly from $x_1, \ldots, x_n$,

$$\begin{aligned}
&\mathbb{P}\left((z_1, t_1) \text{ is picked by both } C_{f_r}, C_{f_\ell} \text{ or by neither}\right)\\
&= \frac{1}{n}\sum_{i=1}^n \mathbb{P}\left((x_i, t_1) \text{ is picked by both } C_{f_r}, C_{f_\ell} \text{ or by neither}\right)\\
&= \frac{1}{n}\sum_{i=1}^n \left(1 - \frac{1}{2}|f_r(x_i) - f_\ell(x_i)|\right)\\
&= 1 - \frac{1}{2}\frac{1}{n}\sum_{i=1}^n |f_r(x_i) - f_\ell(x_i)|\\
&= 1 - \frac{1}{2}d(f_r, f_\ell) \le 1 - \varepsilon/2 \le e^{-\varepsilon/2}
\end{aligned}$$

Substituting,

$$\begin{aligned}
&\mathbb{P}\left(C_{f_r} \text{ and } C_{f_\ell} \text{ pick out different subsets of } (z_1, t_1), \ldots, (z_k, t_k)\right) \\
&= 1 - \mathbb{P}\left((z_1, t_1) \text{ is picked by both } C_{f_r}, C_{f_\ell} \text{ or by neither}\right)^k \\
&\geq 1 - \left(e^{-\varepsilon/2}\right)^k \\
&= 1 - e^{-k\varepsilon/2}
\end{aligned}$$

There are $\binom{m}{2}$ ways to choose f_r and f_ℓ, so

$$\mathbb{P}\left(\text{All pairs } C_{f_r} \text{ and } C_{f_\ell} \text{ pick out different subsets of } (z_1, t_1), \ldots, (z_k, t_k)\right) \geq 1 - \binom{m}{2} e^{-k\varepsilon/2}.$$

What k should we choose so that $1 - \binom{m}{2} e^{-k\varepsilon/2} > 0$? Choose

$$k > \frac{2}{\varepsilon} \log \binom{m}{2}.$$

Then there exist $(z_1, t_1), \ldots, (z_k, t_k)$ such that all C_{f_ℓ} pick out different subsets. But $\{C_f : f \in \mathcal{F}\}$ is VC, so by Sauer's lemma, we can pick out at most $\left(\frac{ek}{V}\right)^V$ out of these k points. Hence, $m \leq \left(\frac{ek}{V}\right)^V$ as long as $k > \frac{2}{\varepsilon} \log \binom{m}{2}$. The latter holds for $k = \frac{2}{\varepsilon} \log m^2$. Therefore,

$$m \leq \left(\frac{e}{V} \frac{2}{\varepsilon} \log m^2\right)^V = \left(\frac{4e}{V\varepsilon} \log m\right)^V,$$

where $m = \mathcal{D}(\mathcal{F}, \varepsilon, d)$. Hence, we get

$$m^{1/V} \leq \frac{4e}{\varepsilon} \log m^{1/V}$$

and defining $m^{1/V} = s$,

$$s \leq \frac{4e}{\varepsilon} \log s.$$

Note that $\frac{s}{\log s}$ is increasing for $s \geq e$ and so for large enough s, the inequality will be violated. We now check that the inequality is violated for $s' = \frac{8e}{\varepsilon} \log \frac{7}{\varepsilon}$. Indeed, one can show that

$$\frac{4e}{\varepsilon} \log \left(\frac{7}{\varepsilon}\right)^2 > \frac{4e}{\varepsilon} \log \left(\frac{8e}{\varepsilon} \log \frac{7}{\varepsilon}\right)$$

since

$$\frac{49}{8e\varepsilon} > \log \frac{7}{\epsilon}.$$

Hence, $m^{1/V} = s \leq s'$ and, thus,

$$\mathcal{D}(\mathcal{F}, \varepsilon, d) \leq \left(\frac{8e}{\varepsilon} \log \frac{7}{\varepsilon}\right)^V.$$