

Chapter 1

Some important models

In this chapter, we describe a few essential discrete probability models. We also formulate various questions about these models and give references to later sections where these questions are partly answered. The list of models described here, which is far from exhaustive, contains mostly models defined on graphs.

1.1 Preliminaries

We start with a brief review of graph theory and Markov chains.

1.1.1 Review of graph theory

Basic definitions An *undirected graph* (or *graph* for short) is a pair $G = (V, E)$ where V is the set of *vertices* (or *nodes* or *sites*) and

$$E \subseteq \{\{u, v\} : u, v \in V\},$$

is the set of *edges* (or *bonds*). See Figure 1.1 for an example. We occasionally write $V(G)$ and $E(G)$ for the vertices and edges of the graph G . The set of vertices V is either finite or countably infinite. Edges of the form $\{u\}$ are called *self-loops*. In general, we do *not* allow E to be a multiset. When E is a multiset, G is called a *multigraph*.

A vertex $v \in V$ is *incident* with an edge $e \in E$ if $v \in e$. The incident vertices of an edge are sometimes called *endvertices*. Two vertices $u, v \in V$ are *adjacent* (or *neighbors*), denoted by $u \sim v$, if $\{u, v\} \in E$. The set of adjacent vertices of v ,

Version: September 9, 2015
Modern Discrete Probability: An Essential Toolkit
Copyright © 2015 Sébastien Roch

*undirected
graph, vertex*

edge, self-loop

*multigraph
incident vertex,
endvertex,
adjacent vertex,
neighborhood,
degree,
isolated vertex,
regular graph,
locally finite
graph*

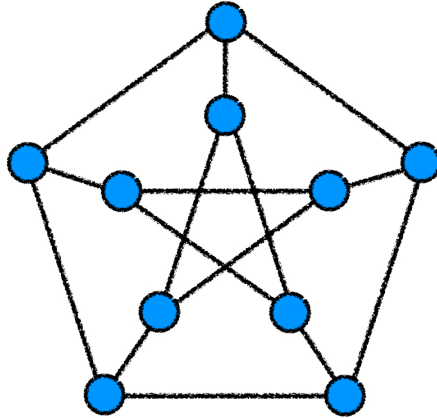


Figure 1.1: Petersen graph.

denoted by $N(v)$, is called the *neighborhood* of v and its size, i.e. $\delta(v) := |N(v)|$, is the *degree* of v . A vertex v with $\delta(v) = 0$ is called *isolated*. A graph is called *d-regular* if all its degrees are d . A countable graph is *locally finite* if all its vertices have a finite degree.

Example 1.1 (Petersen graph). All vertices in the Petersen graph in Figure 1.1 have degree 3, i.e., it is 3-regular. In particular there is no isolated vertex. ◀

A convenient way of specifying a graph is the following matrix representation. Assume the graph $G = (V, E)$ has $n = |V|$ vertices. The *adjacency matrix* A of G is the $n \times n$ matrix defined as

adjacency matrix

$$A_{xy} = \begin{cases} 1 & \text{if } \{x, y\} \in E \\ 0 & \text{o.w.} \end{cases}$$

Example 1.2 (Triangle). The adjacency matrix of a *triangle*, i.e. 3 vertices with all possible non-loop edges, is

triangle

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

◀

Paths, cycles, and trees A *subgraph* of $G = (V, E)$ is a graph $G' = (V', E')$

subgraph,

induced

subgraph

with $V' \subseteq V$ and $E' \subseteq E$. Implied in this definition is the fact that the edges in E' are incident only to V' . The subgraph G' is said to be *induced* if

$$E' = \{\{x, y\} : x, y \in V', \{x, y\} \in E\},$$

i.e., it contains all edges of G between the vertices of V' . In that case the notation $G' := G[V']$ is used. A subgraph is said to be *spanning* if $V' = V$. A subgraph containing all possible non-loop edges between its vertices is called a *complete subgraph* or *clique*.

*spanning
subgraph,
clique*

Example 1.3 (Petersen graph (continued)). The Petersen graph contains no triangle, induced or not. ◀

A *path* in G is a sequence of (not necessarily distinct) vertices $x_0 \sim x_1 \sim \dots \sim x_k$.^{*} The number of edges, k , is called the *length* of the path. If the endvertices x_0, x_k coincide, i.e. $x_0 = x_k$, we call the path a *cycle*. If the vertices are all distinct (except possibly for the endvertices), we say that the path (or cycle) is *self-avoiding*. A self-avoiding path or cycle can be seen as a (not necessarily induced) subgraph of G . We write $u \leftrightarrow v$ if there is a path between u and v . Clearly \leftrightarrow is an equivalence relation. The equivalence classes are called *connected components*. The length of the shortest self-avoiding path connecting two distinct vertices u, v is called the *graph distance* between u and v , denoted by $\rho(u, v)$. It can be checked that the graph distance is a metric.

*path,
path length,
cycle,
self-avoiding
path,
connected
component,
graph distance*

A graph is *connected* if any two vertices are linked by a path, i.e., if $u \leftrightarrow v$ for all $u, v \in V$. Or put differently, if there is only one connected component.

connected graph

Example 1.4 (Petersen graph (continued)). The Petersen graph is connected. ◀

A *forest* is a graph with no self-avoiding cycle. A *tree* is a connected forest. Vertices of degree 1 are called *leaves*. A spanning tree of G is a subgraph which is a tree and is also spanning. It is straightforward to check that a tree on n vertices has $n - 1$ edges. (Proof: induction.) A tree is called *rooted* if it has a single distinguished vertex called the *root*.

*forest, tree,
leaf

rooted tree,
root*

Some standard graphs Here are a few examples of finite graphs.

- *Complete graph K_n* : This graph is made of n vertices with all non-loop edges.
- *Cycle C_n* : The vertex set is $\{0, 1, \dots, n - 1\}$ and two vertices $i \neq j$ are adjacent if and only if $|i - j| = 1$ or $n - 1$.

complete graph

cycle

- *Hypercube* \mathbb{Z}_2^n : The vertex set is $\{0, 1\}^n$ and two vertices $\mathbf{x} \neq \mathbf{y}$ are adjacent if and only if $\|\mathbf{x} - \mathbf{y}\|_1 = 1$. *hypercube*
- *Rooted b-ary tree* $\widehat{\mathbb{T}}_b^\ell$: This graph is a tree with ℓ levels. The unique vertex on level 0 is called the root. For $j = 1, \dots, \ell - 1$, level j has b^j vertices, each of which has exactly one neighbor on level $j - 1$ (its parent) and b neighbors on level $j + 1$ (its children). The b^ℓ vertices on level ℓ are leaves. *b-ary tree*

Here are a few examples of infinite graphs.

- *Infinite d-regular tree* \mathbb{T}_d : This is an infinite tree where each vertex has exactly d neighbors. *infinite degree-d tree*
- *Lattice* \mathbb{L}^d : The vertex set is \mathbb{Z}^d and two vertices $\mathbf{x} \neq \mathbf{y}$ are adjacent if and only if $\|\mathbf{x} - \mathbf{y}\|_1 = 1$. *lattice*

We also consider a few important classes of graphs.

A *bipartite graph* $G = (L, R, E)$ is a graph whose vertex set is composed of the union of two sets $L \cup R$ and whose edge set E is a subset of $\{(\ell, r) : \ell \in L, r \in R\}$. That is, there are no edges between vertices in L and likewise for R . *bipartite graph*

Example 1.5 (Some bipartite graphs). The cycle C_{2n} is a bipartite graph. So is the *complete bipartite graph* $K_{n,m}$ with vertex set $\{\ell_1, \dots, \ell_n\} \cup \{r_1, \dots, r_m\}$ and edge set $\{(\ell_i, r_j) : i \in [n], j \in [m]\}$. *complete bipartite graph*

In a bipartite graph $G = (L, R, E)$, a *perfect matching* is a collection of edges $M \subseteq E$ such that each vertex in $L \cup R$ is incident to exactly one edge M . *perfect matching*

An *automorphism* of a graph $G = (V, E)$ is a bijection ϕ of V to itself that preserves the edges, i.e., such that $\{x, y\} \in E$ if and only if $\{\phi(x), \phi(y)\} \in E$. A graph $G = (V, E)$ is *vertex-transitive* if for any $u, v \in V$ there is an automorphism mapping u to v . *graph automorphism, vertex-transitive graph*

Example 1.6 (Petersen graph (continued)). For any $\ell \in \mathbb{Z}$, a $(2\pi\ell/5)$ -rotation of the planar representation of the Petersen graph in Figure 1.1 corresponds to an automorphism. ◀

Example 1.7 (Trees). The graph \mathbb{T}_d is vertex-transitive. The graph $\widehat{\mathbb{T}}_b^\ell$ has many automorphisms, but it is not vertex-transitive. ◀

Flows Let $G = (V, E)$ be a connected graph with two distinguished, distinct vertex sets, a *source-set* (or *source* for short) $A \subseteq V$ and a *sink-set* (or *sink* for *source, sink,*

*Usually in graph theory a path is called a “walk” but that term has a different meaning in probability theory. *flow, network*

short) Z . Let $c : E \rightarrow \mathbb{R}_+$ be a *capacity function*. A *flow* on the *network* (G, c) from source A to sink Z is a function $f : V \times V \rightarrow \mathbb{R}$ such that:

F1 (*Antisymmetry*) $f(x, y) = -f(y, x), \forall x, y \in V$.

F2 (*Capacity constraint*) $|f(x, y)| \leq c(e), \forall e = \{x, y\} \in E$, and $f(x, y) = 0$ otherwise.

F3 (*Flow-conservation constraint*)

$$\sum_{y:y \sim x} f(x, y) = 0, \quad \forall x \in V \setminus (A \cup Z).$$

For $U, W \subseteq V$ and $F \subseteq E$, let $f(U, W) := \sum_{u \in U, w \in W} f(u, w)$ and $c(F) := \sum_{e \in F} c(e)$. The *strength* of f is $|f| := f(A, A^c)$. Let $F \subseteq E$. We call F a *cutset* separating A and Z if all paths connecting A and Z include an edge in F . Let A_F be the set of vertices not separated from A by F , and similarly for Z_F . *flow strength, cutset*

Lemma 1.8 (Max flow \leq min cut). *For any cutset F separating A from Z , $|f| \leq c(F)$.*

Proof. Note that

$$f(A, A^c) \stackrel{(F3)}{=} f(A, A^c) + \sum_{u \in A_F \setminus A} f(u, V) \stackrel{(F1)}{=} f(A_F, A_F^c) \stackrel{(F2)}{\leq} c(F). \quad (1.1)$$

■

Remarkably, this bound is tight.

Theorem 1.9 (Max-flow min-cut theorem). *Let (G, c) be a finite connected network with source A and sink Z . Then the following holds*

$$\max\{|f| : \text{flow } f \text{ on } (G, c)\} = \min\{c(F) : \text{cutset } F \text{ separating } A \text{ and } Z\}.$$

Proof. Note that, by compactness, the supremum on the l.h.s. is achieved. Let f be an optimal flow. The idea of the proof is to find a matching cutset. An *augmentable path* is a self-avoiding path $x_0 \sim \dots \sim x_k$ with $x_0 \in A$, $x_i \notin A \cup Z$ for all $i \neq 0$ or k , and $f(x_{i-1}, x_i) < c(\{x_{i-1}, x_i\})$ for all i . By maximality of f there cannot be such a path with $x_k \in Z$, otherwise we could push more flow through that path and increase the strength of f —a contradiction. Let $B \subseteq V \setminus (A \cup Z)$ be the set of all final vertices in some augmentable path and let F be the edge set between B and B^c . Note that, by definition of B , all vertices in B can be reached from A without crossing F and that $f(x, y) = c(e)$ for all $e = \{x, y\} \in F$ with $x \in B$ and $y \in B^c$. Furthermore F is a cutset separating A from Z : trivially $A \subseteq B$; $Z \subseteq B^c$ by maximality of f ; and any path from A to Z must exit B and enter B^c through an edge in F . Thus we have equality in (1.1) with $A_F = B$. *augmentable path*

■

Colorings and independent sets A *coloring* of a graph $G = (V, E)$ is an assignment of colors to each vertex in G . In a coloring, two vertices may share the same color. A coloring is *proper* if for every edge e in G the endvertices of e have distinct colors. The smallest number of colors in a proper coloring of a graph G is called the *chromatic number* $\chi(G)$ of G .

coloring,
proper coloring

An *independent set* of G is a subset of vertices $W \subseteq V$ such that all pairs of vertices in W are non-adjacent.

chromatic number
independent set

Directed graphs A *directed graph* (or *digraph* for short) is a pair $G = (V, E)$ where V is a set of *vertices* (or *nodes* or *sites*) and $E \subseteq V^2$ is a set of *directed edges*. A *directed path* is a sequence of vertices x_0, \dots, x_k with $(x_{i-1}, x_i) \in E$ for all $i = 1, \dots, k$. We write $u \rightarrow v$ if there is such a path with $x_0 = u$ and $x_k = v$. We say that $u, v \in V$ *communicate*, denoted by $u \leftrightarrow v$, if $u \rightarrow v$ and $v \rightarrow u$. The \leftrightarrow relation is clearly an equivalence relation. The equivalence classes of \leftrightarrow are called the (*strongly*) *connected components* of G .

directed graph

directed edge,
directed path,
communicating
vertices,
strongly
connected
component

1.1.2 Review of Markov chain theory

Construction Formally, a Markov chain is a stochastic process satisfying the Markov property: given the present, the future is independent of the past. On a countable state space, such a process is characterized by its initial distribution and transition matrix.

For our purposes, it will suffice to “define” a Markov chain through a particular construction. Let V be a finite or countable space. Recall that a *stochastic matrix* on V is a nonnegative matrix $P = (P(i, j))_{i, j \in V}$ satisfying

stochastic matrix

$$\sum_{j \in V} P(i, j) = 1, \quad \forall i \in V.$$

Let μ be a probability measure on V and let P be a stochastic matrix on V . One way to construct a *Markov chain* (X_t) on V with transition matrix P and initial distribution μ is the following:

Markov chain

- Let $X_0 \sim \mu$ and let $(Y(i, n))_{i \in V, n \geq 1}$ be a mutually independent array of random variables with $Y(i, n) \sim P(i, \cdot)$.
- Set inductively $X_n := Y(X_{n-1}, n)$, $n \geq 1$.

So in particular:

$$\mathbb{P}[X_0 = x_0, \dots, X_t = x_t] = \mu(x_0)P(x_0, x_1) \cdots P(x_{t-1}, x_t).$$

We use the notation $\mathbb{P}_x, \mathbb{E}_x$ for the probability distribution and expectation under the chain started at x . Similarly for $\mathbb{P}_\mu, \mathbb{E}_\mu$ where μ is a probability measure.

Example 1.10 (Simple random walk on a graph). Let $G = (V, E)$ be a finite or infinite, locally finite graph. *Simple random walk* on G is the Markov chain on V , started at an arbitrary vertex, which at each time picks a uniformly chosen neighbor of the current state. ◀

Markov property Let (X_t) be a Markov chain with transition matrix P and initial distribution μ . Let $\mathcal{F}_t = \sigma(X_0, \dots, X_t)$. As mentioned above, a fundamental property of Markov chains known as the Markov property is that, given the present, the future is independent of the past. In its simplest form, that can be interpreted as $\mathbb{P}[X_{t+1} = y | \mathcal{F}_t] = \mathbb{P}_{X_t}[X_{t+1} = y] = P(X_t, y)$. More generally (see e.g. [Dur10, Theorem 6.3.1]):

Theorem 1.11 (Markov property). *Let $f : V^\infty \rightarrow \mathbb{R}$ be bounded, measurable and let $F(x) := \mathbb{E}_x[f((X_t)_{t \geq 0})]$, then* *Markov property*

$$\mathbb{E}[f((X_{s+t})_{t \geq 0}) | \mathcal{F}_s] = F(X_s) \quad a.s.$$

Remark 1.12. *We will come back to the strong Markov property in Chapter 3.*

Let (X_t) be a Markov chain with transition matrix P . We define $P^t(x, y) := \mathbb{P}_x[X_t = y]$. An important consequence of the Markov property is the following.

Theorem 1.13 (Chapman-Kolmogorov).

$$P^t(x, z) = \sum_{y \in V} P^s(x, y) P^{t-s}(y, z), \quad s \in \{0, 1, \dots, t\}.$$

Proof. This follows from the Markov property. Note that $\mathbb{P}_x[X_t = z | \mathcal{F}_s] = F(X_s)$ with $F(y) := \mathbb{P}_y[X_{t-s} = z]$ and take \mathbb{E}_x on each side. ■

Example 1.14 (Simple random walk on \mathbb{Z}). Let (X_t) be simple random walk on \mathbb{Z} where $i \sim j$ if $|i - j| = 1$. Then $P(0, x) = 1/2$ if $|x| = 1$. And $P^2(0, x) = 1/4$ if $|x| = 2$ and $P^2(0, 0) = 1/2$. ◀

If we write μ_s for the law of X_s as a row vector, then

$$\mu_s = \mu_0 P^s$$

where here P^s is the matrix product of P by itself s times.

Some more definitions The *transition graph* of a chain is the directed graph on V whose edges are the transitions with nonzero probabilities. A chain is *irreducible* if V is the unique connected component of its transition graph, i.e., if all pairs of states communicate. *transition graph, irreducibility*

Example 1.15 (Simple random walk on a graph (continued)). Simple random walk on G is irreducible if and only if G is connected. ◀

A chain is said to be *aperiodic* if for all $x \in V$

$$\gcd\{t : P^t(x, x) > 0\} = 1.$$

aperiodicity

Example 1.16 (Lazy walk on a graph). A *lazy, simple random walk* on G is a Markov chain such that, at each time, it stays put with probability $1/2$ or chooses a uniformly random neighbor of the current state otherwise. Such a walk is aperiodic. ◀

Let (X_t) be a Markov chain with transition matrix P . A *stationary measure* π is a measure such that *stationary measure*

$$\sum_{x \in V} \pi(x)P(x, y) = \pi(y), \quad \forall y \in V,$$

or in matrix form $\pi = \pi P$. We say that π is a *stationary distribution* if in addition π is a probability measure. *stationary distribution*

Example 1.17 (Simple random walk on \mathbb{L}^d). The measure $\pi \equiv 1$ is stationary for simple random walk on \mathbb{L}^d . ◀

Finite, irreducible always have a unique stationary distribution (see e.g. [LPW06, Corollary 1.17]).

Theorem 1.18 (Existence and uniqueness: finite case). *If P is irreducible and has a finite state space, then it has a unique stationary distribution.*

A transition matrix P is *reversible* w.r.t. a measure η if

$$\eta(x)P(x, y) = \eta(y)P(y, x)$$

reversible chain

for all $x, y \in V$. By summing over y , such a measure is necessarily stationary. By induction, if (X_t) is reversible w.r.t. a stationary distribution π

$$\begin{aligned} \mathbb{P}_\pi[X_0 = x_0, \dots, X_t = x_t] &= \pi(x_0)P(x_0, x_1) \cdots P(x_{t-1}, x_t) \\ &= P(x_1, x_0) \cdots P(x_t, x_{t-1})\pi(x_t) \\ &= \mathbb{P}_\pi[X_0 = x_t, \dots, X_t = x_0], \end{aligned}$$

which explains the name.

Example 1.19 (Simple random walk on \mathbb{L}^d (continued)). The measure $\eta \equiv 1$ is reversible for simple random walk on \mathbb{L}^d . ◀

Example 1.20 (Simple random walk on a graph (continued)). Let (X_t) be simple random walk on a connected graph $G = (V, E)$. Then (X_t) is reversible w.r.t. $\eta(v) := \delta(v)$. Indeed, for all $\{u, v\} \in E$,

$$\delta(u)P(u, v) = \delta(u)\frac{1}{\delta(u)} = 1 = \delta(v)\frac{1}{\delta(v)} = \delta(v)P(v, u).$$

◀

Example 1.21 (Metropolis chain). The Metropolis algorithm modifies a given irreducible, symmetric (i.e., the transition matrix is symmetric) chain Q to produce a new chain P with the same transition graph and a prescribed positive stationary distribution π . The definition is of the new chain is:

$$P(x, y) := \begin{cases} Q(x, y) \left[\frac{\pi(y)}{\pi(x)} \wedge 1 \right], & \text{if } x \neq y, \\ 1 - \sum_{z \neq x} Q(x, z) \left[\frac{\pi(z)}{\pi(x)} \wedge 1 \right], & \text{o.w..} \end{cases}$$

Note that $P(x, z) \leq Q(x, z)$ for all $z \neq x$ so

$$\sum_{z \neq x} P(x, z) \leq 1,$$

and hence P is well-defined as a transition matrix. We claim further that P is reversible w.r.t. π . Suppose $x \neq y$ and $\pi(x) \geq \pi(y)$. Then, by the definition of P , we have

$$\begin{aligned} \pi(x)P(x, y) &= \pi(x)Q(x, y)\frac{\pi(y)}{\pi(x)} \\ &= Q(x, y)\pi(y) \\ &= Q(y, x)\pi(y) \\ &= P(y, x)\pi(y), \end{aligned}$$

where we used the symmetry of Q . ◀

Convergence and mixing time A key property of Markov chains is that, under suitable assumptions, they converge to a stationary regime (see e.g. [Dur10, Theorem 6.6.4]).

Theorem 1.22 (Convergence to stationarity). *Suppose P is irreducible, aperiodic and has stationary distribution π . Then, for all x, y ,*

$$P^t(x, y) \rightarrow \pi(y),$$

as $t \rightarrow +\infty$.

We will be interested in quantifying the speed of this convergence. For probability measures μ, ν on V , let their *total variation distance* be

$$\|\mu - \nu\|_{\text{TV}} := \sup_{A \subseteq V} |\mu(A) - \nu(A)|.$$

*total variation
distance*

Lemma 1.23 (Alternative definition of total variation distance). *It holds that*

$$\|\mu - \nu\|_{\text{TV}} = \frac{1}{2} \sum_{x \in V} |\mu(x) - \nu(x)|.$$

Proof. Let $B := \{x : \mu(x) \geq \nu(x)\}$. Then, for any $A \subseteq V$,

$$\mu(A) - \nu(A) \leq \mu(A \cap B) - \nu(A \cap B) \leq \mu(B) - \nu(B),$$

and similarly $\nu(A) - \mu(A) \leq \nu(B^c) - \mu(B^c)$. The two bounds are equal so $|\mu(A) - \nu(A)| \leq \mu(B) - \nu(B)$, which is achieved at $A = B$. Also

$$\mu(B) - \nu(B) = \frac{1}{2} [\mu(B) - \nu(B) + \nu(B^c) - \mu(B^c)] = \frac{1}{2} \sum_{x \in V} |\mu(x) - \nu(x)|.$$

■

To quantify convergence to stationarity, we define

$$d(t) := \sup_{x \in V} \|P^t(x, \cdot) - \pi(\cdot)\|_{\text{TV}}.$$

Lemma 1.24. *$d(t)$ is non-increasing in t .*

Proof. Note that

$$\begin{aligned} d(t+1) &= \sup_{x \in V} \sup_{A \subseteq V} |P^{t+1}(x, A) - \pi(A)| \\ &= \sup_{x \in V} \sup_{A \subseteq V} \left| \sum_z P(x, z) (P^t(z, A) - \pi(A)) \right| \\ &\leq \sup_{x \in V} \sum_z P(x, z) \sup_{A \subseteq V} |P^t(z, A) - \pi(A)| \\ &\leq \sup_{z \in V} \sup_{A \subseteq V} |P^t(z, A) - \pi(A)| \\ &= d(t), \end{aligned}$$

where on the second line we used that P is a stochastic matrix. ■

The following concept will play a key role here.

Definition 1.25 (Mixing time). *For a fixed $\varepsilon > 0$, the mixing time is defined as*

mixing time

$$t_{\text{mix}}(\varepsilon) := \inf\{t \geq 0 : d(t) \leq \varepsilon\}.$$

1.2 Percolation

There are several types of percolation processes. We mostly focus on bond percolation.

Definition 1.26 (Bond percolation). *Let $G = (V, E)$ be a finite or infinite graph. The bond percolation process on G with density $p \in [0, 1]$, whose measure is denoted by \mathbb{P}_p , is defined as follows: each edge of G is independently set to open with probability p , otherwise it is set to closed. Write $x \Leftrightarrow y$ if $x, y \in V$ are connected by a path all of whose edges are open. The open cluster of x is*

bond percolation

open cluster

$$\mathcal{C}_x := \{y \in V : x \Leftrightarrow y\}.$$

We will mostly consider bond percolation on \mathbb{L}^d or \mathbb{T}_d .

Typical questions regarding bond percolation (and pointers to some answers) include:

- For which values of p is there an infinite open cluster? How many infinite clusters are there? What is the probability that y is in the open cluster of x ?
 - Lattice: Sections 2.1.3, 4.3.6 6.2.5.
 - Trees: Sections 2.2.3, 3.1.4, 5.3.1.

1.3 Random graphs

There are many different models of random graphs. We focus on the celebrated Erdős-Rényi model. We will also encounter preferential attachment graphs.

Erdős-Rényi graphs The Erdős-Rényi random graph is defined as follows.

Definition 1.27 (Erdős-Rényi graphs). *Let $V = [n]$ and $p \in [0, 1]$. The Erdős-Rényi graph $G = (V, E)$ on n vertices with density p is defined as follows: for each pair $x \neq y$ in V , the edge $\{x, y\}$ is in E with probability p independently of all other edges. We write $G \sim \mathbb{G}_{n,p}$ and we denote the corresponding measure by $\mathbb{P}_{n,p}$.*

Erdős-Rényi graph

Typical questions regarding Erdős-Rényi graphs (and pointers to some answers) include:

- What is the probability of observing a particular subgraph, e.g., a triangle?
 - Sections 2.2.2, 4.3.5.
- Is G connected? If not, how large are the components?
 - Sections 2.2.4, 5.3.3.
- How are the degrees distributed?
 - Section 4.2.2.
- What is the typical chromatic number?
 - Section 3.2.3.

Preferential attachment graphs In the definition of preferential attachment graphs, we consider *multigraphs*, i.e., we allow parallel edges.

Definition 1.28 (Preferential attachment graph). *The preferential attachment process produces a sequence of graphs $(G_t)_{t \geq 1}$ as follows. We start at time 1 with two vertices, denoted v_0 and v_1 , connected by an edge. At time t , we add vertex v_t with a single edge connecting it to an old vertex, which is picked proportionally to its degree. We write $(G_t)_{t \geq 1} \sim \text{PA}_1$. For $m > 1$, we run the previous process for time nm and then collapse vertices $km, km-1, \dots, (k-1)m + 1$ to make vertex k .*

*preferential
attachment
graphs*

Typical questions regarding preferential attachment graphs (and pointers to some answers) include:

- How are the degrees distributed?
 - Section 3.2.5.

1.4 Markov and Gibbs random fields

We start with Gibbs random fields.

Definition 1.29 (Gibbs random field). *Let S be a finite set and let $G = (V, E)$ be a finite graph. Denote by \mathcal{K} the set of all cliques of G . A positive probability measure μ on $\mathcal{X} := S^V$ is called a Gibbs random field if there exist clique potentials ϕ_K :*

*Gibbs random
field*

$S^K \rightarrow \mathbb{R}$, $K \in \mathcal{K}$, such that

$$\mu(x) = \frac{1}{\mathcal{Z}} \exp \left(\sum_{K \in \mathcal{K}} \phi_K(x_K) \right),$$

where x_K is x restricted to the vertices of K and \mathcal{Z} is a normalizing constant.

The following example introduces the primary Gibbs random field we will consider.

Example 1.30 (Ising model). For $\beta > 0$, the (ferromagnetic) Ising model with inverse temperature β is the Gibbs random field with $S := \{-1, +1\}$, $\phi_{\{i,j\}}(\sigma_{\{i,j\}}) = \beta \sigma_i \sigma_j$ and $\phi_K \equiv 0$ if $|K| \neq 2$. The function $\mathcal{H}(\sigma) := -\sum_{\{i,j\} \in E} \sigma_i \sigma_j$ is known as the Hamiltonian. The normalizing constant $\mathcal{Z} := \mathcal{Z}(\beta)$ is called the partition function. The states $(\sigma_i)_{i \in V}$ are referred to as spins. Ising model
Hamiltonian,
partition
function,
spin

Typical questions regarding Gibbs random field (and pointers to some answers) include:

- How fast is correlation decaying?

Proof of Hammersley-Clifford To be written. See [Gri10a, Theorem 7.12].

1.5 Random walks on networks

Recall the following definition.

Definition 1.31 (Simple random walk on a graph). Let $G = (V, E)$ be a finite or countable, locally finite graph. Simple random walk on G is the Markov chain on V , started at an arbitrary vertex, which at each time picks a uniformly chosen neighbor of the current state. simple random
walk on a graph

We generalize the definition slightly by adding weights to the edges.

Definition 1.32 (Random walk on a network). Let $G = (V, E)$ be a finite or countable, locally finite graph. Let $c : E \rightarrow \mathbb{R}_+$ be a positive edge weight function on G . We call $\mathcal{N} = (G, c)$ a network. We assume that for all $u \in V$

$$c_u := \sum_{v: e=\{u,v\} \in E} c(e) < +\infty.$$

network,
random walk
on a network

Random walk on network \mathcal{N} is the Markov chain on V , started at an arbitrary vertex, which at each time picks a neighbor of the current state proportionally to the weight of the corresponding edge, i.e., the transition matrix is given by

$$P(u, v) = \begin{cases} \frac{c(\{u, v\})}{c_u} & \text{if } \{u, v\} \in E \\ 0 & \text{o.w.} \end{cases}$$

Any countable, reversible Markov chain can be seen as a random walk on a network (not necessarily locally finite) by setting $c(e) := \pi(x)P(x, y) = \pi(y)P(y, x)$ for all $e = \{x, y\} \in E$.

Typical questions regarding random walks on networks (and pointers to some answers) include:

- *How long does it take to visit all vertices at least once or a particular subset of vertices for the first time?*
 - This question is related to the notions of commute time and hitting time, which are defined in Section 3.1.1.
 - General results: Sections 3.1.2, 3.3.1, 3.3.2, 3.3.3.
 - Special chains: birth-and-death chains (Section 3.3.2); trees (Section 3.3.2); complete graph (Section 3.3.3).
 - Application to branching processes: Section 5.1.
- *How often does the walk return to its starting point?*
 - This question is related to the notion of recurrence, which is defined in Section 3.1.1.
 - General results: Sections 3.3.2, 3.3.3, 6.1.4.
 - Special chains: lattice (Section 3.3.3); trees (Section 3.3.3).
- *How fast does it approach stationarity?*
 - This question is related to the notion of mixing time, which is defined in Section 1.1.2.
 - General results: lower bounds (Sections 2.3.4, 6.1.1, 6.2.1); upper bounds (Sections 4.4.1, 4.4.3, 6.1.1, 6.2.1).
 - Special chains: cycle (Sections 4.4.2, 6.1.2, 6.2.2); trees (Sections 4.4.2, 6.2.2); hypercube (Sections 4.4.2, 6.1.2, 6.2.2).

1.6 Interacting particles on graphs

We consider a particular class of Markov chains known as interacting particle systems. We focus on finite graphs.

Glauber dynamics We begin the Glauber dynamics of the Ising model.

Definition 1.33 (Glauber dynamics for the Ising model). *Let μ_β be the Ising model with inverse temperature $\beta > 0$ on a graph $G = (V, E)$. The (single-site) Glauber dynamics is the Markov chain on $\mathcal{X} := \{-1, +1\}^V$ which at each time:*

- selects a site $i \in V$ uniformly at random, and
- updates the spin at i according to μ_β conditioned on agreeing with the current state at all sites in $V \setminus \{i\}$.

*Glauber
dynamics*

Specifically, for $\gamma \in \{-1, +1\}$, $i \in \Lambda$, and $\sigma \in \mathcal{X}$, let $\sigma^{i,\gamma}$ be the configuration σ with the spin at i being set to γ . Let $n = |V|$ and $S_i(\sigma) := \sum_{j \sim i} \sigma_j$. Because the Ising measure factorizes, the nonzero entries of the transition matrix are

$$Q_\beta(\sigma, \sigma^{i,\gamma}) := \frac{1}{n} \cdot \frac{e^{\gamma\beta S_i(\sigma)}}{e^{-\beta S_i(\sigma)} + e^{\beta S_i(\sigma)}}.$$

Theorem 1.34. *The Glauber dynamics is reversible w.r.t. μ_β .*

Proof. This chain is clearly irreducible. For all $\sigma \in \mathcal{X}$ and $i \in V$, let

$$S_{\neq i}(\sigma) := \mathcal{H}(\sigma^{i,+}) + S_i(\sigma) = \mathcal{H}(\sigma^{i,-}) - S_i(\sigma).$$

We have

$$\begin{aligned} \mu_\beta(\sigma^{i,-}) Q_\beta(\sigma^{i,-}, \sigma^{i,+}) &= \frac{e^{-\beta S_{\neq i}(\sigma)} e^{-\beta S_i(\sigma)}}{\mathcal{Z}(\beta)} \cdot \frac{e^{\beta S_i(\sigma)}}{n[e^{-\beta S_i(\sigma)} + e^{\beta S_i(\sigma)}]} \\ &= \frac{e^{-\beta S_{\neq i}(\sigma)}}{n\mathcal{Z}(\beta)[e^{-\beta S_i(\sigma)} + e^{\beta S_i(\sigma)}]} \\ &= \frac{e^{-\beta S_{\neq i}(\sigma)} e^{\beta S_i(\sigma)}}{\mathcal{Z}(\beta)} \cdot \frac{e^{-\beta S_i(\sigma)}}{n[e^{-\beta S_i(\sigma)} + e^{\beta S_i(\sigma)}]} \\ &= \mu_\beta(\sigma^{i,+}) Q_\beta(\sigma^{i,+}, \sigma^{i,-}). \end{aligned}$$

That concludes the proof. ■

Typical questions regarding the Glauber dynamics (and pointers to some answers) include:

- How quickly does the chain approach μ_β as a function of β ? This question is related to the notion of mixing time, as discussed above in the context of random walks on networks. The answer of course depends on the graph.
 - General results: Section 4.4.4.
 - Special graphs: empty graph (Sections 4.4.2, 6.1.2, 6.2.2); complete graph (Section 6.2.3); expander graphs (Section 6.2.3).

Voter model Our final model:

Definition 1.35 (Voter model). *Let $S := \{0, 1\}$ and let $G = (V, E)$ be a finite, connected graph. The voter model on G is the Markov chain (η_t) on S^V which, at each time t , picks a uniform site $i \in V$ as well as a uniform neighbor $j \sim i$ and sets $\eta_t(i) := \eta_{t-1}(j)$.* *voter model*

Typical questions regarding the voter model (and pointers to some answers) include:

- How long does it take to reach one of the two *absorbing states*, i.e., the all-0 and all-1 configurations?
 - Section 4.5.1.

Bibliographic remarks

For an introduction to graph theory, see for example [Die10] or [Bol98]. Markov chain theory is covered in details in [Dur10, Chapter 6]. For a more gentle introduction, see e.g. [Dur12, Chapter 1] or [Res92, Chapter 2]. See also [LPW06, Chapters 1-3].