intersection of the $N(x_u)$, $u \in U$ and $A_l$. Then

$$\prod_{u\in U}(p_{ul}-\varepsilon) \le |Y|m^{-1} \le \prod_{u\in U}(p_{ul}+\varepsilon) \tag{9.8}$$

Let $x_1, \ldots, x_r$ be a normal partial copy. We say it is destroyed by $x_{r+1} \in A_{r+1}$ if $x_1, \ldots, x_{r+1}$ is a partial copy but is not normal. We claim at most $2s\varepsilon m$ vertices $x_{r+1}$ can destroy $x_1, \ldots, x_r$. How can this occur? Let $l > r+1$ be adjacent to $r+1$ and let $U, Y$ be as above (looking only at $x_1, \ldots, x_r$). Then $Y \cap N(x_{r+1})$ would need to be either too big or too small. If more than $2s\varepsilon m$ vertices $x_{r+1}$ destroyed $x_1, \ldots, x_r$ then there would be a set $X \subset A_{r+1}$ of size at least $m\varepsilon$ of such $x_{r+1}$, all with the same $l$ and with either all $Y \cap N(x_{r+1})$ too big or all too small. Assume the former, the latter being similar. Then $d(X, Y) > p_{r+1,l} + \varepsilon$. But $|Y| \ge m\varepsilon$ by our choice of $\kappa$. From $\varepsilon$-regularity $|X| \le m\varepsilon$, as claimed.

The $N$ choices of $x_i \in A_i$, $1 \le i \le s$ for which $x_i, x_j$ are adjacent in $G$ whenever $i, j$ are adjacent in $H$ fall into two categories. There are at most $2s^2\varepsilon m^s$ choices such that $x_{r+1}$ destroys $x_1, \ldots, x_r$ for some $r$. The other choices are bounded in number between $m^s \prod(p_{ij} - \varepsilon)$ and $m^s \prod(p_{ij} + \varepsilon)$, the products over $i, j$ adjacent in $H$. Let $f(\varepsilon)$ denote the maximum distance between either of these products and $\prod p_{ij}$. We can then set $\gamma_2 = 2s^2\varepsilon + f(\varepsilon)$.

■

## 9.5 GRAPHONS

As in Section 9.3 we set $N_G(H)$ denote the number of labelled copies of $H$ as a (not necessarily induced) subgraph of $G$. We set $t(H, G) = N_G(H)n^{-a}$ where $H, G$ have $a, n$ vertices respectively. This may naturally be interpreted as the proportion of $H$ in $G$, $0 \le t(H, G) \le 1$ tautologically.

**Definition 5** *A sequence of graphs $G_n$ is called a limit sequence, if $\lim_{n\to\infty} t(H, G_n)$ exists for all finite graphs $H$.*

**Definition 6** *Two limit sequences $G_n, G'_n$ are called equivalent if $\lim_{n\to\infty} t(H, G_n) = \lim_{n\to\infty} t(H, G'_n)$ for all finite graphs $H$. A* graphon *is an equivalence class of limit sequences.*

A graphon is a subtle object, an abstract limit of a convergent (by Definition 5) sequences of graphs. (We call a limit sequence $G_n$ a graphon even though, technically, the graphon is the equivalence class.) It is *not* itself an infinite graph, though it may seem like one. It reflects the properties of very large graphs (formally, in a limit sense) of similar nature. The excellent book (Lovász 2012) serves as a general reference to graphons.

Surprisingly, and integral to the strength of this concept, there is a good characterization of graphons. Let $W : [0, 1] \times [0, 1] \to [0, 1]$ be a Lebesgue measurable function with $W(x, y) = W(y, x)$ for all $x, y \in [0, 1]$. For each positive integer $n$ we define a random graph, denoted $G(n, W)$ on vertex set $1, \ldots, n$ as follows:

1. Select $x_1, \ldots, x_n \in [0, 1]$ uniformly and independently.
2. For $i \neq j$ let $\{i, j\}$ be an edge of $G(n, W)$ with probability $W(x_i, x_j)$, and let the events that $\{i, j\}$ are edges be mutually independent.

As an important example, when $W$ is the constant function $W(x, y) = p$, $G(n, W)$ is simply $G(n, p)$. We call $W$ *checkered* if it splits into constant valued rectangles. More precisely, let $K \geq 1$, let $a_i \geq 0$ for $1 \leq i \leq K$ with $\sum_{i=1}^{K} a_i = 1$. Decompose $[0, 1]$ into intervals $I_i$, $1 \leq i \leq K$, of length $a_i$. Then define $W(x, y) = W(y, x) = p_{ij}$ for $x \in I_i, y \in I_j, x \leq y$. For such $W$, $G(n, W)$ is basically a random multipartite graph with the vertex set split into sets $V_i$ of size $\sim na_i$ and all $\{x, y\}$ with $x \in V_i$, $y \in V_j$ being adjacent with independent probability $p_{ij}$.

Let $H$ be a graph on $1, \ldots, s$. Set

$$c(H, W) = \int \prod_{\{i,j\} \in H} W(x_i, x_j) \tag{9.9}$$

where the integral is over $x_1, \ldots, x_s \in [0, 1]$ and the null product is interpreted as one. We leave as an exercise that with probability one the sequence $G(n, W)$ is a limit sequence with

$$\lim_{n \to \infty} t(H, G(n, W)) = c(H, W) \tag{9.10}$$

We say that a graphon $G_n$ is represented by $W$ if

$$\lim_{n \to \infty} t(H, G_n) = c(H, W) \tag{9.11}$$

for every finite $H$. Observe, from Property $P_1^p(s)$ of Theorem 9.3.2, that a sequence $G_n$ is a graphon represented by the constant function $W(x, y) = p$ if and only if $G_n$ is a quasirandom graph sequence with parameter $p$, as given by Definition 4.

**Theorem 9.5.1** *Every graphon is represented by some $W$.*

The proof of Theorem 9.5.1 requires some techniques slightly beyond the scope of this chapter. Rather, we prove the following weaker version.

**Theorem 9.5.2** *Let $\kappa > 0$ and positive integer $L$ be given. Let $G_n$ be an arbitrary graphon. Then there exists a checkered $W$ such that*

$$|\lim_{n \to \infty} t(H, G_n) - c(H, W)| \leq \kappa \tag{9.12}$$

*for all $H$ with $s \leq L$ vertices.*

**Proof.** Let $\varepsilon$ be a small positive real and $t$ a large positive integer, as described more fully below. For each $G_n$ apply the Regularity Lemma (Theorem 9.4.1) to give an $\varepsilon$-regular partition $(V_0, V_1, \ldots, V_k)$ with $t \leq k \leq T = T(\varepsilon, t)$. Take a subsequence on which $k$ is a constant. Further take a subsequence on which the set of $\{i, j\}$ for which $(V_i, V_j)$ is an $\varepsilon$-regular pair is the same. Further take a subsequence such that $d(V_i, V_j)$ approaches a limit $p_{ij}$ for all $0 \le i, j \le t$ and such that the proportion of vertices in $V_i$ approaches a limit $\alpha_i$ for all $0 \le i \le t$. Now define a checkered $W$ by splitting $[0,1]$ into intervals $I_i$ of length $\alpha_i$, $0 \le i \le t$ and letting $W$ take the constant value $p_{ij}$ on $I_i \times I_j$. Let $H$ be a graph on vertex set $1, \ldots, s$ with $s \le L$. We compare $\lim t(H, G_n)$ and $c(H, W)$. Let $\psi : V(H) \to V(G_n)$ with $\psi(i) \in V_{x_i}$. We say $x_1, \ldots, x_s \in \{0, 1, \ldots, t\}$ is normal (else, abnormal) if the $x_i$ are distinct and nonzero and all $(V_{x_i}, V_{x_j})$ are $\varepsilon$-regular. The proportion of abnormal $\psi$ is then at most $s\varepsilon$ (some $\psi(i) \in V_0$) plus $\binom{s}{2}\varepsilon$ (some $(V_{x_i}, V_{x_j})$ are not $\varepsilon$-regular) which is at most $L^2\varepsilon$. We can make this arbitrarily small by adjusting $\varepsilon$. Now suppose $x_1, \ldots, x_s$ is normal. Let $N(x_1, \ldots, x_s)$ denote the number of choices of $v_i \in V_{x_i}$, $1 \le i \le s$, such that $v_i, v_j$ are adjacent in $G$ whenever $i, j$ are adjacent in $H$. From Theorem 9.4.4 $N(x_1, \ldots, x_s)m^{-L}$ differs from $\prod p_{ij}$ (product over adjacent $i, j$) by at most $\gamma$, where $m$ is the size of each $V_{x_i}$. Summing over all normal $x_1, \ldots, x_s$ the contribution to $Nn^{-L}$ differs from the contribution to $c(H, W)$ by at most $\gamma$. From (9.7) we may make $\gamma$ arbitrarily small by choosing appropriately small $\varepsilon$. The total difference between $t(H, G_n)$ and $c(H, W)$ is then at most $L^2\varepsilon + \gamma$. For any given positive $\kappa$ we may find $\varepsilon$ so that this is less than $\kappa$. ■

Among the applications of graphons is the replacement of asymptotic questions on graphs $G_n$ with analytic questions on functions $W$. We content ourself with a typical example.

Let $b$ be the minimal real number so that there exist $G_n$ with $0.7\binom{n}{2} + o(n^2)$ edges and $b\binom{n}{3} + o(n^3)$ triangles. Let $b'$ be the minimum of $\int W(x,y)W(x,z)W(y,z)$ $(x, y, z \in [0,1])$ given that $\int W(x,y) = 0.7$. (We leave as an exercise that the minima $b, b'$ are attained.) Both tough questions. We will show that they are the *same* question, that $b = b'$.

The easy part is $b \le b'$. Let $W$ be such that $c(K_2, W) = 0.7$ and $c(K_3, W) = b'$. Then the random sequence $G_n \sim G(n, W)$ has, with probability one, $0.7\binom{n}{2} + o(n^2)$ edges and $b'\binom{n}{3} + o(n^3)$ triangles. Hence the minimal possible $b$ has $b \le b'$.

For the opposite direction we first give a natural topological result:

**Theorem 9.5.3** *Any sequence $G_n$ contains a subsequence which is a limit sequence in the sense of Definition 5.*

**Proof.** Place all finite graphs into a countable list $H_1, H_2, \ldots$ and set $t_i(G) = t(H_i, G)$. Let $SEQ_0$ denote the original sequence $G_n$. As all $t_1(G) \in [0,1]$ we find a subsequence $SEQ_1$ on which $t_1(G)$ converges. Given $SEQ_{i-1}$ we find a subsequence of it, denoted $SEQ_i$, on which $t_i(G)$ converges. Employ diagonalization, letting $SEQ_\omega$ be that sequence whose $i$-th term is the $i$-th term of $SEQ_i$. For each $i$, $SEQ_\omega$ is a subsequence of $SEQ_i$ except for possibly the first $i-1$ terms and hence $t_i(G)$ converges. ■

Now let $G_n$ be any sequence with $0.7\binom{n}{2} + o(n^2)$ edges and $b\binom{n}{3} + o(n^3)$ triangles. Apply Theorem 9.5.3 to find a limit sequence with the same property. Now apply Theorem 9.5.1 to find $W$ representing that limit sequence. That $W$ has $c(K_2, W) = 0.7$ and $c(K_3, W) = b$. Thus the minimal possible $b'$ has $b' \le b$.

## 9.6 EXERCISES

1. By considering a random bipartite three-regular graph on $2n$ vertices obtained by picking three random permutations between the two color classes, prove that there is a $c > 0$ such that for every $n$ there exists a $(2n, 3, c)$-expander.

2. Let $G = (V, E)$ be an $(n, d, \lambda)$-graph, suppose $n$ is divisible by $k$, and let $C : V \to \{1, 2, \ldots, k\}$ be a coloring of $V$ by $k$ colors, so that each color appears precisely $n/k$ times. Prove that there is a vertex of $G$ which has a neighbor of each of the $k$ colors, provided $k\lambda \leq d$.

3. Let $G = (V, E)$ be a graph in which there is at least one edge between any two disjoint sets of size $a+1$. Prove that for every set $Y$ of $5a$ vertices, there is a set $X$ of at most $a$ vertices, such that for every set $Z$ satisfying $Z \cap (X \cup Y) = \emptyset$ and $|Z| \leq a$, the inequality $|N(Z) \cap Y| \geq 2|Z|$ holds.

4. Prove that for every $\epsilon > 0$ there exists an $n_0 = n_0(\epsilon)$ so that for every $(n, n/2, 2\sqrt{n})$- graph $G = (V, E)$ with $n > n_0$, the number of triangles $M$ in $G$ satisfies $|M - n^3/48| \leq \epsilon n^3$.

5. Let $\varepsilon > 0$, $p \in (0, 1)$, $\lambda > 0$. Let $G \sim G(n, p)$, with vertex set $V_n$. Show that the following property has limiting probability one as $n \to \infty$: $(A_n, B_n)$ is $\varepsilon$-regular for *all* disjoint $A_n, B_n \subset V_n$ with $|A| \geq n\lambda$ and $|B| \geq n\lambda$.

6. Combine Turán's Theorem with the Regularity Lemma to prove the following result, due to Erdős, Simonovits and Stone: For every fixed graph $H$ of chromatic number $r > 1$ and every $\varepsilon > 0$, there is an $n_0 = n_0(H, \varepsilon)$ so that if $n > n_0$ then any simple graph with $n$ vertices and at least $(1 - \frac{1}{r-1} + \varepsilon)\binom{n}{2}$ edges contains a copy of $H$.

7. Let $t'(H, G)$ denote the number of induced copies of $H$ in $G$, i.e., the number of vertex subsets $S$ such that $G|_S$ is isomorphic to $H$. Show that $G_n$ is a limit sequence if and only if $\lim_{n\to\infty} t'(H, G_n)$ exists for all finite graphs $H$.

8. Prove Theorem 9.3.2.

9. Prove that the minima $b, b'$ given in Section 9.5 are actually attained.

10. Let $G = G_n$ be a sequence of bipartitie graphs with designated parts $T_n, B_n$ each of size $n$. Let $p \in (0, 1)$ and assume $\lim_{n\to\infty} d(T_n, B_n) = p$. Call such a sequence bipartite quasirandom with parameter $p$ if for all $\varepsilon > 0$ the pair $(T_n, B_n)$ is $\varepsilon$-regular for $n$ sufficiently large. State and prove a result analogous to Theorem 9.3.2, giving equivalent notions for bipartite quasirandomness.

11. Prove that for all $H, W$ and $\varepsilon > 0$ there exists $\alpha > 0$ such that
$$\Pr[|t(H, G(n, W)) - c(H, W)| > \varepsilon] < 2e^{-n\alpha} \tag{9.13}$$
Deduce that with probability one the sequence $G(n, W)$ is a limit sequence satisfying (9.10)

# *THE PROBABILISTIC LENS:*
# *Random Walks*

A *vertex-transitive* graph is a graph $G = (V, E)$ such that for any two vertices $u, v \in V$ there is an automorphism of $G$ that maps $u$ into $v$. A *random walk* of length $l$ in $G$ starting at a vertex $v$ is a randomly chosen sequence $v = v_0, v_1, \ldots, v_l$, where each $v_{i+1}$ is chosen, randomly and independently, among the neighbors of $v_i$ $(0 \leq i < l)$.

The following theorem states that for every vertex-transitive graph $G$, the probability that a random walk of even length in $G$ ends at its starting point is at least as big as the probability that it ends at any other vertex. Note that the proof requires almost no computation. We note also that the result does not hold for general regular graphs, and the vertex transitivity assumption is necessary.

**Theorem 1** *Let $G = (V, E)$ be a vertex-transitive graph. For an integer $k$ and for two (not necessarily distinct) vertices $u, v$ of $G$, let $P^k(u, v)$ denote the probability that a random walk of length $k$ starting at $u$ ends at $v$. Then, for every integer $k$ and for every two vertices $u, v \in V$,*

$$P^{2k}(u, u) \geq P^{2k}(u, v)\,.$$

**Proof.** We need the following simple inequality, sometimes attributed to Chebyshev.

**Claim 9.6.1** *For every sequence $(a_1, \ldots, a_n)$ of $n$ reals and for any permutation $\pi$ of $\{1, \ldots, n\}$,*

$$\sum_{i=1}^{n} a_i a_{\pi(i)} \leq \sum_{i=1}^{n} a_i^2\,.$$

**Proof.** The inequality follows immediately from the fact that

$$\sum_{i=1}^{n} a_i^2 - \sum_{i=1}^{n} a_i a_{\pi(i)} = \frac{1}{2} \sum_{i=1}^{n} (a_i - a_{\pi(i)})^2 \geq 0\,.$$

■

Consider, now, a random walk of length $2k$ starting at $u$. By summing over all the possibilities of the vertex the walk reaches after $k$ steps we conclude that for every vertex $v$,

$$P^{2k}(u,v) = \sum_{w \in V} P^k(u,w) P^k(w,v) = \sum_{w \in V} P^k(u,w) P^k(v,w)\,, \tag{1}$$

where the last equality follows from the fact that $G$ is an undirected regular graph.

Since $G$ is vertex-transitive, the two vectors $(P^k(u,w))_{w \in V}$ and $(P^k(v,w))_{w \in V}$ can be obtained from each other by permuting the coordinates. Therefore, by the claim above, the maximum possible value of the sum in the right-hand side of (1) is when $u = v$, completing the proof of the theorem. ■

*Part II*

# *TOPICS*