▶ Lecture 14 EN.553.744 Prof. Luana Ruiz

Today: - what if my graph grows, what if too large?
↳ graph limits: graphons

→ We understand the different ways a graph can be perturbed, and we know how to design stable GNNs.

↳ 1) but what if my graph grows?
2) or, what if it is too large and I don't have enough resources to train on it?
(recall a GNN forward pass requires $O(LK|\mathcal{E}|)$ complexity)

For 1): can we measure how close two large graphs are? b) If we can, and if the GNN is continuous wrt this metric, then we are good. d)
what does it mean for graphs to be close?
↳ convergence to a common limit a)

For 2) can we sample a subgraph? c)

↳ random graph model

We have more questions than answers! Let's take them on one at a time.

a) What does it mean for graphs to be close?

We are statisticians/probabilists; the prevailing perspective is that we consider graphs to be close if sampled subgraphs have similar distributions, or, equivalently, if they have similar subgraph counts

↳ can be measured using homomorphism densities

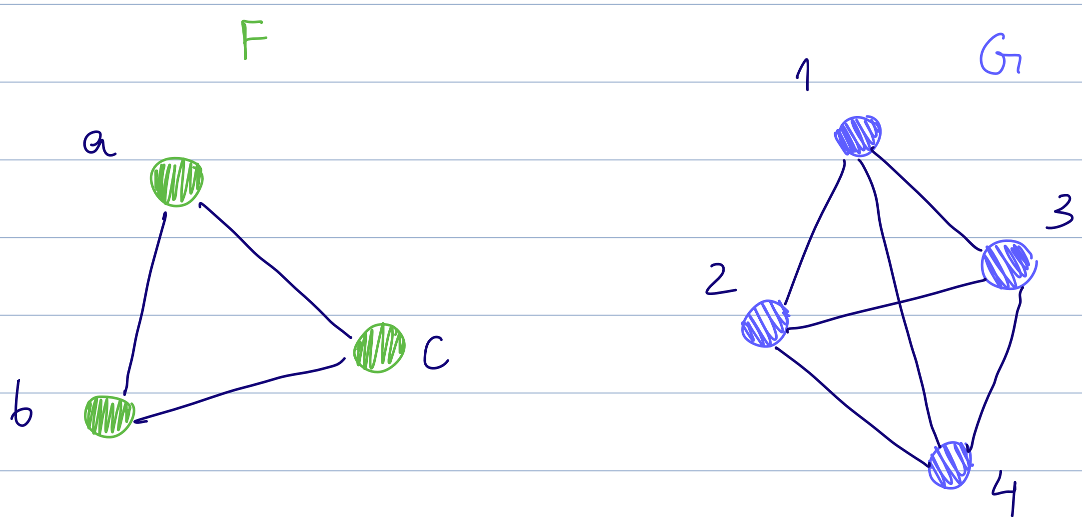
Recall from Lecture 11:

(DEF) Graph homomorphism: Let $G = (V, E)$ & $F = (V', E')$. A homomorphism from F to G is a map $\gamma : V' \to V$:

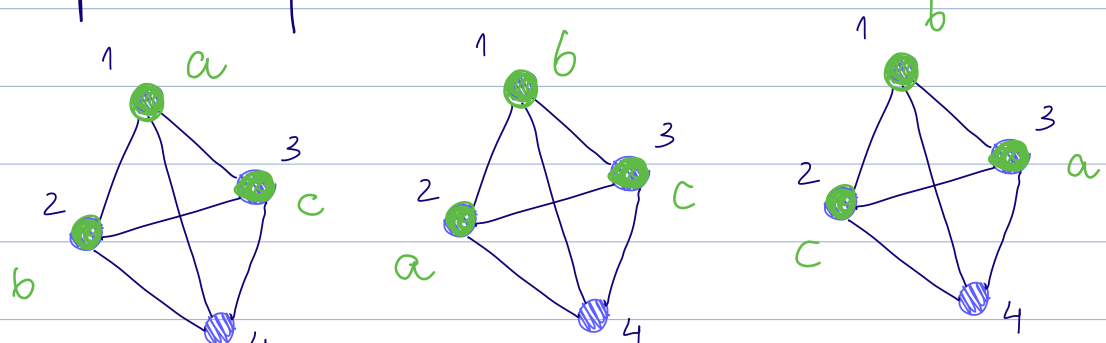
$$(i,j) \in E' \Rightarrow (\gamma(i), \gamma(j)) \in E$$

I.e., homomorphisms are adjacency preserving (of F) maps

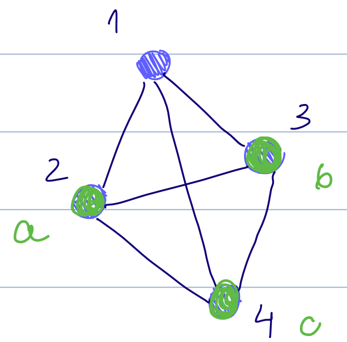
E.g.:



Homomorphisms from F to G:



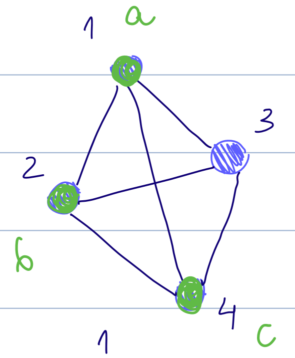
etc...



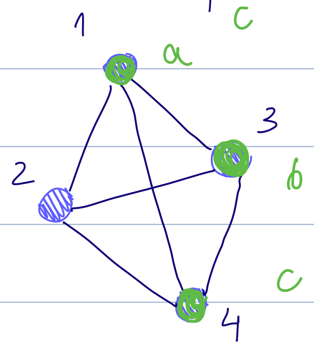
. . .

(6 in total)

(plus other permutations of a, b, c) (6 total)



=



=

We will denote the total number of homomorphisms from F to G hom(F, G). In the example, hom(F, G) = 24

(DEF) Homomorphism density: the hom. density from F to G, denoted t(F, G) is:

$$t(F, G) = \frac{\text{hom}(F, G)}{|V|^{|V'|}}$$

In the example, $t(F,G) = \frac{24}{4^3} = \frac{3}{8}$

$\Rightarrow$ For now, we will say two graphs G_1 & G_2 are close if, for all motifs F (undirected, unweighted graphs, w/ one edge per node pair & no self loops), $t(F,G_1) \approx t(F,G_2)$

- Convergent graph sequences & graphons (Lovász, Chayes, Borgs, Vesztergombi) 2008-onwards

(DEF) Let $(G_n)_n$ be a graph sequence such that $\lim_{n\to\infty} t(F,G_n)$ exists for all motifs F.

Then, this sequence is convergent and its limit is given a graphon.

(DEF) A graphon is a symmetric, bounded, measurable function $W: [0,1]^2 \to [0,1]$
$([0,1] \times [0,1])$

I.e., it is a bounded kernel

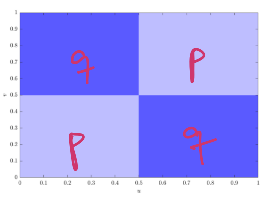
Obs. 1: A more general definition is $W: \Omega \times \Omega \to [0,1]$, where Ω is some sample space endowed with some probability measure f. But, since we can always map Ω into $[0,1]$ using a measure-preserving map (as long as CDF of f is strictly monotone $\to$ exercise to verify), we'll stick with $[0,1]$ as our node sample space.

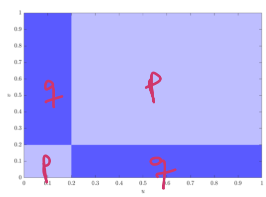
Obs. 2: The codomain of W may be $[0,B]$, $B<+\infty$, but it is typical to have $B=1$ since $W(x,y)$ often represents a probability.

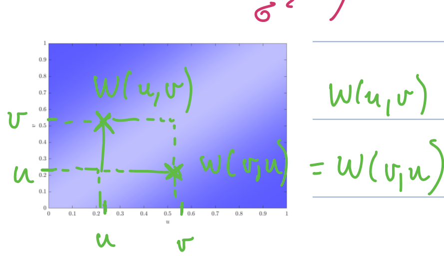
The easiest way to think of a graphon is as a graph with an uncountable number of nodes $u \in [0,1]$, and edges (u,v) with weights $W(u,v)$.

$$W(u,v) = \exp\left(\frac{(u-v)^2}{\sigma^2}\right)$$

E.g.







(DEF) Graphon homomorphism density

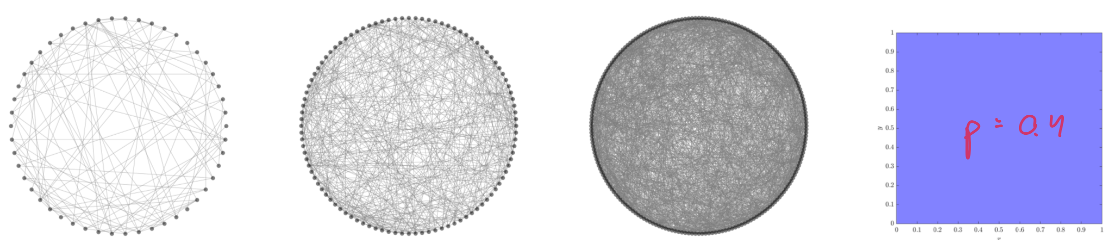
The density of homomorphisms from a graph motif $F=(V',E')$ to a graphon W is denoted $t(F,W)$ and defined as:

$$\boxed{t(F,W) = \iint_{[0,1]^{|V'|}} \prod_{(i,j)\in E'} W(u_i,u_j) \prod_{i\in V'} du_i}$$

If $W(u,v) < 1 \;\; \forall\, u, v,$ $t(F,W)$ can be interpreted as the probability of sampling the motif F from the graphon W.

$$t(F, G_n) \xrightarrow{n\to\infty} t(F,W) \;\; \forall F \iff (G_n)_n \text{ is a conv. graph sequence w/ limit } W$$

E.g.: G_{50} G_{100} G_{200} W



(a) 50 nodes (b) 100 nodes (c) 200 nodes (d) Graphon

This notion of convergence is called left convergence as it deals with left homomorphisms $t(F, G_n)$, $t(F, W)$

This is not the only way to define/identify (dense) convergent graph sequences. Another definition based on the convergence of min-cuts (micro-canonical ground state energy in physics) exists, based on right homomorphisms ($t(G_n, F)$, $t(W, F)$). For dense graphs, left & right convergence are equivalent.

b) Can we measure how close two large graphs are?

While $t(F, G)$ gives us a way to define convergence, it is difficult to use it to measure the distance between graphs, as we'd have to compute $t(F, G_1)$ & $t(F, G_2)$ for all motifs F.

▶ Cut distance for graphs

(1) Graphs G & G' with same n & same node sets V, V' (i.e., same node labeling)

We define two distances:

L_1 norm or edit distance: $d_1(G, G') = \|A - A'\|_1$

where A, A' are the (unweighted) adjacencies of G, G'

Cut distance: $d_\square = \|A - A'\|_\square$,

where $\|\cdot\|_\square$ is the cut norm:

$$\|B\|_\square = \max_{S \subseteq [n], T \subseteq [n]} \left| \sum_{s \in S, t \in T} B_{ts} \right|$$

I.e., given a (graph) matrix B, $\|B\|_\square$ is "the size" (total number of edges, or sum of edge weights) of the max cut.

(2) Graphs G & G' with the same number of nodes:

Cut distance: $\hat{\delta}_{\square}(G, G') = \min_{P \in \mathcal{P}} \| A - P^T A' P \|_{\square}$

I.e., we must look at all possible permutations of the nodes of G' (while fixing the node labels in G)

(3) Graphs G_n & G_m with $\neq$ number of nodes ($n \neq m$)

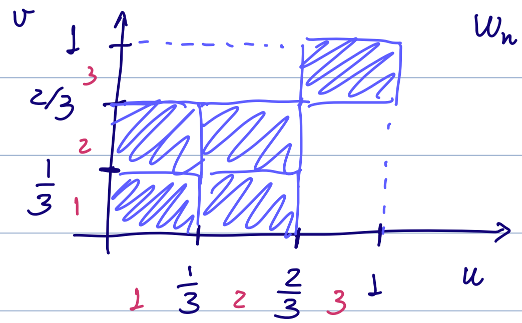
Define $W_n(u,v) = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} \mathcal{I}(u \in I_i) \mathcal{I}(v \in I_j)$

where $\mathcal{I}$ is the indicator function and I_j is defined as:

$$I_j = \begin{cases} \left[\frac{j-1}{n}, \frac{j}{n} \right[, & 1 \leq j \leq n-1 \\ \left[\frac{n-1}{n}, 1 \right] , & j = n \end{cases}$$

E.g.:

$$\begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & & & \\ 2 & & & \\ 3 & & & \end{array} \quad A_n = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



W_n is called the graphon induced by G_n.

Back to the cut distance:

Using transformation (*), we can represent both G_n and G_m as kernels on $[0,1]^2$, so they are compatible objects. But now we need a notion of cut norm for kernels.

Cut norm: Let W be a kernel in $[0,1]^2$.
Its cut norm is defind as:

$$\|W\|_\square = \sup_{S,T \subseteq [0,1]} \left| \iint_{S\times T} W(u,v)\,du\,dv \right|$$

We're not done yet! We need to take into account "node relabelings" (permutations of $I_1, I_2, \dots$).

So for two Kernels (graphons), we define the cut metric:

$$\delta_\square(W, W') = \inf_\phi \| W^\phi - W' \|_\square$$

where $W^\phi(u,v) = W(\phi(u), \phi(v))$ and ϕ are measure-preserving bijections.