Interacting Particle Systems on Random Graphs

F. Capannoli, F. den Hollander *

Abstract

The present overview of interacting particle systems on random graphs collects the notes of a mini-course given by the authors at the Brazilian School of Probability, 5–9 August 2024, in Salvador, Bahia, Brazil. The content is a personal snapshot of an interesting area of research at the interface between probability theory, combinatorics, statistical physics and network science that is developing rapidly.

Keywords: Interacting particle systems, stochastic Ising model, voter model, contact process, sparse and dense random graphs, space-time scaling, phase transitions.

MSC 2020: 05C80; 60C05; 60K35; 82B26.

Acknowledgement: The authors are supported by the Netherlands Organisation for Scientific Research (NWO) through Gravitation Grant NETWORKS-024.002.003. FC is also supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 945045.

*Mathematical Institute, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands.

Contents

0 An overview: structure and goals **3**

1 LECTURE 1: Interacting Particle Systems on $\mathbb{Z}^d$, $d \geq 1$. General properties and three examples **3**

1.1 Spin-flip systems 4
1.2 Shift-invariant attractive systems 5
1.3 Three examples 5
1.4 Partial ordering and coupling 6
1.5 Convergence to equilibrium 6
1.6 Properties of the three examples 9
1.7 The Cox-Greven finite systems scheme 11

2 LECTURE 2: The Stochastic Ising Model (SIM) **14**

2.1 SIM on graphs 14
2.2 SIM on the complete graph 15
2.3 SIM on random graphs 16
2.4 SIM on the Erdős-Rényi random graph 17
2.5 SIM on the inhomogeneous Erdős-Rényi random graph 18
2.6 SIM on sparse graphs 18
2.7 SIM on the configuration model 19

3 LECTURE 3: The Voter Model (VM) **23**

3.1 VM duality and graphical representation 23
3.2 VM on the complete graph 27
3.3 VM on the random regular graph 30
3.4 VM meeting times and coalescence times 34
3.5 VM on the random regular graph with random rewiring 37

4 LECTURE 4: The Contact Process (CP) **38**

4.1 CP on graphs 38
4.2 CP on the complete graph 38
4.3 CP on the configuration model 39
4.4 CP on the Erdős-Rényi random graph 40
4.5 CP on the preferential attachment model 40
4.6 CP on tree-like random graphs 41
4.7 CP on general finite graphs 41

0 An overview: structure and goals

In the present overview we focus on *interacting particle systems on random graphs*. This is a relatively new research area in which progress is rapid, problems are challenging, and new panoramas unfold. It lies at the interface between *probability theory*, *combinatorics*, *statistical physics* and *network science*, and as such is captivating.

The overview is structured as follows:

- Section 1: Lecture 1. Background and motivation for Interacting Particle Systems (IPS) on $\mathbb{Z}^d$, $d \geq 1$. Key questions and core tools. Phase transitions.
- Section 2: Lecture 2. Stochastic Ising Model (SIM) on random graphs.
- Section 3: Lecture 3. Voter Model (VM) on random graphs.
- Section 4: Lecture 4. Contact Process (CP) on random graphs.

In Lectures 2–4, IPS on *four classes of random graphs* will be considered:

- Homogeneous Erdős-Rényi Random Graph (HER).
- Inhomogeneous Erdős-Rényi Random Graph (IER).
- Configuration Model (CM).
- Preferential Attachment Model (PAM).

The goal of the overview is to:

- Sketch what is known and not known about IPS on random graphs.
- Highlight the role of sparse versus dense graphs.
- Exhibit the relevant time scales for critical phenomena and identify how these depend on the size of the graph.
- List some open problems and indicate some lines of future research.

Standard references for IPS on $\mathbb{Z}^d$ are Liggett [38, 39]. Standard references for random graphs are van der Hofstad [34, 35]. In what follows, proofs of theorems are sometimes included but often not. For details the reader is referred to the relevant references that are given.

1 LECTURE 1: Interacting Particle Systems on $\mathbb{Z}^d$, $d \geq 1$. General properties and three examples

As an area of research, IPS started in the 1970s, with pioneers Spitzer, Dobrushin, Harris, Holley, Stroock, Liggett, Griffeath, Durrett. Over the years, IPS has turned out to be a fertile breeding ground for the development of new ideas and techniques in mathematical statistical physics, including *graphical representation*, *coupling*, *duality* and *correlation inequalities*.

1.1 Spin-flip systems

We start by defining what an IPS is. We focus on *spin-flip systems*, which constitute a particularly tractable class. Within this class we focus on *three examples*:

Stochastic Ising Model (SIM)
Voter Model (VM)
Contact Process (CP)

Standard references for IPS on $\mathbb{Z}^d$ are Liggett [38, 39]. For most of the results to be described in this section, references can be found in these monographs.

Definition. An Interacting Particle System (IPS) is a Markov process $\xi = (\xi_t)_{t\geq 0}$ on the state space

$$\Omega = \{0,1\}^{\mathbb{Z}^d}, \qquad d \geq 1,$$

where

$$\xi_t = \{\xi_t(x)\colon\ x \in \mathbb{Z}^d\}$$

denotes the configuration at time t, with $\xi_t(x) = 1$ or 0 meaning that there is a 'particle' or a 'hole' at site x at time t, respectively. Alternative interpretations are:

$$\begin{array}{rcl} 1 & = & \text{spin-up/democrat/infected} \\ 0 & = & \text{spin-down/republican/healthy.} \end{array}$$

♠

The configuration changes over time, which models how:

- magnetic atoms flip up and down as a result of noise,
- two political parties evolve in an election campaign,
- a virus spreads through a population.

The evolution is specified via a set of *local transition rates*

$$c(x,\eta), \qquad x \in \mathbb{Z}^d,\ \eta \in \Omega,$$

playing the role of the rate at which the state at site x changes in the configuration η, i.e.,

$$\eta \to \eta^x$$

with η^x the configuration obtained from η by changing the state at site x (either $0 \to 1$ or $1 \to 0$). Since there are only two possible states at each site, the IPS is called a *spin-flip system*.

If $c(x,\eta)$ depends on η only via $\eta(x)$, the value of the spin at site x, then ξ consists of independent spin-flips. In general, however, the rate to flip the spin at x depends on the spins located in the neighbourhood of x (possibly even on all spins). This dependence models an *interaction* between the spins at different sites.

In order for ξ to be well-defined, some restrictions must be placed on the local transition rates: $c(x,\eta)$ must depend only weakly on the states at far away sites (formally, $\eta \mapsto c(x,\eta)$ is continuous in the product topology) and must be not too large (formally, bounded away from infinity in some appropriate sense). See Liggett [38, Chapter I] for details.

1.2 Shift-invariant attractive systems

A typical assumption is that

$$c(x,\eta) = c(x+y,\tau_y\eta), \qquad y \in \mathbb{Z}^d,$$

with τ_y the shift of space over y, i.e.,

$$(\tau_y\eta)(x) = \eta(x-y), \qquad x \in \mathbb{Z}^d.$$

This property says that the flip rate at x only depends on the configuration η seen relative to x, which is natural when the interaction between spins is homogeneous in space. Spin-flip systems with this property are called *shift-invariant*.

Another useful assumption is that the interaction favours spins that are alike, i.e.,

$$\eta \preceq \eta' \to \begin{cases} c(x,\eta) \le c(x,\eta') & \text{if } \eta(x) = \eta'(x) = 0, \\ c(x,\eta) \ge c(x,\eta') & \text{if } \eta(x) = \eta'(x) = 1, \end{cases}$$

where $\preceq$ denotes the partial order in Ω. This property says that the spin at x flips up faster in η' than in η when η' is everywhere larger than η, and flips down slower. In other words, the dynamics preserves $\preceq$. Spin-flip systems with this property are called *attractive*.

1.3 Three examples

1. SIM: *Stochastic Ising model*.

   This model is defined on $\Omega = \{-1,1\}^{\mathbb{Z}^d}$ with rates

   $$c(x,\eta) = \exp\left[-\beta\eta(x)\sum_{y\sim x}\eta(y)\right], \qquad \beta \ge 0,$$

   which means that spins prefer to align with the majority of the neighbouring spins.

2. VM: *Voter model*.

   This model is defined on $\Omega = \{0,1\}^{\mathbb{Z}^d}$ with rates

   $$c(x,\eta) = \frac{1}{2d}\sum_{y\sim x} 1_{\{\eta(y)\neq\eta(x)\}},$$

   which means that sites choose a random neighbour at rate 1 and adopt the opinion of that neighbour.

3. CP: *Contact process*.

   This model is defined on $\Omega = \{0,1\}^{\mathbb{Z}^d}$ with rates

   $$c(x,\eta) = \begin{cases} \lambda \sum\limits_{y\sim x}\eta(y), & \text{if } \eta(x) = 0, \\ 1, & \text{if } \eta(x) = 1, \end{cases} \qquad \lambda \ge 0,$$

   which means that infected sites become healthy at rate 1 and healthy sites become infected at rate λ times the number of infected neighbours.

It is easy to check that all three models are shift-invariant and attractive. Below we will discuss each model in some detail. We will see that the properties shift-invariant and attractive allow for a number of interesting conclusions concerning their *equilibrium*, as well as their convergence to equilibrium.

1.4 Partial ordering and coupling

Given two probability measures P, P' on $\Omega = \{0,1\}^{\mathbb{Z}^d}$ with Borel σ-algebra $\mathcal{B}$, we say that P' *stochastically dominates* P, and write $P \preceq P'$, if

$$P(A) \leq P'(A) \text{ for all } A \subseteq \mathcal{B} \text{ non-decreasing},$$

where A non-decreasing means

$$x \in A \implies A \supseteq \{y \in \Omega\colon\ x \preceq y\},$$

and $\preceq$ is a partial ordering on Ω (in our case the natural partial ordering induced by the ordering of the components at each site). Equivalently, P' stochastically dominates P if

$$\int_\Omega f \,\mathrm{d}P \leq \int_\Omega f \,\mathrm{d}P' \text{ for all } f\colon\ \Omega \to \mathbb{R} \text{ measurable, bounded and non-decreasing},$$

where f non-decreasing means

$$x \preceq y \implies f(x) \leq f(y).$$

A *coupling* of P, P' on Ω is any joint probability measure $\hat{P}$ on $\Omega \times \Omega$ with marginals P, P'. Strassen's theorem (see Lindvall [40]) says that if $P \preceq P'$, then there exists a coupling $\hat{P}$ of (P, P') such that

$$\hat{P}\{(x, x') \in \Omega \times \Omega\colon\ x \preceq x'\} = 1.$$

In fact, on a partially ordered space Ω the following three statements are equivalent:

1. $P \preceq P'$,
2. $\int_\Omega f \,\mathrm{d}P \leq \int_\Omega f \,\mathrm{d}P'$ for all f measurable, bounded and non-decreasing,
3. $\exists\, \hat{P}\colon\ \hat{P}(\hat{X} \preceq \hat{X}') = 1$,

where X, X' have marginal laws P, P' and $(\hat{X}, \hat{X}')$ has joint law $\hat{P}$.

1.5 Convergence to equilibrium

Write [0] and [1] to denote the configurations $\eta \equiv 0$ and $\eta \equiv 1$, respectively. These are the smallest, respectively, the largest configurations in the natural *partial order* on Ω, and hence

$$[0] \preceq \eta \preceq [1], \qquad \eta \in \Omega.$$

Since the dynamics preserves the partial order (see below), we obtain information about what happens when the system starts from any $\eta \in \Omega$ by comparing with what happens when it starts from [0] or [1].

An IPS can be described by a *semigroup of transition kernels*

$$(P_t)_{t \geq 0}.$$

Formally, P_t is an operator acting on $C_b(\Omega)$, the space of bounded continuous functions on Ω, as

$$(P_t f)(\eta) = \mathbb{E}_\eta[f(\xi_t)], \qquad \eta \in \Omega,\ f \in C_b(\Omega).$$

If this definition holds on a *dense* subset of $C_b(\Omega)$, then it *uniquely* determines P_t. Note that P_0 is the identity and that $P_{t+s} = P_t \circ P_s$ for all $s, t \geq 0$ (where $\circ$ denotes composition). To see the latter, note that

$$(P_0 f)(\eta) = \mathbb{E}^\eta[f(\xi_0)] = f(\eta),$$

where $\mathbb{P}^\eta$ is the law of ξ with initial position η:

$$\mathbb{P}^\eta(\xi_\cdot \in C_b(\Omega) \colon \xi_0 = \eta) = 1.$$

Moreover, for $t \geq 0$ let $\mathcal{F}_t$ be the σ-algebra generated by $(\xi_s)_{0 \leq s \leq t}$. Then

$$\begin{aligned}(P_{t+s} f)(\eta) &= \mathbb{E}^\eta[f(\xi_{t+s})] = \mathbb{E}^\eta[\mathbb{E}^\eta[f(\xi_{t+s})] \mid \mathcal{F}_t] = \mathbb{E}^\eta[\mathbb{E}^{\xi_t}[f(\xi_s)]] \\ &= \mathbb{E}^\eta[(P_t f)(\xi_s)] = ((P_t \circ P_s) f)(\eta)\,,\end{aligned}$$

where the third equality uses the Markov property of ξ.

Formally, we can write $P_t = \mathrm{e}^{tL}$ with L the *generator* of the IPS:

$$(Lf)(\eta) = \sum_{x \in \mathbb{Z}^d} c(x, \eta)[f(\eta^x) - f(\eta)].$$

Alternatively, the semigroup can be viewed as acting on the space of probability measures μ on Ω via the *duality relation*

$$\int_\Omega f \,\mathrm{d}(\mu P_t) = \int_\Omega (P_t f)\,\mathrm{d}\mu, \qquad f \in C_b(\Omega).$$

Lemma 1.1 *Let $(P_t)_{t \geq 0}$ denote the semigroup of transition kernels that is associated with $\xi = (\xi_t)_{t \geq 0}$. Write $\delta_\eta P_t$ to denote the law of ξ_t conditional on $\xi_0 = \eta$ (which is a probability distribution on Ω). Then*

$$\begin{aligned} t &\mapsto \delta_{[0]} P_t \quad \textit{is stochastically increasing,} \\ t &\mapsto \delta_{[1]} P_t \quad \textit{is stochastically decreasing.}\end{aligned}$$

Proof. For $t, h \geq 0$,

$$\begin{aligned}\delta_{[0]} P_{t+h} &= (\delta_{[0]} P_h) P_t \succeq \delta_{[0]} P_t, \\ \delta_{[1]} P_{t+h} &= (\delta_{[1]} P_h) P_t \preceq \delta_{[1]} P_t,\end{aligned}$$

where we use that

$$\delta_{[0]} P_h \succeq \delta_{[0]}, \qquad \delta_{[1]} P_h \preceq \delta_{[1]}, \qquad h \geq 0,$$

and we use Strassen's theorem in combination with the coupling representation that goes with the partial order. (Strassen's theorem says that stochastic ordering is equivalent to the existence of an ordered coupling.) ∎

Corollary 1.2 *Both*

$$\begin{aligned}\underline{\nu} &= \lim_{t \to \infty} \delta_{[0]} P_t \;= \textit{lower stationary law,} \\ \overline{\nu} &= \lim_{t \to \infty} \delta_{[1]} P_t \;= \textit{upper stationary law,}\end{aligned}$$

exist as probability distributions on Ω and are equilibria for the dynamics. Any other equilibrium π satisfies $\underline{\nu} \preceq \pi \preceq \overline{\nu}$.

Proof. This is immediate from Lemma 1.1 and the sandwich

$$\delta_{[0]}P_t \preceq \delta_\eta P_t \preceq \delta_{[1]}P_t, \qquad \eta \in \Omega,\, t \geq 0.$$

∎

The class of all equilibria for the dynamics is a convex set in the space of signed bounded measures on Ω. An element of this set is called *extremal* when it is not a proper linear combination of any two distinct elements in the set, i.e., is not of the form

$$p\nu_1 + (1-p)\nu_2, \qquad p \in (0,1),\, \nu_1 \neq \nu_2.$$

Lemma 1.3 *Both $\underline{\nu}$ and $\overline{\nu}$ are extremal.*

Proof. We give the proof for $\overline{\nu}$ only. Suppose that

$$\overline{\nu} = p\nu_1 + (1-p)\nu_2, \qquad \nu_1 \neq \nu_2,\, p \in (0,1).$$

Since ν_1 and ν_2 are equilibria, Corollary 1.2 gives

$$\int_\Omega f \mathrm{d}\nu_1 \leq \int_\Omega f \mathrm{d}\overline{\nu}, \qquad \int_\Omega f \mathrm{d}\nu_2 \leq \int_\Omega f \mathrm{d}\overline{\nu},$$

for any f non-decreasing. Since

$$\int_\Omega f \mathrm{d}\overline{\nu} = p\int_\Omega f \mathrm{d}\nu_1 + (1-p)\int_\Omega f \mathrm{d}\nu_2$$

and $p \in (0,1)$, both inequalities must be equalities. Integrals of non-decreasing functions determine the measure w.r.t. which is being integrated, and so it follows that $\nu_1 = \overline{\nu} = \nu_2$.

To see why the family $\{\int_\Omega f \,\mathrm{d}\nu\colon\ f\colon\ \Omega \to \mathbb{R}$ non-decreasing$\}$ determines the measure ν, it suffices to consider functions of the form $f(\eta) = 1_{\eta \supseteq \xi}$, $\eta \in \Omega$, with ξ running over Ω. For such functions $\int_\Omega f \,\mathrm{d}\nu = \nu(A_\xi)$ with $A_\xi = \{\eta \in \Omega\colon\ \eta \supseteq \xi\}$ a non-increasing subset of Ω, and so the claim follows from the monotone class theorem, which says that the sigma-algebra of all events is the same as the sigma-algebra of monotone events. ∎

Corollary 1.4 *The following three properties are equivalent (for shift-invariant and attractive spin-flip systems):*

1. *ξ is ergodic (i.e., $\delta_\eta P_t$ has the same limiting distribution as $t \to \infty$ for all η).*
2. *There is a unique stationary distribution.*
3. *$\underline{\nu} = \overline{\nu}$.*

Proof. The claim is obvious in view of the sandwich of the configurations between [0] and [1]. ∎

REMARK: If $\underline{\nu} \neq \overline{\nu}$, then there is no guarantee that $\lim_{t\to\infty} \mu P_t = \nu$ exists for arbitrary μ. In fact, stronger assumptions than attractiveness are needed to make that happen. We do know that any convergent subsequence has a limit ν such that $\underline{\nu} \preceq \nu \preceq \overline{\nu}$.

1.6 Properties of the three examples

Example 1: Stochastic Ising Model. For $\beta = 0$, $c(x, \eta) = 1$ for all x and η, in which case the dynamics consists of independent spin-flips, up and down at rate 1. In that case $\overline{\nu} = \underline{\nu} = (\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_{+1})^{\otimes \mathbb{Z}^d}$. For $\beta > 0$, the dynamics has a tendency to align the spins. For small β this tendency is weak, for large β it is strong. It turns out that in $d \geq 2$ there is a *critical value* $\beta_d \in (0, \infty)$ such that

$$\begin{aligned} \beta \leq \beta_d\colon &\quad \underline{\nu} = \overline{\nu}, \\ \beta > \beta_d\colon &\quad \underline{\nu} \neq \overline{\nu}. \end{aligned}$$

The proof uses the so-called Peierls argument. In the first case there is a *unique ergodic equilibrium*, which depends on β and is denoted by ν_β. In the second case there are *two extremal equilibria*, both of which depend on β and are denoted by

$$\begin{aligned} \nu_\beta^+ &= \text{plus state} \quad \text{with} \quad m_\beta^+ = \textstyle\int_\Omega \eta(0)\nu_\beta^+(\mathrm{d}\eta) > 0, \\ \nu_\beta^- &= \text{minus-state} \quad \text{with} \quad m_\beta^- = \textstyle\int_\Omega \eta(0)\nu_\beta^-(\mathrm{d}\eta) < 0, \end{aligned}$$

which are called the magnetised states. Note that ν_β^+ and ν_β^- are images of each other under the swapping of $+1$'s and -1's and so $m_\beta^+ = -m_\beta^- = m_\beta$.

It can be shown that in $d = 2$ all equilibria are a convex combination of ν_β^+ and ν_β^-, while in $d \geq 3$ *other equilibria are possible* as well (e.g. not shift-invariant) when β is large enough. It turns out that $\beta_1 = \infty$, i.e., in $d = 1$ the SIM is ergodic for all $\beta > 0$. It is known that $\beta_2 = \frac{1}{2}\log(1 + \sqrt{2})$.

Example 2: Voter Model. Note that [0] and [1] are both *traps* for the dynamics (if all sites have the same opinion, then no change of opinion occurs), and so

$$\underline{\nu} = \delta_{[0]}, \qquad \overline{\nu} = \delta_{[1]}.$$

It turns out that in $d = 1, 2$ these are the only extremal equilibria, while in $d \geq 3$ there is a 1-*parameter family of extremal equilibria*

$$(\nu_\rho)_{\rho \in [0,1]}$$

with ρ the density of 1's, i.e., $\nu_\rho(\eta(0) = 1) = \rho$. This fact is remarkable because the VM has no parameter. For $\rho = 0$ and $\rho = 1$ these equilibria coincide with $\delta_{[0]}$ and $\delta_{[1]}$, respectively.

REMARK: The dichotomy $d = 1, 2$ versus $d \geq 3$ is directly related to simple random walk being recurrent in $d = 1, 2$ and transient in $d \geq 3$. This property has to do with the fact that the VM is *dual* to a system of coalescing random walks.

Example 3: Contact Process. Note that [0] is a trap for the dynamics (if all sites are healthy, then no infection will ever occur), and so

$$\underline{\nu} = \delta_{[0]}.$$

For small λ infection is transmitted slowly, for large λ rapidly. It turns out that in $d \geq 1$ there is a *critical value* $\lambda_d \in (0, \infty)$ such that

$$\begin{aligned} \lambda \leq \lambda_d\colon &\quad \overline{\nu} = \delta_{[0]} \quad = \text{extinction, no epidemic}, \\ \lambda > \lambda_d\colon &\quad \overline{\nu} \neq \delta_{[0]} \quad = \text{survival, epidemic}. \end{aligned}$$

Lemma 1.5 (Liggett [38], Durrett [30].)
(i) $2d\lambda_d \geq 1$.
(ii) $d\lambda_d \leq \lambda_1$.
(iii) $\lambda_1 < \infty$.

Note that (i–iii) combine to yield that $0 < \lambda_d < \infty$ for all $d \geq 1$, so that the phase transition occurs at a *non-trivial* value of the infection rate parameter. Here is a proof of (i) and (ii).

Proof. (i) Pick A_0 finite and consider the CP in dimension d with parameter λ starting from the set A_0 as the set of infected sites. Let $A = (A_t)_{t\geq 0}$ with A_t the set of infected sites at time t. Then

$$\begin{aligned} &|A_t| \text{ decreases by 1 at rate } |A_t|, \\ &|A_t| \text{ increases by 1 at rate } \leq 2d\lambda|A_t|, \end{aligned}$$

where the latter holds because each site in A_t has at most $2d$ non-infected neighbours. Now consider the two random processes $X = (X_t)_{t\geq 0}$ with $X_t = |A_t|$ and $Y = (Y_t)_{t\geq 0}$ given by the birth-death process on $\mathbb{N}_0$ that moves at rate n from n to $n-1$ (death) and at rate $(2d\lambda)n$ from n to $n+1$ (birth), both starting from $n_0 = |A_0|$. Then X and Y can be coupled such that

$$\hat{P}(X_t \leq Y_t \ \forall\, t \geq 0) = 1,$$

where $\hat{P}$ denotes the coupling measure. Note that $n = 0$ is a trap for both X and Y. If $2d\lambda < 1$, then this trap is hit with probability 1 by Y, i.e., $\lim_{t\to\infty} Y_t = 0$ a.s., and hence also by X, i.e., $\lim_{t\to\infty} X_t = 0$ a.s. Therefore $\nu_\lambda = \delta_0$ when $2d\lambda < 1$. Consequently, $2d\lambda_d \geq 1$.

(ii) The idea is to couple two CP's that live in dimensions 1 and d. Again, let $A = (A_t)_{t\geq 0}$ with A_t the set of infected sites at time t of $\mathrm{CP}_d(\lambda)$, the CP in dimension d with parameter λ, this time starting from $A_0 = \{0\}$. Let $B = (B_t)_{t\geq 0}$ be the same as A, but for $\mathrm{CP}_1(\lambda d)$, the CP in dimension 1 with parameter λd, starting from $B_0 = \{0\}$.

Define the projection $\pi_d\colon\ \mathbb{Z}^d \to \mathbb{Z}$ as

$$\pi_d(x_1, \ldots, x_d) = x_1 + \cdots + x_d.$$

We will construct a coupling $\hat{P}$ of A and B such that

$$\hat{P}(B_t \subseteq \pi_d(A_t) \ \forall\, t \geq 0) = 1.$$

From this we get

$$P(A_t \neq \emptyset \mid A_0 = \{0\}) = P\big(\pi_d(A_t) \neq \emptyset \mid A_0 = \{0\}\big) \geq P\big(B_t \neq \emptyset \mid B_0 = \{0\}\big),$$

which implies that if A dies out, then also B dies out. In other words, if $\lambda \leq \lambda_d$, then $\lambda d \leq \lambda_1$, which implies that $d\lambda_d \leq \lambda_1$ as claimed. The construction of the coupling is as follows. Fix $t \geq 0$. Suppose that $B_t \subseteq \pi_d(A_t)$. For each $y \in B_t$ there is at least one $x \in A_t$ with $y = \pi_d(x)$. Pick one such x for every y (e.g. choose the closest up or the closest down). Now couple:

- If x becomes healthy, then y becomes healthy too.
- If x infects any of the d sites $x - e_i$ with $i = 1, \ldots, d$, then y infects $y - 1$.
- If x infects any of the d sites $x + e_i$ with $i = 1, \ldots, d$, then y infects $y + 1$.

It is much harder to prove (iii). One way is to compare the CP with directed percolation in two dimensions (= space × time). See Durrett [30] for details. ∎

REMARK: Sharp estimates are available for λ_1, but these require heavy machinery. Numerically, $\lambda_1 \approx 1.6494$. A series expansion of λ_d in powers of $1/2d$ is known up to several orders, but again the proof is very technical.

1.7 The Cox-Greven finite systems scheme

As a prelude to Lectures 2–4, in which we take a closer look at SIM, VM, CP on *finite random graphs*, we describe what is known about these processes on a large *finite torus* in $\mathbb{Z}^d$,

$$\Lambda_N = [0, N)^d \cap \mathbb{Z}^d, \qquad N \in \mathbb{N},$$

endowed with periodic boundary conditions. The behaviour on Λ_N is different from that on $\mathbb{Z}^d$. In particular, there is an N-dependent *characteristic time scale* α_N on which the process notices that Λ_N differs from $\mathbb{Z}^d$, resulting in *different behaviour for short, moderate and long times*. A systematic study was initiated in Cox, Greven [23], Cox, Greven, Shiga [24, 25].

WARNING: The text in the remainder of this section is technical.

• **SIM ON THE TORUS.** Since $|\Lambda_N| < \infty$, we have

$$\underline{\nu}^N = \overline{\nu}^N = \nu_\beta^N \quad \text{with} \quad \int_\Omega \eta(0)\nu_\beta^N(\mathrm{d}\eta) = 0 \qquad \forall\, \beta \in (0, \infty),$$

i.e., on any finite lattice eventually the average magnetisation vanishes. An interesting question is: How long does it take the SIM to loose its magnetisation and what does it do along the way?

Let

$$\mathcal{M}_t^N = \frac{1}{|\Lambda_N|} \sum_{x \in \Lambda_N} \xi_t^N(x)$$

denote the magnetisation at time t. Suppose that the law of ξ_0^N is the restriction to Λ_N of the equilibrium measure ν_β^- on $\mathbb{Z}^d$, which has magnetisation m_β^-.

Theorem 1.6 (Cox, Greven [23], Bovier, Eckhoff, Gayrard, Klein [10].)
(a) For $\beta < \beta_d$ and any $T_N \to \infty$,

$$\lim_{N\to\infty} \mathcal{L}\big[\mathcal{M}_{T_N}^N\big] = \delta_0.$$

(b) For $\beta > \beta_d$,

$$\lim_{N\to\infty} \mathcal{L}\big[\mathcal{M}_{s\alpha_N}^N\big] = m_\beta^{Z_s}, \qquad Z_0 = -,$$

where $(Z_s)_{s\geq 0}$ is the Markov chain on $\{-,+\}$ jumping at rate 1, and α_N is the average crossover time between the magnetisations associated with ν_β^- and ν_β^+ on $\mathbb{Z}^d$ restricted to Λ_N.

For $\beta > \beta_d$ it can further be shown that $(\xi_s^N \alpha_N)_{s\geq 0}$ converges in distribution to $\nu_\beta^{Z_s}$ as $N \to \infty$.

The computation of α_N is hard and belongs to the area of *metastability*. It is expected that

$$\alpha_N = \exp\big[\kappa_d(\beta) N^{d-1}(1 + o(1))\big]$$

with $\kappa_d(\beta)$ the free energy of the so-called *Wulff droplet* of volume $\frac{1}{2}$ in $\mathbb{R}^d$ representing the barrier between ν_β^-, ν_β^+. The proof remains a challenge. See Schonmann, Shlosman [52] and Bovier, den Hollander [12, Chapter 22] for more background.

• **VM ON THE TORUS.** Since $|\Lambda_N| < \infty$, we have

$$\underline{\nu}^N = [0]_N, \qquad \overline{\nu}^N = [1]_N,$$

i.e., on any finite lattice eventually consensus is reached. An interesting question is: How long does it take the VM to reach consensus and what does it do along the way?

Let

$$\mathcal{O}_t^N = \frac{1}{|\Lambda_N|} \sum_{x\in\Lambda_N} \xi_t^N(x)$$

denote the fraction of 1-opinions at time t. Suppose that the law of ξ_0^N is the restriction to Λ_N of a shift-invariant and ergodic probability measure on $\mathbb{Z}^d$ with mean $\theta \in [0, 1]$.

Theorem 1.7 (Cox, Greven [23].)
(a) For $d = 1, 2$ and any $T_N \to \infty$,

$$\lim_{N\to\infty} \mathcal{L}\big[\mathcal{O}_{T_N}^N\big] = (1 - \theta)\delta_0 + \theta\delta_1.$$

(b) For $d \geq 3$,

$$\lim_{N\to\infty} \mathcal{L}\big[\mathcal{O}_{s\alpha_N}^N\big] = Z_s, \qquad Z_0 = \theta,$$

where $\alpha_N = |\Lambda_N|$ and $(Z_s)_{s\geq 0}$ is the Fisher-Wright diffusion on $[0, 1]$ with diffusion constant $1/G_d$, the inverse of the average number of visits to 0 of simple random walk on $\mathbb{Z}^d$.

A heuristic explanation of Theorem 1.7 is as follows. The VM is *dual* to a system of coalescing random walks, in the sense that the evolution of the genealogy of the opinions in the fomer is the *time reversal* of the latter. Since simple random walk is recurrent on $\mathbb{Z}^d$ for $d = 1, 2$ and transient for $d \geq 3$, the dichotomy between (a) and (b) is plausible. For duality and its relation to graphical representations of IPS, see Liggett [38, Chapter I].

• **CP ON THE TORUS.** Since $|\Lambda_N| < \infty$, we have

$$\underline{\nu}^N = \overline{\nu}^N = [0]_N \qquad \forall\, \lambda \in (0, \infty),$$

i.e., on a finite lattice every infection eventually becomes extinct, irrespective of the infection rate. An interesting question is the following: Starting from $[1]_N$, how long does it take the CP to reach $[0]_N$? In particular, we want to know the extinction time

$$\tau_{[0]_N} = \inf\{t \geq 0 \colon\, \xi_t^N = [0]_N\}.$$

We expect this time to grow slowly with N when $\lambda < \lambda_d$ and rapidly with N when $\lambda > \lambda_d$, where λ_d is the critical infection threshold for the infinite lattice $\mathbb{Z}^d$.

Let

$$\mathcal{I}_t^N = \frac{1}{|\Lambda_N|} \sum_{x \in \Lambda_N} \xi_t^N(x)$$

denote the fraction of infected vertices at time t. Suppose that $\xi_0^N = [1]_N$.

Theorem 1.8 (Cox, Greven [23].)
(a) For $\lambda < \lambda_d$ and any $T_N \to \infty$,

$$\lim_{N \to \infty} \mathcal{L}\big[\mathcal{I}_{T_N}^N\big] = \delta_0.$$

(b) For $\lambda > \lambda_d$,

$$\lim_{N \to \infty} \mathcal{L}\big[\mathcal{I}_{s\alpha_N}^N\big] = Z_s, \qquad Z_0 = 1,$$

where $\alpha_N = \mathbb{E}_{[1]_N}(\tau_{[0]_N})$ and $(Z_s)_{s \geq 0}$ is the Markov chain on $\{0,1\}$ that jumps from 1 to 0 at rate 1 and is absorbed in 0.

Theorem 1.9 (Durrett, Liu [31], Durrett, Schonmann [32], Mountford [43, 44].)
There exist $C_-(\lambda), C_+(\lambda) \in (0,\infty)$ such that

$$\lambda < \lambda_d: \qquad \lim_{N \to \infty} \frac{\alpha_N}{\log |\Lambda_N|} = C_-(\lambda),$$

$$\lambda > \lambda_d: \qquad \lim_{N \to \infty} \frac{\log \alpha_N}{|\Lambda_N|} = C_+(\lambda).$$

In the subcritical phase the extinction time grows *logarithmically* fast with the volume of Λ_N, while in the supercritical phase it grows *exponentially* fast. This is a rather dramatic dichotomy. Here is a heuristic explanation.

- *Subcritical phase*: When $\lambda < \lambda_c$, the infection cannot sustain itself in the long run. Each infected site has a higher tendency of becoming healthy than of becoming infected. Hence, the number of infected sites decreases over time, and the extinction time scales logarithmically with the system size because the infection dies out only when the last infected site has disappeared.

- *Supercritical phase*: When $\lambda > \lambda_c$, the infection can sustain itself and even grow. Each infected site has a higher tendency of becoming infected than of becoming healthy. Hence, the infection can create a large cluster of infected sites that persists for a long time. The extinction time scales exponentially with the system size because the infection dies out only after all sites become healthy in a short straight run. Many attempts are required to achieve this run, all of which have a small success probability. This implies the memoryless property of the extinction time, from which the exponential scaling follows.

Rough polynomial bounds on α_N are available in $d = 1$ at $\lambda = \lambda_1$ (Duminil-Copin, Tassion, Teixeira [29]).

2 LECTURE 2: The Stochastic Ising Model (SIM)

2.1 SIM on graphs

Let $G = (V, E)$ be a finite connected non-oriented graph. Ising spins are attached to the vertices V and interact with each other along the edges E (see Figure 1).

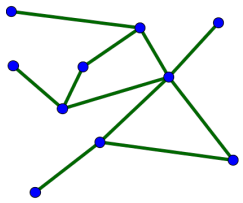

Figure 1: A finite connected non-oriented graph.

1. The energy associated with the configuration $\sigma = (\sigma_i)_{i\in V} \in \Omega = \{-1, +1\}^V$ is given by the Hamiltonian

$$H(\sigma) = -J \sum_{(i,j)\in E} \sigma_i\sigma_j - h\sum_{i\in V} \sigma_i$$

where $J > 0$ is the *ferromagnetic interaction strength* and $h > 0$ is the *external magnetic field*.

2. Spins flip according to *Glauber dynamics* $(\sigma_t^G)_{t\geq 0}$,

$$\forall\, \sigma \in \Omega\ \forall\, j \in V\colon\ \ \sigma \to \sigma^j \text{ at rate } \mathrm{e}^{-\beta[H(\sigma^j)-H(\sigma)]_+}$$

where σ^j is the configuration obtained from σ by flipping the spin at vertex j, and $\beta > 0$ is the *inverse temperature*.

3. The *Gibbs measure*

$$\mu(\sigma) = \frac{1}{\Xi}\,\mathrm{e}^{-\beta H(\sigma)}, \qquad \sigma \in \Omega,$$

is the reversible equilibrium of this dynamics.

4. Three sets of configurations play a central role (see Figure 2):

$$\begin{aligned} \mathbf{m} &= \text{metastable state}\\ \mathbf{c} &= \text{crossover state}\\ \mathbf{s} &= \text{stable state.} \end{aligned}$$

Definition. (a) The stable state is the set of configurations having minimal energy:

$$\mathbf{s} = \Big\{\sigma \in \Omega\colon\ H(\sigma) = \min_{\zeta\in\Omega} H(\zeta)\Big\}.$$

(b) The metastable state is the set of configurations not in $\mathbf{s}$ that lie at the bottom of the next deepest valley:

$$\mathbf{m} = \big\{\sigma \in \Omega \setminus \mathbf{s}\colon\ V_\sigma = \max_{\zeta\in\Omega\setminus\mathbf{s}} V_\zeta\big\}$$

with V_ζ the minimal amount a path from ζ needs to climb in energy in order to reach an energy $< H(\zeta)$. (c) The crossover state $\mathbf{c}$ is the set of configurations realising the min-max for paths connecting $\mathbf{m}$ and $\mathbf{s}$. ♠

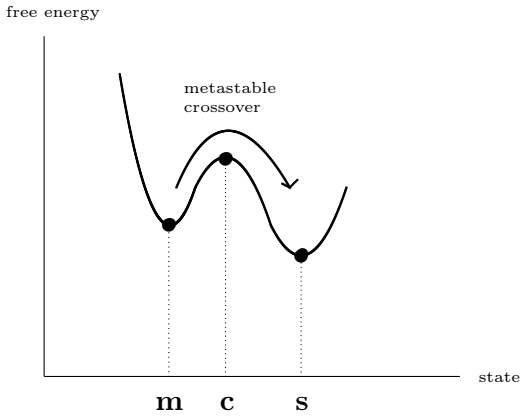

Figure 2: Caricature picture of the free energy landscape [free energy = energy − entropy].

2.2 SIM on the complete graph

Let us see what happens on the complete graph with N vertices. This is a mean-field setting.

The ferromagnetic interaction strength is chosen to be $J = N^{-1}$. It can be shown that the empirical magnetisation

$$m_t^N = \frac{1}{N} \sum_{i \in [N]} (\sigma_t^N)_i$$

performs a continuous-time random walk on the $2N^{-1}$-grid in $[-1, +1]$, in a potential that is given by the finite-volume free energy per vertex

$$f_{\beta,h}^N(m) = -\tfrac{1}{2}m^2 - hm + \beta^{-1} I^N(m)$$

with an entropy term

$$I^N(m) = -\frac{1}{N} \log \binom{N}{\frac{1+m}{2}N}.$$

In the limit $N \to \infty$, the empirical magnetisation performs a Brownian motion on $[-1, +1]$, in a potential that is given by the infinite-volume free energy per vertex

$$f_{\beta,h}(m) = -\tfrac{1}{2}m^2 - hm + \beta^{-1} I(m)$$

with

$$I(m) = \tfrac{1}{2}(1+m)\log(1+m) + \tfrac{1}{2}(1-m)\log(1-m),$$

where a redundant shift by $-\log 2$ is dropped. The above formulas describe what is called the *Curie-Weiss model with Glauber dynamics*.

Theorem 2.1 (Bovier, Eckhoff, Gayrard, Klein [10].)
If $\beta > 1$ and $h \in (0, \chi(\beta))$, then

$$\mathbb{E}^{\mathrm{CW}}_{\mathbf{m}_N^-}(\tau_{\mathbf{m}_N^+}) = K\, \mathrm{e}^{N\Gamma}[1 + o(1)], \qquad N \to \infty,$$

where $\mathbf{m}_N^-, \mathbf{m}_N^+$ are sets of configurations for which the discrete magnetisations tends to the continuum magnetisations m_-^*, m_+^*,

$$\Gamma = \beta\,[f_{\beta,h}(m^*) - f_{\beta,h}(m_-^*)]$$

$$K = \pi\beta^{-1} \sqrt{\frac{1+m^*}{1-m^*}\, \frac{1}{1-m_-^{*2}}\, \frac{1}{[-f_{\beta,h} =''(m^*)] f''_{\beta,h}(m_-^*)}}$$

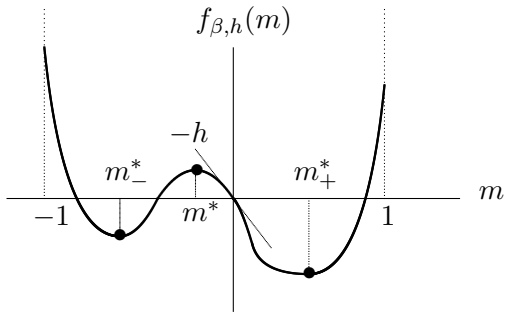

Figure 3: The free energy per vertex $f_{\beta,h}(m)$ at magnetisation m (caricature picture with $\mathbf{m} = m_-^*$, $\mathbf{c} = m^*$, $\mathbf{s} = m_+^*$).

and

$$\chi(\beta) = \sqrt{1 - \tfrac{1}{\beta}} - \tfrac{1}{2\beta} \log \left[\beta \left(1 + \sqrt{1 - \tfrac{1}{\beta}} \right)^2 \right].$$

The conditions on β, h guarantee that $f_{\beta,h}$ has a double-well shape (see Figure 3) and represents the parameter regime for which metastable behaviour occurs (see Figure 4).

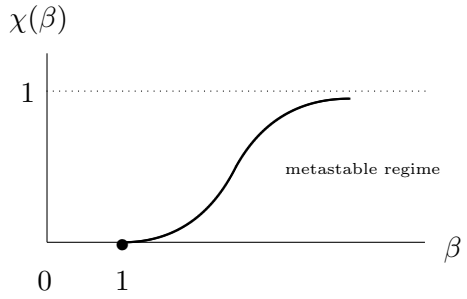

Figure 4: Metastable regime for the parameters β, h.

The expression for the average crossover time in Theorem 2.1 is called the *Kramers formula*.

2.3 SIM on random graphs

We want to investigate what can be said when the complete graph is replaced by a random graph. Our target will be to derive *Arrhenius laws*, i.e.,

$$\mathbb{E}_{\mathbf{m}}[\tau_{\mathbf{s}}] = K\, \mathrm{e}^{N\Gamma}[1 + o(1)], \qquad N \to \infty,\ \beta \text{ fixed},$$

$$\mathbb{E}_{\mathbf{m}}[\tau_{\mathbf{s}}] = K\, \mathrm{e}^{\beta\Gamma}[1 + o(1)], \qquad \beta \to \infty, N \text{ fixed}.$$

In general Γ, K are *random* and are hard to identify. In fact, in what follows we will mostly have to content ourselves with bounds on these quantities and with convergence in probability under the law of the random graph. In Sections 2.4–2.5 we focus on *dense* homogeneous and inhomogeneous Erdős-Rényi random graphs, and derive an Arrhenius law of the first type. In Sections 2.6–2.7 we focus on *sparse* graphs, both deterministic and random, and on configuration models, and derive an Arrhenius law of the second type.

2.4 SIM on the Erdős-Rényi random graph

Theorem 2.2 (den Hollander, Jovanovski [36]).
On the Erdős-Rényi random graph with N vertices, for $J = 1/pN$, $\beta > 1$ and $h \in (0, \chi(\beta))$,

$$\mathbb{E}^{\mathrm{ER}}_{\mathbf{m}^-_N}(\tau_{\mathbf{m}^+_N}) = N^{E_N}\, \mathbb{E}^{\mathrm{CW}}_{\mathbf{m}^-_N}(\tau_{\mathbf{m}^+_N}), \qquad N \to \infty,$$

where E_N is a random exponent that satisfies

$$\lim_{N\to\infty} \mathbb{P}_{\mathrm{ER_N(p)}}\big(|E_N| \le \tfrac{11}{6}\tfrac{\beta}{p}(m^* - m_-)\big) = 1,$$

with $\mathbb{P}_{\mathrm{ER_N(p)}}$ the law of the random graph.

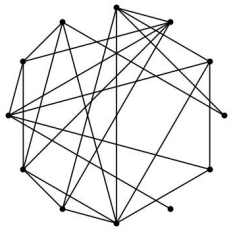

Figure 5: Erdős-Rényi random graph (ERRG): take the complete graph with N vertices and retain edges with probability $p \in (0, 1)$.

Apart from a polynomial error term, the crossover time is the same on the Erdős-Rényi random graph (see Figure 5) as on the complete graph, after the *change of interaction* from $J = 1/N$ to $J = 1/pN$. The asymptotic estimate of the crossover time is uniform in the starting configuration drawn from the set $\mathbf{m}^-_N$. Note that J needs to be *scaled* up by a factor $1/p$ in order to allow for a comparison with the Curie-Weiss model: in the Erdős-Rényi model every spin interacts with $\sim pN$ spins rather than N spins. The critical value in equilibrium changes from 1 to $1/p$ (Bovier, Gayrard [11]). On the complete graph the prefactor is constant and computable. On the Erdős-Rényi random graph it is random and more involved.

The proof of Theorem 2.2 follows the pathwise approach to metastability (see Bovier, den Hollander [12]). In particular, the empirical magnetisation $(m^N_t)_{t\ge 0}$ is monitored on a mesoscopic space-time scale. The difficulty is that the lumping technique typical for mean-field settings is no longer available: after projection the Markov property is lost. The way around this problem is via coupling: sandwich $(m^N_t)_{t\ge 0}$ between two Curie-Weiss models with a perturbed magnetic field h^N, tending to h as $N \to \infty$. The computations are rather elaborate and are beyond the scope of the present overview.

The following theorem provides a *refinement of the prefactor*.

Theorem 2.3 (Bovier, Marello, Pulvirenti [14]).
For $\beta > 1$, $h > 0$ small enough and $s > 0$,

$$\lim_{N\to\infty} \mathbb{P}_{\mathrm{ER}_N(p)}\left(C_1 \mathrm{e}^{-s} \le \frac{\mathbb{E}^{\mathrm{ER}}_{\mathbf{m}^-_N}(\tau_{\mathbf{m}^+_N})}{\mathbb{E}^{\mathrm{CW}}_{\mathbf{m}^-_N}(\tau_{\mathbf{m}^+_N})} \le C_2 \mathrm{e}^{s}\right) \ge 1 - k_1 \mathrm{e}^{-k_2 s^2},$$

where $k_1, k_2 > 0$ are absolute constants, and $C_1 = C_1(p, \beta)$ and $C_2 = C_2(p, \beta, h)$.

This theorem shows that the prefactor is a tight random variable, and hence constitutes a considerable sharpening of Theorem 2.2. The proof of Theorem 2.3 uses the potential-theoretic approach to metastability.

The local homogeneity of the Erdős-Rényi random graph again plays a crucial role: it turns out that the exact same test functions and test flows that are employed in relevant variational estimates work for the Curie-Weiss model and can be used to give sharp upper and lower bounds on the average crossover time. The better control on the prefactor comes at a price: the magnetic field has to be taken small enough; the dynamics starts from the last-exit biased distribution on $\mathbf{m}_N^-$ for the transition from $\mathbf{m}_N^-$ to $\mathbf{m}_N^+$, rather than from an arbitrary configuration in $\mathbf{m}_N^-$.

Proofs rely on elaborate techniques: isoperimetric inequalities, concentration estimates, capacity estimates, coupling techniques, coarse-graining techniques. These techniques exploit the fact that in the dense regime the Erdős-Rényi random graph is *locally homogeneous*.

2.5 SIM on the inhomogeneous Erdős-Rényi random graph

Theorem 2.3 can be extended to the inhomogeneous ERRG. The Hamiltonian becomes

$$H(\sigma) = -\sum_{(i,j)\in E} J_{ij}\sigma_i\sigma_j - h\sum_{i\in V}\sigma_i$$

with $J_{ij} > 0$ independent random variables. An example is Bernoulli with probability $r(\frac{i}{N}, \frac{j}{N})$, where

$$r(x,y), \qquad x,y \in [0,1],$$

is a continuous reference graphon. A special case is the rank-1 choice $r(x,y) = v(x)v(y)$ for some weight function $v(x)$, $x \in [0,1]$, which corresponds to the Chung-Lu Random Graph. See Bovier, den Hollander, Marello, Pulvirenti, Slowik [13] for further details.

2.6 SIM on sparse graphs

The ERRG is a dense graph. We next consider sparse graphs. Given a finite connected non-oriented multigraph

$$G = (V, E),$$

the Hamiltonian is

$$H(\sigma) = -\frac{J}{2}\sum_{(i,j)\in E}\sigma_i\sigma_j - \frac{h}{2}\sum_{i\in V}\sigma_i, \qquad \sigma \in \Omega,$$

where $J > 0$ is the ferromagnetic pair potential and $h > 0$ is the magnetic field. We write $\mathbb{P}_\sigma^{G,\beta}$ to denote the law of $(\sigma_t^G)_{t\geq 0}$ given $\sigma_0^G = \sigma$. The upper indices G, β exhibit the dependence on the underlying graph G and the interaction strength β between neighbouring spins.

It is easy to check that $\mathbf{s} = \{\boxplus\}$ for all G because $J, h > 0$. For general G, however, $\mathbf{m}$ is not a singleton, but we will be interested in those G for which the following hypothesis is satisfied (see Figure 6):

$$\text{(H)} \qquad \mathbf{m} = \{\boxminus\}.$$

The energy barrier between $\boxminus$ and $\boxplus$ is

$$\Gamma^\star = H(\mathcal{C}^\star) - H(\boxminus),$$

where $\mathcal{C}^\star = \mathbf{c}$ is the set of *critical configurations realising the min-max* for the crossover from $\boxminus$ to $\boxplus$, all of which have the same energy.

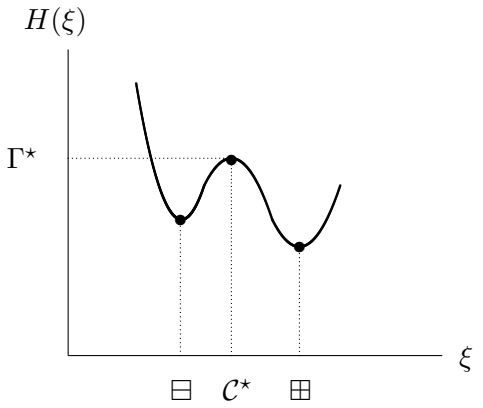

Figure 6: Schematic picture of H and $\boxminus, \boxplus$ and $\Gamma^\star, \mathcal{C}^\star$.

Theorem 2.4 (Bovier, den Hollander [12]).
There exists a $K^\star \in (0, \infty)$, called prefactor, such that

$$\lim_{\beta\to\infty} \mathrm{e}^{-\beta\Gamma^\star}\, \mathbb{E}^{G,\beta}_{\boxminus}(\tau_{\boxplus}) = K^\star.$$

The validity of Theorem 2.4 does not rely on the details of the graph G, provided it is finite, connected and non-oriented. For concrete choices of G, the task is to identify the *critical triplet* (see Figure 6)

$$(\mathcal{C}^\star, \Gamma^\star, K^\star).$$

For deterministic graphs this task has been successfully carried out for a large number of examples. However, for random graphs the triple is random, and identification represents a very serious challenge. In what follows we focus on the CM.

2.7 SIM on the configuration model

The CM is a sparse graph that can be generated via a simple pairing algorithm (see Figure 7).

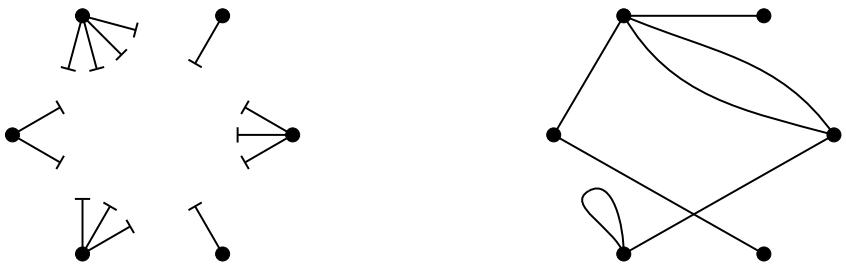

Figure 7: Configuration model with 6 vertices and degrees $(1, 3, 1, 3, 2, 4)$ after randomly pair half-edges.

WARNING: The text in the remainder of this section is technical. We go over it in leaps to sketch the main picture.

In order to state our main theorems, we need some notations and definitions.

1. Fix $N \in \mathbb{N}$. With each vertex $i \in [N]$ we associate a random degree d_i, in such a way that

$$(d_i)_{i\in[N]}$$

are i.i.d. with probability distribution f conditional on the event $\{\sum_{i\in[N]} d_i = \text{even}\}$. Consider a uniform matching of the half-edges, leading a multi-graph CM_N satisfying the requirement that the degree of vertex i is d_i for $i \in [N]$. The total number of edges is $\frac{1}{2}\sum_{i\in[N]} d_i$.

2. Throughout the sequel we write $\mathbb{P}_N$ to denote the law of the random multi-graph CM_N generated by the Configuration Model.

3. To avoid degeneracies we assume that

$$\begin{aligned} d_{\min} &= \min\{k \in \mathbb{N}\colon\ f(k) > 0\} \geq 3, \\ d_{\mathrm{ave}} &= \sum_{k\in\mathbb{N}} k f(k) < \infty, \end{aligned}$$

i.e., all degrees are at least three and the average degree is finite. In this case CM_N is connected with high probability (whp), i.e., with probability tending to 1 as$= N \to \infty$.

4. Along the way we need a technical function that allows us to quantify certain properties of the energy landscape, which we introduce next. Later we provide the underlying heuristics. For $x \in (0, \frac{1}{2}]$ and $\delta \in (1, \infty)$, define (see Figure 8)

$$\begin{aligned} I_\delta(x) = \inf\Big\{ y \in (0,x]\colon\ 1 < x^{x(1-1/\delta)}\,(1-x)^{(1-x)(1-1/\delta)} \\ \times (1-x-y)^{-(1-x-y)/2}\,(x-y)^{-(x-y)/2}\,y^{-y}\Big\}. \end{aligned}$$

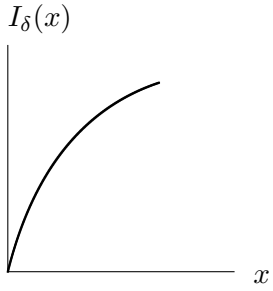

Figure 8: Plot of the function $x \mapsto I_\delta(x)$ for $\delta = 6$.

Main theorems. The following results are taken from Dommers, den Hollander, Jovanovski, Nardi [28]. We want to prove Hypothesis (H) and also to identify the critical triplet for CM_N, which we henceforth denote by $(\mathcal{C}_N^\star, \Gamma_N^\star, K_N^\star)$, in the limit as $N \to \infty$.

Our first theorem settles Hypothesis (H) for small h/J. Suppose that

$$\frac{h}{J} < \frac{2I_{d_{\mathrm{ave}}}\left(\frac{1}{2}\right) - \frac{1}{2}\left(1 - 4I_{d_{\min}}\left(\frac{1}{2}\right)\right)^2\left(1 - 2I_{d_{\min}}\left(\frac{1}{2}\right)\right)^{-1}}{\left(\frac{1}{d_{\mathrm{ave}}} + \frac{1}{2}\right)}.$$

Theorem 2.5 *If the above inequality is satisfied, then*

$$\lim_{N\to\infty} \mathbb{P}_N\big(\mathrm{CM}_N \text{ satisfies (H)}\big) = 1.$$

Our second and third theorem provide upper and lower bounds on $\Gamma_N^\star$. Label the vertices of the graph in order of increasing degree. Let $\gamma\colon \boxminus \to \boxplus$ be the path that successively flips the vertices $1,\dots,N$ (in that order), and for $M \in [N]$ let $\ell_M = \sum_{i\in[M]} d_i$.

Theorem 2.6 *Define*

$$\bar{M} = \bar{M}\left(\frac{h}{J}\right) = \min\left\{M \in [N]\colon\ \frac{h}{J} \geq \ell_{M+1}\left(1 - \frac{\ell_{M+1}}{\ell_N}\right) - \ell_M\left(1 - \frac{\ell_M}{\ell_N}\right)\right\},$$

and note that $\bar{M} < N/2$. Then with high probability

$$\Gamma_N^\star \leq \Gamma_N^+, \qquad \Gamma_N^+ = J\ell_{\bar{M}}\Big(1 - \frac{\ell_{\bar{M}}}{\ell_N}\Big) - h\bar{M} \pm O\big(\ell_N^{3/4}\big).$$

Theorem 2.7 *Define*

$$\tilde{M} = \min\left\{M \in [N]\colon\ \ell_M \geq \tfrac{1}{2}\ell_N\right\}.$$

Then with high probability

$$\Gamma_N^\star \geq \Gamma_N^-, \qquad \Gamma_N^- = J\, d_{\mathrm{ave}}\, I_{d_{\mathrm{ave}}}\left(\tfrac{1}{2}\right) N - h\tilde{M} - o(N).$$

Corollary 2.8 *Under Hypothesis* (H)*, Theorems 2.6–2.7 yield*

$$\lim_{\beta\to\infty} \mathbb{P}_{\boxminus}^{G,\beta}\left(\mathrm{e}^{\beta(\Gamma_N^- - \varepsilon)} \leq \tau_{\boxplus} \leq \mathrm{e}^{\beta(\Gamma_N^+ + \varepsilon)}\right) = 1 \quad \forall\, \varepsilon > 0.$$

REMARK: For simple degree distributions, like Dirac or power law, the quantities $\bar{M}$, $\ell_{\bar{M}}$, $\tilde{M}$ can be computed explicitly.

The bounds in Theorems 2.6–2.7 are tight in the limit of large degrees. Indeed, by the law of large numbers we have that

$$\ell_N \frac{\ell_{\bar{M}}}{\ell_N}\left(1 - \frac{\ell_{\bar{M}}}{\ell_N}\right) \leq \tfrac{1}{4}\ell_N = \tfrac{1}{4} d_{\mathrm{ave}}\, N\,[1 + o(1)]\,.$$

Hence

$$\frac{\Gamma_N^+}{\Gamma_N^-} = \frac{\frac{1}{4}d_{\mathrm{ave}}\,[1 + o\,(1)] - \frac{h}{J}\frac{\bar{M}}{N} + o(1)}{d_{\mathrm{ave}} I_{d_{\mathrm{ave}}}\left(\frac{1}{2}\right) - \frac{h}{J}\frac{\tilde{M}}{N} - o(1)}.$$

In the limit as $d_{\mathrm{ave}} \to \infty$ we have $I_{d_{\mathrm{ave}}}(\frac{1}{2}) \to \frac{1}{4}$, in which case the above ratio tends to 1.

Discussion.

1. The integer $\bar{M}$ has the following interpretation. The path $\gamma\colon \boxminus \to \boxplus$ is obtained by flipping (-1)-valued vertices to $(+1)$-valued vertices in order of increasing degree. Up to fluctuations of size $o(N)$, the energy along γ increases for the first $\bar{M}$ steps and decreases for the remaining $N - \bar{M}$ steps.

2. The integer $\tilde{M}$ has the following interpretation. To obtain our lower bound on $\Gamma_N^\star$ we consider configurations whose $(+1)$-valued vertices have total degree at most $\frac{1}{2}\ell_N$. The total number of $(+1)$-valued vertices in such type of configurations is at most $\tilde{M}$.

3. If we consider all sets on CM_N that are of total degree $x\ell_N$ and share $y\ell_N$ edges with their complement, then $I_\delta(x)$ represents (a lower bound on) the least value for y such that the average number of such sets is at least 1. In particular, for smaller values of y this average number is exponentially small.

4. We believe that Hypothesis (H) holds as soon as

$$0 < h < (d_{\min} - 1)J,$$

i.e., we believe that in the limit as $\beta \to \infty$ followed by $N \to \infty$ this choice of parameters corresponds to the metastable regime of our dynamics, i.e., the regime where $(\boxminus, \boxplus)$ is a metastable pair.

5. The scaling behaviour of $\Gamma_N^\star, K_N^\star$ as $N \to \infty$, as well as the geometry of $\mathcal{C}_N^\star$, are hard to capture. Here are some conjectures put forward in Dommers, den Hollander, Jovanovski, Nardi [28].

Conjecture 2.9 *There exists a $\gamma^\star \in (0, \infty)$ such that*

$$\lim_{N\to\infty} \mathbb{P}_N\Big(\big|N^{-1}\Gamma_N^\star - \gamma^\star\big| > \delta\Big) = 0 \qquad \forall\, \delta > 0.$$

Conjecture 2.10 *There exists a $c^\star \in (0, 1)$ such that*

$$\lim_{N\to\infty} \mathbb{P}_N\Big(\big|N^{-1}\log|\mathcal{C}_N^\star| - c^\star\big| > \delta\Big) = 0 \qquad \forall\, \delta > 0.$$

Conjecture 2.11 *There exists a $\kappa^\star \in (1, \infty)$ such that*

$$\lim_{N\to\infty} \mathbb{P}_N\Big(\big||\mathcal{C}_N^\star|\, K_N^\star - \kappa^\star\big| > \delta\Big) = 0 \qquad \forall\, \delta > 0.$$

6. It is shown in Dommers [27] that for a random regular graph with degree $r \geq 3$, there exist constants $0 < \gamma_-^\star(r) < \gamma_+^\star(r) < \infty$ such that

$$\lim_{N\to\infty} \lim_{\beta\to\infty} \mathbb{E}_N\left(\mathbb{P}_{\boxminus}^{\mathrm{CM}_N}\left(\mathrm{e}^{\beta N\gamma_-^\star(r)} \leq \tau_{\boxplus} \leq \mathrm{e}^{\beta N\gamma_+^\star(r)}\right)\right) = 1$$

when $\frac{h}{J} \in (0, C_0\sqrt{r})$ for some constant $C_0 \in (0, \infty)$ that is small enough. Moreover, there exist constants $C_1 \in (0, \frac{1}{4}\sqrt{3})$ and $C_2 \in (0, \infty)$ (depending on C_0) such that

$$\gamma_-^\star(r) \geq \tfrac{1}{4}Jr - C_1 J\sqrt{r}, \qquad \gamma_+^\star(r) \leq \tfrac{1}{4}Jr + C_2 J\sqrt{r}.$$

These results are derived without Hypothesis (H), but it is shown that Hypothesis (H) holds as soon as $r \geq 6$.

3 LECTURE 3: The Voter Model (VM)

In this lecture we focus on the VM on the *regular random graph*. We analyse how the fraction of discordant edges evolves over time, in the limit as the size of the graph tends to infinity, on *three time scales*: short, moderate, and long. We also analyse what happens when the edges of the random regular graph are *randomly rewired* while the VM is running. It will turn out that the graph dynamics has several interesting consequences. Most of what is written below is taken from Avena, Baldasso, Hazra, den Hollander, Quattropani [2, 3].

Given a connected graph $G = (V, E)$, the voter model is the Markov process $(\xi_t)_{t\geq 0}$ on state space $\{0,1\}^V$ where each vertex carries opinion 0 or 1, at rate 1 selects one of the neighbouring vertices uniformly at random, and adopts its opinion. Write $\xi_t = \{\xi_t(i)\colon\ i \in V\}$ with $\xi_t(i)$ the opinion at time t of vertex i. We analyse the evolution of the *fraction of discordant edges*

$$\mathcal{D}_t^N = \frac{|D_t^N|}{M}, \qquad D_t^N = \big\{(i,j) \in E\colon\ \xi_t(i) \neq \xi_t(j)\big\},$$

where $N = |V|$ and $M = |E|$. This is an interesting quantity because it monitors the *size of the boundary between the two opinions*.

The *consensus time* is defined as

$$\tau_{\text{cons}} = \inf\{t \geq 0\colon\ \xi_t(i) = \xi_t(j)\,\forall\, i,j \in V\}.$$

For finite graphs we know that $\tau_{\text{cons}} < \infty$ with probability 1, either at $[0]_N$ or at $[1]_N$. The interest lies in determining the *relevant time scale* on which consensus is reached, and *how* it is reached. Via time reversal, the voter model is *dual* to a system of coalescing random walks, describing the genealogy of the opinions, as shown in the next section. This section is an *intermezzo*, after which we will return to the random regular graph.

3.1 VM duality and graphical representation

In this section we give the *graphical representation* of the VM and indicate how *duality* is obtained via *time reversal*. One of the reasons why voter models have been studied intensively is the fact that they represent a class of interacting particle systems for which the dual is simple, namely, a *system of coalescing random walks*. This allows for a rephrasing of problems regarding voter models as problems regarding systems of coalescing random walks, which are often easier to deal with.

There are two ways to define a duality relation between two Markov processes $X = (X_t)_{t\geq 0}$ and $Y = (Y_t)_{t\geq 0}$ with state space $\mathcal{X}$ and $\mathcal{Y}$, respectively: *analytically* or *graphically*. The latter considers the time reversal of the graphical representation of the process and deduces information of the original process by using the evolution of the reversed process (which is often easier to study). The former needs the identification of a bounded measurable function $H\colon\ \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ such that

$$\mathbb{E}^x[H(X_t, y)] = \mathbb{E}^y[H(x, Y_t)], \qquad x \in \mathcal{X},\ y \in \mathcal{Y},\ t \geq 0.$$

If this is the case, then we say that X and Y are dual to one another with respect to the duality function H. Below we give a summary of both forms of duality. An important reference for interacting particle systems on graphs, particularly for the duality between the voter model and coalescing random walks on graphs, is [1, Chapter 14].

Definition of the general voter model. Let us first recall the definition of the voter model in terms of its transition rates. We consider the continuous-time voter model $(\eta_t)_{t\geq 0}$ with state space $\mathcal{X} = W^V$, where $V = [n] = \{1, \dots, n\}$, $n \in \mathbb{N}$, is the vertex set of a finite graph G, while W is a finite alphabet of admissible opinions with $|W| = o \in \{2, \dots, n\}$. We will be mainly interested in the case $o = 2$, with a given initial configuration of opinions $\eta_0 \in \mathcal{X}$, and in the case $o = n$, which corresponds to the setting in which every site has a different opinion. For each current state $\eta \in \mathcal{X}$, the only allowed transitions are the ones to the states $\eta^{x\leftarrow y} \in \mathcal{X}$ defined as

$$\eta^{x\leftarrow y}(z) = \begin{cases} \eta(y) & \text{if } z = x, \\ \eta(z) & \text{if } z \neq x, \end{cases}$$

for some $x, y \in [n]$ such that $\eta(y) \neq \eta(x)$. These transitions describe the events in which, starting from a configuration η, the voter at site x adopts the opinion of the voter at site y, while all other sites retain their opinions. We can write the generator L of the process as follows: for any $f \in D(\mathcal{X})$ (= the set of Lipschitz functions on $\mathcal{X}$) and $\eta \in \mathcal{X}$,

$$(Lf)(\eta) = \sum_{\eta' \in \mathcal{X}} \omega(\eta, \eta') \left[f(\eta') - f(\eta) \right] = \sum_{x \in [n]} \sum_{y \in [n]} c_y(x, \eta) \left[f(\eta^{x\leftarrow y}) - f(\eta) \right],$$

where the values $\omega(\eta, \eta') \geq 0$ represent the rates at which transitions $\eta \to \eta'$ occur for $\eta, \eta' \in \mathcal{X}$. The only possible transitions are those of the type $\eta \to \eta^{x\leftarrow y}$ for some sites x, y with different opinions.

Thus, we renamed the sum over the configurations η' of $\omega(\eta, \eta^{x\leftarrow y})$ as the sum over the sites and opinions of the rates $c_y(x, \eta)$, emphasising the dependence on the parameters x, y, η. Since we want that, for every pair of sites x, y, transitions $\eta \to \eta^{x\leftarrow y}$ occur with a rate proportional to the number of neighbours y of x having opinion $\eta(y)$, we set

$$c_y(x, \eta) = \sum_{z \in [n]} p(x, z)\, 1_{\{\eta(z) = \eta(y), \eta(x) \neq \eta(y)\}},$$

where $p(\cdot, \cdot)$ are the jump probabilities of an irreducible continuous-time random walk on $[n]$, in particular, $p(x, y) \geq 0$ for all $x, y \in [n]$ and $\sum_{y \in [n]} p(x, y) = 1$ for all $x \in [n]$. In other words, a site x waits an exponential time of parameter 1, after which it adopts the value of the voter seen at a site z that is chosen with probability $p(x, z)$. It follows immediately that the underlying random walk structure given by $p(\cdot, \cdot)$ uniquely determines the process. We mention that in the case where $W = \{0, 1\}$, due to the invariance under relabelling of the two opinions, the rates can be written as $c(x, \eta) = \sum_{y:\ \eta(x) \neq \eta(y)} p(x, y)$, and represent the rate of the transition $\eta \to \eta_x$, where $\eta_x(z) = 1 - \eta(x)$ if $z = x$ and $\eta_x(z) = \eta(z)$ otherwise.

Analitic duality. In order to state the duality relation analytically, let us for the moment consider the case in which $W = \{0, 1\}$, i.e., the voter model over $X = W^{[n]}$ is a spin-flip system. Once the duality relation is stated in this setting, the natural generalisation to a finite number of opinions o will be straightforward. For a detailed discussion of duality for spin-flip systems, we refer the reader to Liggett [38, 39]. Let us denote by $(A_t)_{t\geq 0}$ the dual Markov process with respect to the voter model $(\eta_t)_{t\geq 0}$ with state space X. Since we are considering a spin-flip system, the state space of the dual $(A_t)_{t\geq 0}$ is taken to be

$$Y = \{A\colon\ A \text{ finite subset of } [n]\},$$

which is finite for every $n \in \mathbb{N}$. Therefore the dual process $(A_t)_{t\geq 0}$ is actually a Markov chain on Y, and can be interpreted as the locations in time of a collection of independent continuous-time random walks on $[n]$ that coalesce every time two of them meet at the same site. It follows that $|A_t|$ can only decrease as t increases. The duality function H that will take this into account is

$$H(\eta, A) = 1_{\{\eta(x)=1 \ \forall x\in A\}}, \qquad \eta \in X,\ A \in Y.$$

The chief reason for this choice is related to the fact that we want to look at consensus states, i.e., η's such that $\eta(x) = \eta(y)$ for all $x, y \in [n]$. It is not difficult to see that, by using the latter duality function in combination with the generator and the coefficients for $k = 2$, we get that the rates of $(A_t)_{t\geq 0}$ for the transitions $A \to B$, $A, B \in Y$, are

$$q(A, B) = \sum_{x\in A} \sum_{\substack{y\in[n]:\\ (A\setminus\{x\})\cup\{y\}=B}} p(x, y).$$

The interpretation is the following. Each $x \in A$ is removed from A at rate 1 and is replaced by y with probability $p(x, y)$. Moreover, when an attempt is made to place a point y at a site that is already occupied, then the two points coalesce into one. In our system of coalescing random walks, this means that each random walk independently has an exponential waiting time of parameter 1, moves according to $p(x, y)$, and when moving to an occupied site coalesces with the occupier. Thus, the semigroup of each random walk is given by

$$P^t(x, y) = \mathrm{e}^{-t} \sum_{n\in\mathbb{N}_0} \frac{t^n}{n!}\, p^n(x, y), \qquad x, y \in [n], t \geq 0.$$

All the previous statements can be generalised to a setting with k opinions with $k \in \{2, \dots, n\}$ in the following way. The dual state space is taken to be $Y^{(k)} = Y^{k-1}$, the product of $k-1$ copies of Y. Consequently, the dual process is of the form $A_t^{(k)} = (A_1, \dots, A_{k-1})_t$. A good reference for an analytical description of duality for interacting particle systems is López and Sanz [42]. Here, the interpretation is slightly different from the case $k = 2$. In the latter, $A \in Y$ represents the set of sites at which there were particles, while $A^{(k)} = (A_1, \dots, A_{k-1}) \in Y^{(k)}$ represents the positions A_i of the particles that trace back to opinion i (this will become clear in the graphical representation), possibly with $A_i = \emptyset$ for some $i \in \{1, \dots, k-1\}$. The dual function $H\colon\ X \times Y^{(k)} \to \mathbb{R}$ now reads

$$H(\eta, A^{(k)}) = 1_{\{\eta(x)=i \ \forall\, x\in A_i \ \forall\, i\in\{1,\dots,k-1\}\}},$$

and the duality relation becomes

$$\mathbb{P}^{\eta}\big(\eta_t(x) = i \ \forall\, x \in A_i \ \forall\, i \in \{1, \dots, k-1\}\big) = \mathbb{P}^{A^{(k)}}\big(\eta(x_t) = i \ \forall\, x_t \in (A_i)_t \ \forall\, i \in \{1, \dots, k-1\}\big)$$

for every initial configuration $\eta \in X$ and every initial state $A^{(k)} \in Y^{(k)}$.

Let us consider the case in which $k = n$, and take $A^{(k)}$ such that $A_i = \{i\}$ for all sites, i.e., we place a continuous-time random walk at every site. Fix $\eta = \eta_0 \in X$ to be the configuration in which each site has its personal opinion, say vertex $i \in [n]$ has opinion i, for which $|[n]| = |W|$. If one considers the configuration $\hat{\eta}_t$ defined as

$$\hat{\eta}_t(i) = \eta_0((A_i)_t), \qquad i \in [n],\ t > 0,$$

then the duality relation says that $\hat{\eta}_t$ has the same distribution as the state η_t of the voter model at time t with initial configuration η_0. Define the *consensus time* as

$$\tau_{\text{cons}} = \inf\{t \geq 0 \colon\ \eta_t(x) = \eta_t(y)\ \forall\, x, y \in V\}$$

and the *coalescence time* as

$$\tau_{\text{coal,n}} = \inf\{t \geq 0 \colon\ \text{ all } n \text{ particles have coalesced into one}\}.$$

In particular, it follows that the consensus time has the same law of the coalescence time, i.e.,

$$\mathbb{E}[\tau_{\text{cons}}] = \mathbb{E}[\tau_{\text{coal},n}].$$

It follows that if we consider the same model with $2 \leq k < n$ opinions and a given initial configuration η_0 of them among $[n]$, then the distribution of $\tau_{\text{cons,k}}$ will be stochastically dominated by the distribution of $\tau_{\text{coal},n}$ because $\{\tau_{\text{cons,k}} > t\} \supseteq \{\tau_{\text{cons}} > t\}$, so $\mathbb{P}^{\eta_0}(\tau_{\text{cons,k}} \leq t) \leq \mathbb{P}(\tau_{\text{cons}} \leq t)$ for all $t \geq 0$. (With $\tau_{\text{cons,k}}$ we mean the consensus time of the voter model with k opinions, with the convention that $\tau_{\text{cons,n}} = \tau_{\text{cons}}$.) Thus,

$$\mathbb{E}^{\eta_0}[\tau_{\text{cons,k}}] \leq \mathbb{E}[\tau_{\text{coal},n}], \qquad 2 \leq k < n.$$

Moreover, it can be proved that, in the case where $k = 2$ and the initial distribution μ_u is given by the product measure of parameter $u \in (0, 1)$ Bernoulli random variables,

$$2u(1-u)\,\mathbb{E}[\tau_{\text{coal},n}] \leq \mathbb{E}^{\mu_u}[\tau_{\text{cons},2}] \leq \mathbb{E}[\tau_{\text{coal},n}], \qquad u \in (0, 1).$$

Graphical representation of the dual process. We conclude this section by giving the duality principle using the graphical representation. Start with the same setting as above: voter model with state space $X = W^{[n]}$, $|W| = k$, defined by its generator and its rates. Consider the graph $\{(j, t) \colon\ j \in [n], t \geq 0\}$ and independent rate-1 Poisson processes $(N_i(t))_{t \geq 0}$, $i \in [n]$. The dynamics is the following: if $\bar{t}$ is an event of the clock N_i for some $i \in [n]$, then draw an arrow from $(\bar{t}, j)$ to $(\bar{t}, i)$, where $j \in [n]$ is chosen with probability $p(i, j)$. These transition probabilities coincide with the ones given above. In other words, an event represented by an arrow $j \to i$ means that at time $\bar{t}$ the voter at site i decides to adopt the opinion of the voter at site j.

Given any initial configuration $\eta_0 \in X$, we let the opinions flow upwards, starting at time $t = 0$, and any time they encounter the base of an arrow they follow its direction, changing the opinion that is at the tip of the arrow. In the case of two-opinions (0 and 1), this construction can be seen as a percolation process where a fluid is placed at $t = 0$ in the 1-sites of η_0 and flowing up the structure: the arrows are the pipes and the tips are the dams (see Durrett [30]). An example of this graphical representation is given in Figure 9 for the 8-cycle graph with $k = n$.

Let us now fix a time horizon $t_0 > 0$ and position a walk in (i, t_0) for all sites $i \in [n]$. We let these walks evolve independently as follows: they move downwards through the graph $\{(j, t_0 - t) \colon\ j \in [n], t \in [0, t_0]\}$, and any time they encounter the tip of an arrow they follow it in the opposite direction. Furthermore, if one of them moves to a site already occupied by another walk, then the two walks coalesce into a single (independent) one. Alternatively, the process can be described as follows: each of the walks waits an exponential time of parameter 1 and, given the current position $x \in [n]$, moves to $y \in [n]$ with probability $p(x, y)$. The

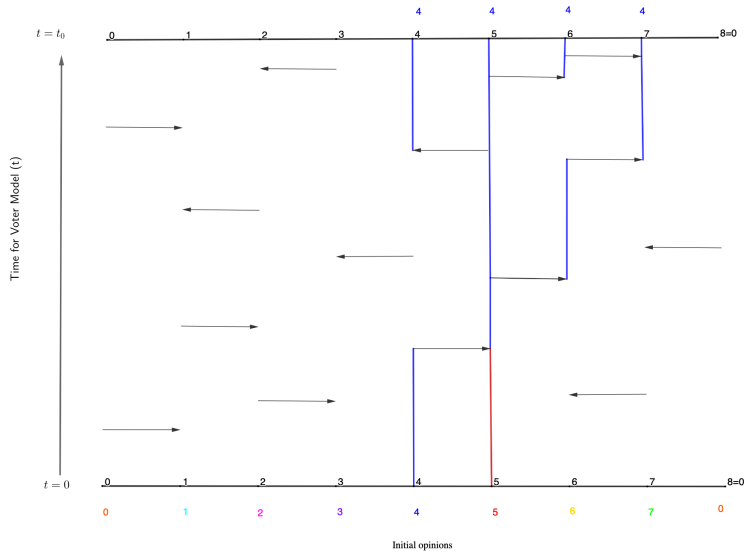

Figure 9: Realisation of the voter model on the 8-cycle graph where the initial configuration is given by a different opinion for each site: site $i \in [n]$ has opinion i. Times flows vertically up to a finite horizon $t_0 > 0$, and the black horizontal arrows represent the events of the independent Poisson processes $(N_i(t))_{t\geq 0}$ described above. In this picture we highlight the evolution of opinion 4 that at time t_0 is shared among the sites $\{4, 5, 6, 7\}$.

same for the coalescing condition, i.e., each time at least two walks meet at the same site they coalesce into a single partcile. Denote now by $A_t^{(k),t_0} = (A_1^{t_0}, \ldots, A_n^{t_0})_t$ the resulting system of n coalescing random walks (CRWs) evolving as above, where, for each $i \in [n]$, $(A_i)_t^{t_0}$ is the position of the walk starting in (i, t_0) at time t, in particular, $A_0^{(k),t_0} = [n]$. Following the previous example, Figure 10 gives a realisation of the CRWs system in the same 8-cycle graph shown in Figure 9.

Given this construction it is immediate to see that the opinion held by a vertex i at time t_0 can be derived by *tracing back in time* the path of the random walkers up to time $t = 0$. Thus,

$$\eta_{t_0}(i) = \eta_0((A_i)_{t_0}^{t_0}) \qquad \forall\, i \in [n],\, \forall\, t_0 > 0.$$

With the latter we derived the same result obtained using the duality relation. All the other results regarding the equivalence in distribution between the coalescing time and the consensus time with $k = n$ opinions, i.e. $\eta_0 = [n]$, follow directly.

3.2 VM on the complete graph

As a prelude we look at the VM on the complete graph, for which computations can be carried through explicitly. Indeed, the number of 1-opinions at time t, given by

$$O_t^N = \sum_{i\in V} \xi_t(i),$$

performs a continuous-time nearest-neighbour random walk on the set $\{0, \ldots, N\}$ with transition rates

$$n \to n+1 \quad \text{at rate } n(N-n)\tfrac{1}{N-1},$$

$$n \to n-1 \quad \text{at rate } (N-n)n\tfrac{1}{N-1}.$$

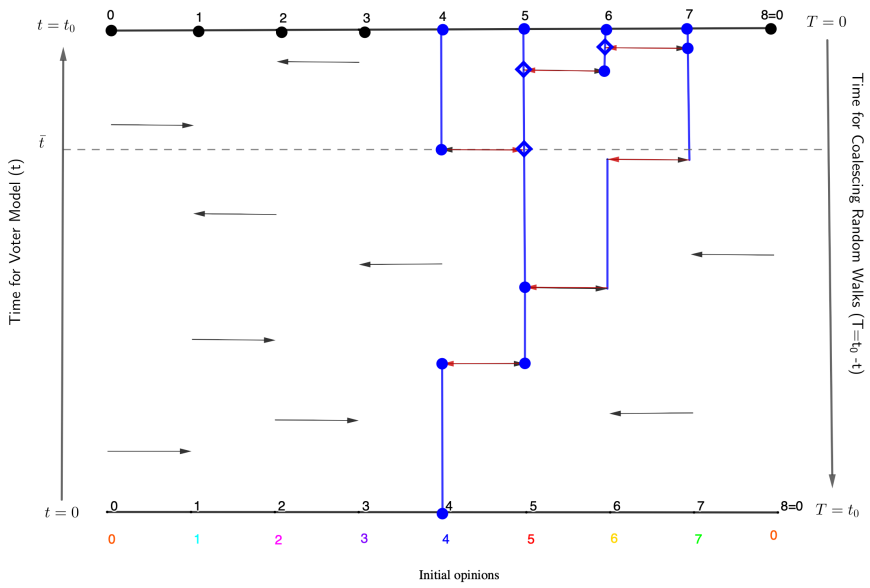

Figure 10: Realisation of a coalescing random walk system, starting at $t = t_0$, as dual process of the voter model in Figure 9. In the figure the large dots in the graph at t_0 represent the initial positions of the walkers, the blue dots are the ones for which we derived the whole trajectory up to time $t = 0$, and coincide with the sites in which we had a local consensus with respect to opinion 4 in the voter model. The blue dots in all the other positions describe the evolution of the position of the (blue) walks, while the blue diamonds $\diamond$ mean that in such a time-space location a clustering occurred, i.e., two walks coalesced into one. Moreover, the red horizontal arrows shows the direction to be followed by the walks, which are the opposite w.r.t. the direction of the original (black) arrows of the voter model. Note that by time $\bar{t}$ all walks started in $\{4, 5, 6, 7\}$ coalesced into one, meaning that the opinion of each vertex in the latter set at time $t = t_0$ is the same of the one held by vertex 4 at time $t = 0$.

This is the same as the Moran model from population genetics. Put $\mathcal{O}_t^N = \frac{1}{N} O_t^N$ for the *fraction of* 1*-opinions at time* t. The following is a well-known fact.

Lemma 3.1 *The process*

$$(\mathcal{O}_{sN}^N)_{s\geq 0}$$

converges in law as $N \to \infty$ *to the* Fisher-Wright diffusion $(\chi_s)_{s\geq 0}$ *on* $[0,1]$ *given by*

$$\mathrm{d}\chi_s = \sqrt{2\chi_s(1-\chi_s)}\,\mathrm{d}W_s,$$

where $(W_s)_{s\geq 0}$ *is standard Brownian motion.*

The number of discordant edges equals

$$D_t^N = \frac{O_t^N(N - O_t^N)}{2}.$$

Recall that $\mathcal{D}_t^N = \frac{1}{M} D_t^N$ denotes the *fraction of discordant edges at time* t, with $M = \binom{N}{2}$ for the complete graph. Since

$$\mathcal{D}_t^N = \frac{O_t^N(N - O_t^N)}{N(N-1)} = \frac{N}{N-1}\,\mathcal{O}_t^N(1 - \mathcal{O}_t^N),$$

it follows that

$$(\mathcal{D}_{sN}^N)_{s\geq 0}$$

converges in law as $N \to \infty$ to the process

$$\big(\chi_s(1-\chi_s)\big)_{s\geq 0}.$$

In the mean-field setting of the complete graph, the fraction of discordant edges is the *product* of the fractions of the two opinions. The latter property *fails* on non-complete graphs, in particular, on random graphs.

We close this section with a proof of Lemma 3.1. The proof is instructive because it shows how computations with generators can be useful.

Proof. It is enough to prove convergence of generators (see Ethier, Kurtz [33, Chapter 10]). The rescaled voter model on the complete graph is the birth-death process with state space $\{0, \frac{1}{N}, \ldots, 1 - \frac{1}{N}, 1\}$ and infinitesimal generator $\hat{L}_N$ given by

$$(\hat{L}_N f)\left(\tfrac{i}{N}\right) = N\,\tfrac{i}{N}\,(N-1)\big[f\left(\tfrac{i+1}{N}\right) + f\left(\tfrac{i-1}{N}\right) - 2f\left(\tfrac{i}{N}\right)\big], \quad i \in \{0, \ldots, N\}.$$

This can be seen by conditioning on the number of steps by the non-rescaled process and using the definition of generator. We have to prove that $\hat{L}_N$ converges as $N \to \infty$ to the generator of the FW-diffusion given by

$$(Lf)(y) = y(1-y)\,f''(y), \quad y \in [0,1],$$

i.e.,

$$\lim_{N\to\infty} (\hat{L}_N f)\left(\tfrac{i}{N}\right) = (Lf)(y) \quad \text{when} \quad \lim_{N\to\infty} \frac{i}{N} = \lim_{N\to\infty} \frac{i_N}{N} = y.$$

We have to specify which set of test functions f we consider. Let $C([0,1])$ be the set of $\mathbb{R}$-valued continuous functions on the unit interval, and define

$$C_0([0,1]) = \{f \in C([0,1]) \colon\ f(0) = f(1) = 0\}.$$

Since the 'local speed' of the FW-diffusion is given by the diffusion function $g_{FW} \colon\ [0,1] \to [0,\infty)$ with $g_{FW}(y) = y(1-y)$, it is possible to show that the domain $D(L)$ of the generator L of the FW-diffusion is a subset of $C_0([0,1])$. For general Markov processes it is not easy to characterise $D(L)$, but it suffices to consider the action of L on a subset $K(L) \subset D(L)$ that is large enough to maintain the generality of the argument. We refer to $K(L)$ as a core of the generator and require that $K(L)$ is dense in $C_0([0,1])$ with respect to the supremum norm. Then the required generality is obtained via continuous extension. In our case it suffices to choose

$$K(L) = \{f \in C_0([0,1]) \colon\ f \text{ is infinitely differentiable}\}.$$

This choice enables us to use the Taylor expansion of any test function up to any order. Using the Taylor expansion of f around i/N up to second order in the generator, we get

$$\begin{aligned}
&f\left(\tfrac{i+1}{N}\right) + f\left(\tfrac{i-1}{N}\right) - 2f\left(\tfrac{i}{N}\right) \\
&= \left(\tfrac{i+1}{N} - \tfrac{i}{N}\right) f'\left(\tfrac{i}{N}\right) + \left(\tfrac{i-1}{N} - \tfrac{i}{N}\right) f'\left(\tfrac{i}{N}\right) + N^{-2} f''\left(\tfrac{i}{N}\right) + \mathcal{O}(N^{-3}) \\
&= N^{-2} f''\left(\tfrac{i}{N}\right) + \mathcal{O}(N^{-3}).
\end{aligned}$$

Therefore

$$(\hat{L}_N f)\left(\tfrac{i}{N}\right) = N^2\, \tfrac{i}{N}\, \left(1 - \tfrac{i}{N}\right) \left[N^{-2} f''\left(\tfrac{i}{N}\right) + \mathcal{O}(N^{-3})\right], \quad i \in \{0, \dots, N\}.$$

Hence, when $\lim_{N\to\infty} \frac{i}{N} = \lim_{N\to\infty} \frac{i_N}{N} = y$, thanks to the continuity of $f \in K(L)$, we get

$$\lim_{N\to\infty} (\hat{L}_N f)\left(\tfrac{i}{N}\right) = (Lf)(y).$$

■

3.3 VM on the random regular graph

Consider the regular random graph $G_{d,N} = (V, E)$ of degree $d \geq 3$, consisting of

$$|V| = N \text{ vertices},$$

$$|E| = M = \frac{dN}{2} \text{ edges}.$$

Denote the law of $G_{d,N}$ by $\mathbb{P}$. Chen, Choi, Cox [26] consider the fraction of 1-opinions at time t,

$$\mathcal{O}_t^N = \frac{1}{N} \sum_{i \in V} \xi_t(i),$$

and show that

$$(\mathcal{O}_{sN}^N)_{s \geq 0}$$

converges in law as $N \to \infty$ to the Fisher-Wright diffusion $(\chi_s)_{s\geq 0}$ given by

$$\mathrm{d}\chi_s = \sqrt{2\theta_d \chi_s(1-\chi_s)}\,\mathrm{d}W_s,$$

where $(W_s)_{s\geq 0}$ is standard Brownian motion, and

$$\theta_d = \frac{d-2}{d-1}.$$

plays the role of a *diffusion constant*.

Main theorems. For $u \in (0,1)$, let $\mathbf{P}_u$ be the law of $(\mathcal{D}_t^N)_{t\geq 0}$ starting from $[\text{Bernoulli}(u)]^N$. For $\eta \in \{0,1\}^V$, let $\mathbf{P}$ be the law of $(\mathcal{D}_t^N)_{t\geq 0}$ starting from η.

Theorem 3.2 *Fix $u \in (0,1)$. Then, for any $t_N \in [0,\infty)$,*

$$\left| \mathbf{E}_u\left[\mathcal{D}_{t_N}^N\right] - 2u(1-u)\, f_d(t_N)\, \mathrm{e}^{-2\theta_d \frac{t_N}{N}} \right| \xrightarrow{\mathbb{P}} 0,$$

where

$$f_d(t) = \mathbf{P}^{\mathcal{T}_d}(\tau_{\text{meet}} > t),$$

with $\mathbf{P}^{\mathcal{T}_d}$ the law of two independent random walks on the infinite d-regular tree $\mathcal{T}_d$ starting from the endpoints of an edge.

The profile function f_d is given by

$$f_d(t) = \sum_{k=0}^{\infty} \mathrm{e}^{-2t} \frac{(2t)^k}{k!} \sum_{l > \lfloor \frac{k-1}{2} \rfloor} \binom{2l}{l} \frac{1}{l+1} \left(\frac{1}{d}\right)^{l+1} \left(\frac{d-1}{d}\right)^{l},$$

and satisfies $f_d(0) = 1$ and $f_d(\infty) = \theta_d$. Note that Theorem 3.2 shows three times scales (see Figures 11–12):

short:	$t_N \ll N$,
moderate:	$t_N \asymp N$,
long:	$t_N \gg N$.

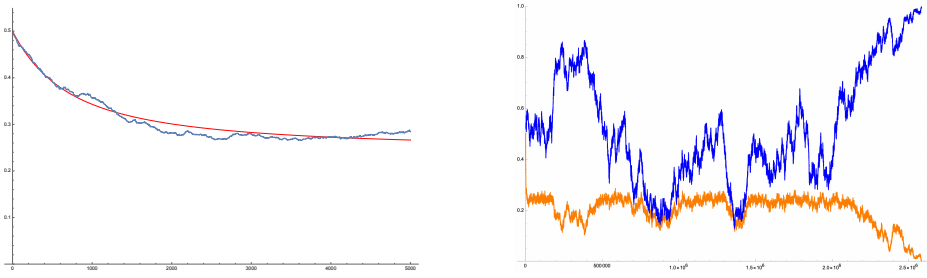

Figure 11: A single simulation for $N = 1000$, $d = 3$, $u = 0.5$. *Left*: In blue the fraction of discordant edges up to $t = 5$, in red the function $t \mapsto 2u(1-u)\, f_d(t)$. *Right*: In blue the fraction of 1-opinions up to consensus, in orange the fraction of discordant edges up to consensus.

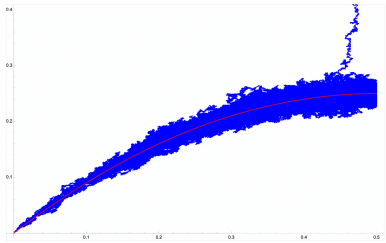

Figure 12: Scatter plot for the same simulation: the fraction of discordant edges versus the fraction of the minority opinion. The piece sticking out corresponds to short times. The curve in red is $x \mapsto x(1-x)$, which says that the fraction of discordant edges is close to the product of the fractions of the two opinions.

Theorem 3.3 *The following concentration properties hold:*

(i) Let t_N be such that $t_N/N \to 0$. Then, for every $\varepsilon > 0$,

$$\sup_{\eta \in \{0,1\}^V} \mathbf{P}_\eta \left(\left| \mathcal{D}_{t_N}^N - \mathbf{E}_\eta[\mathcal{D}_{t_N}^N] \right| > \varepsilon \right) \xrightarrow{\mathbb{P}} 0.$$

(ii) Let t_N be such that $t_N/N \to s \in (0,\infty)$. Then, for every $u \in (0,1)$,

$$\sup_{x \in [0,1]} \left| \mathbf{P}_u \left(\mathcal{D}_{t_N}^N \le x \right) - \mathbf{P}_u \left(\chi_s(1-\chi_s) \le x \right) \right| \xrightarrow{\mathbb{P}} 0.$$

Theorem 3.4 *Fix $u \in (0,1)$. Then, for every $\delta, \epsilon > 0$,*

$$\mathbf{P}_u \Big(\sup_{0 \le t \le N^{1-\delta}} \left| \mathcal{D}_t^N - \mathbf{E}_u[\mathcal{D}_t^N] \right| > \varepsilon \Big) \xrightarrow{\mathbb{P}} 0.$$

The proofs of Theorems 3.2–3.4 are based on the classical notion of *duality* between the voter model and a collection of coalescent random walks. A crucial role is played by properties of coalescing random walks that hold in *mean-field geometries*. In particular, in Oliveira [48] it is shown that

$$\lim_{N \to \infty} \frac{\mathbf{E}[\tau_{\text{coal}}]}{\mathbf{E}[\tau_{\text{meet}}^{\pi \otimes \pi}]} = 2,$$

where τ_{coal} is the coalescence time of N random walks each starting from a different vertex, $\mathbf{E}$ is expectation w.r.t. these random walks, while $\tau_{\text{meet}}^{\pi \otimes \pi}$ is the meeting time of two random walks independently starting from the stationary distribution π. A further discussion of this result is postponed to Section 3.4.

1. On time scales $o(\log N)$, below the typical distance between two vertices, the analysis proceeds by coupling two random walks on the d-regular random graph with two random walks on the d-regular tree, both starting from adjacent vertices. Because the tree is regular, the distance of the two random walks can be viewed as the distance to the origin of a single biased random walk on $\mathbb{N}_0$ starting from 1. Note that the same does not hold when the tree is not regular.

2. On time scale $\Theta(\log N)$, the scale of the typical distance between two vertices, the coupling argument is combined with a finer control of the two random walks on the d-regular random graph.

Lemma 3.5 *There is a sequence of random variables $(\theta_{d,N})_{N\in\mathbb{N}}$ converging to θ_d such that*

$$\lim_{N\to\infty} \sup_{t\geq 0} \left| \frac{\mathbf{P}(\tau_{\text{meet}}^{\pi\otimes\pi} > t)}{\exp[-2\theta_{d,N}(t/N)]} - 1 \right| = 0 \quad \textit{in probability.}$$

Lemma 3.5, together with a first-moment argument, is enough to compute the evolution of the expected number of discordant edges on every time scale.

3. In order to obtain concentration, a much deeper analysis is required. Roughly, in order to have proper control on the correlations between the discordant edges, we must analyse a dual system of random walks whose number grows with N. An upper bound is derived for the number of meetings of a poly-logarithmic number of independent random walks evolving on the random graph for a time $N^{1-o(1)}$. This is exploited to derive an upper bound for the deviation from the mean that is exponentially small in N and uniform in time. This upper bound can be translated into a concentration estimate by taking a union bound.

OPEN PROBLEMS:

- We expect that Theorems 3.2–3.3 can be extended to non-regular sparse random graphs. We do not have a conjecture on how the function f_d and the diffusion constant θ_d modify in this more general setting.
- We expect that Theorem 3.3 can be strengthened to the statement that, for every $u \in (0,1)$, every t_N such that $t_N/N \to 0$ and every $C_N \to \infty$,

$$\mathbf{P}_u\big(\left|\mathcal{D}_{t_N}^n - \mathbf{E}_u[\mathcal{D}_{t_N}^N]\right| > C_N\sqrt{t_N/N}\,\big) \xrightarrow{\mathbb{P}} 0.$$

Directed graphs. For directed sparse random graphs more can be said. The setting is the *directed configuration model* with prescribed in-degees $d^{\text{in}} = (d_i^{\text{in}})_{i=1}^N$ and out-degees $d^{\text{out}} = (d_i^{\text{out}})_{i=1}^N$ in which directed half-edges are matched randomly.

Theorem 3.6 (Avena, Capannoli, Hazra, Quattropani [4], Capannoli [18]).
Under mild conditions on the in-degrees d^{in} and the out-degrees d^{out}, the same scaling applies and an explicit formula can be derived both for the profile function $f_{d^{\text{in}},d^{\text{out}}}$ and for the diffusion constant $\theta_{d^{\text{in}},d^{\text{out}}}$.

For instance, if $d^{\text{in}} = d^{\text{out}}$ (= Eulerian graph), then

$$\theta_{d^{\text{in}},d^{\text{out}}} = \left(\tfrac{m_2}{m_1^2} - 1 + \sqrt{1-\tfrac{1}{m_1}}\right)^{-1}$$

with m_1, m_2 the first and the second moment of the limit of the empirical degree distribution.

From the fact that $\theta_{d^{\text{in}},d^{\text{out}}}$ is an explicit function of d^{in} and d^{out}, it is possible to analyse its behaviour as a function of both. Relevant questions are: Does the consensus time speed up or slow down when, in the setting of constant out-degrees (or in-degrees), the variability of the in-degrees (or out-degrees) is increased? What is the effect of positive or negative correlation between the in-degrees and out-degrees? If certain features of the degree sequences are constrained, then what are the guiding principles to minimise or maximise the consensus time?

3.4 VM meeting times and coalescence times

Discordance and meeting times. As show in Section 3.1, the graphical representation allows us to write the probability that a fixed edge is discordant at time t in terms of the meeting time of random walks. Let $D_t = \{e = (x, y) \in E\colon\ \eta_t(x) \neq \eta_t(y)\}$ be the set of discordant edges at time t for the voter model with initial distribution of opinions given by the product measure of independent Bern(u), $u \in (0, 1)$. Moreover, define

$$\tau_{\text{meet}}^{(x,y)} = \inf\{t \geq 0\colon\ X_t^x = Y_t^y\}$$

to be the first meeting time of two independent random walks X, Y having initial position $(X_0, Y_0) = (x, y)$, $x, y \in V$. Then, by duality,

$$\begin{aligned}\mathbf{P}_u(e \in D_t) &= \mathbf{P}\big(X_s^x \neq Y_s^y\ \forall\, 0 \leq s \leq t\big) \times \mathbf{P}_u\big(\eta_0(X_t^x) \neq \eta_0(Y_t^y)\big)\\ &= \mathbf{P}\big(\tau_{\text{meet}}^{(x,y)} > t\big) \times 2u(1-u).\end{aligned}$$

(See Figure 13.)

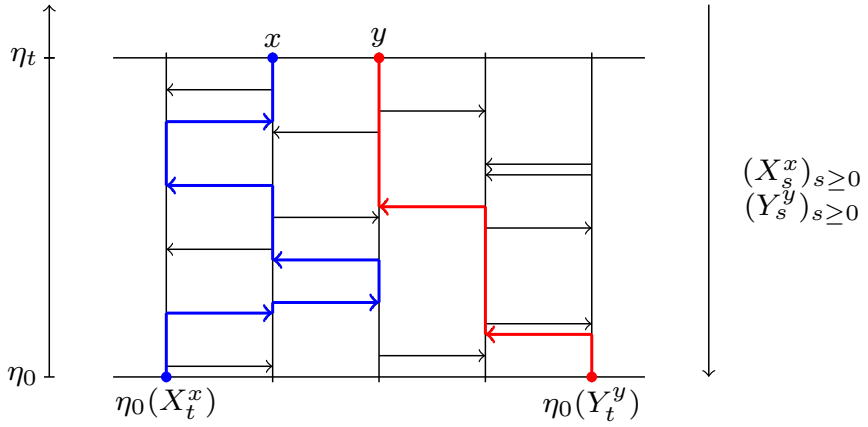

Figure 13: Discordance versus meeting.

Coalescence times. Recall the remarks below Theorems 3.2–3.4. Let $m(G) = \mathbb{E}[\tau_{\text{meet}}^{\pi\times\pi}]$ be the expected first meeting time of two independent random walks starting from their stationary distribution π over a finite connected graph $G = (V, E)$. Recall that τ_{coal} is the first time such that a system of n random walks coalesce into a single one, where $n = |V|$. On the complete graph with n vertices it can be proved (see Aldous and Filll [1, Chapter 14]) that

$$\frac{\tau_{\text{coal}}}{m(G)} = \frac{\tau_{\text{coal}}}{(n-1)/2} \stackrel{d}{=} \sum_{i=2}^{n} Z_i,$$

where $(Z_i)_{i\in\mathbb{N}\setminus\{1\}}$ are independent random variables with law

$$Z_i \stackrel{d}{=} \text{Exp}\Big(\frac{1}{\binom{i}{2}}\Big), \quad i \in \mathbb{N} \setminus \{1\}.$$

In particular,

$$\mathrm{Law}\left(\frac{\tau_{\mathrm{coal}}}{m(G)}\right) \underset{n\to\infty}{\longrightarrow} \mathrm{Law}\left(\sum_{i\in\mathbb{N}\setminus\{1\}} Z_i\right)$$

and

$$\frac{\mathbb{E}[\tau_{\mathrm{coal}}]}{m(G)} \underset{n\to\infty}{\longrightarrow} 2.$$

As explained in Oliveira [48], Z_i is the time it takes for a system with i random walk to evolve to a system with $i-1$ random walks, rescaled by the expected meeting time of two random walks. For $i=2$, this is just the (rescaled) meeting time of a pair of random walks, which is an exponential random variable with mean 1. For $i>2$, we are looking at the first meeting time among $\binom{i}{2}$ pairs of random walks. It turns out that these pairwise meeting times are independent. Since the minimum of k independent exponential random variables with mean μ is an exponential random variable with mean μ/k, we deduce that Z_i is exponential with mean $1/\binom{i}{2}$. Such an interpretation comes from the Kingman coalescent [37]: a pure death process taking values in the collection of partitions of $[n]$ such that if the partition has j sets, then at rate $\binom{j}{2}$ two randomly chosen sets in the partition are joined together.

The aim of Oliveira [48] was to find a large class of graphs such that the latter equations hold. The class of graphs for which this mean-field-type behaviour had been proved earlier are: the discrete tori $(\mathbb{Z}/m\mathbb{Z})^d$ with $d\geq 2$ fixed and $m\gg 1$ (see Cox [22]), and large random d-regular graphs (see Cooper, Frieze and Radzik [21]). Oliveira [48] showed that the equations are satisfied for any finite, transitive and irreducible graph such that the mixing time of a single random walk occurs much faster than $m(G)$. Heuristically, in a short amount of time all but two random walks coalesce, and by that time the two remaining random walks have mixed well and will have to meet from stationarity.

Meeting times on random geometries. We aim to compare the meeting times on random regular graphs and inhomogeneous random graphs via the so-called annealed random walk. On the d-regular random graph one can exploit the locally-tree-like nature of the environment, which leads one to study the observable on a deterministic d-regular tree. Therefore, the problem is reduced to the study of a biased random walk on $\mathbb{Z}$, representing the graph distance of the two random walks, with jump probabilities $1/d$ and $(d-1)/d$. By studying random walks on locally-tree-like random graphs, such as the Configuration Model, the Erdős-Rényi rando graph and the Chung-Lu random graph (see e.g. van der Hofstad [34], Bollobás [7]), one may be interested in certain observables, such as the first meeting time, of random walks on Galton-Watson trees related to such geometries.

Fix $n\in\mathbb{N}$, and let $G=([n],E)$ be a locally-tree-like random graph with vertex set $V=[n]=\{1,\dots,n\}$. If we want to study τ_{meet} on a Galton-Watson tree having a non-trivial offspring distribution, in our case generated by G, then the problem becomes much harder, as now τ_{meet} becomes a *random variable* with respect to the realisation of the tree. For this reason, we cannot rewrite the problem in terms of a single biased walk on $\mathbb{N}_0$. One of the ways to deal with the two very dependent sources of randomness (the random graph given by the Galton-Watson tree and the stochastic process given by two independent random walks) is to *average out over the environment*. More precisely, let $\mathbf{P}=\mathbf{P}^G$ be the *quenched* law of the two random walks, i.e., the *random measure* that depends on the realisation of the random graph

G, and let $\mathbb{P}$ be the law of G. Let

$$\mathbb{E}[\mathbf{P}(\cdot)] = \sum_{G \in \mathfrak{G}_n} \mathbf{P}^{\mathcal{G}}(\cdot)\, \mathbb{P}(\mathcal{G} = G)$$

be the *annealed* law of two random walks on $\mathcal{G}$, where $\mathfrak{G}_n$ is the collection of all possible graphs with n vertices. Typically, in this setting, one can attempt to compute the annealed law of τ_{meet}, i.e., $\mathbb{E}[\mathbf{P}(\tau_{\text{meet}} \leq t)]$, $t \in \mathbb{N}_0$, and then try to recover the quenched result by a concentration argument.

Consider the CM G with degree sequence $\mathbf{d} = (d_1, \dots, d_n)$. It consists in a random graph with fixed vertex set $[n]$ and each vertex $x \in [n]$ has a pre-assigned degree d_x. We will assume that the degree sequence $\mathbf{d}$ satisfies some regularity conditions (see van der Hofstad [34]). It can be shown that, under these assumptions on the degree sequence, the resulting graph local limit is an unimodular Galton-Watson tree $\mathcal{T}_o$, whose root o has offspring distribution $p = (p_k)_{k \in \mathbb{N}}$ given by

$$p_k = \mathbb{P}(D_n = k) = \frac{1}{n} \sum_{x \in [n]} 1_{d_x = k},$$

where D_n is the degree of an uniformly chosen vertex, while the other sites have offspring distribution given by the size-biased law

$$\mu(k) = \frac{(k+1)\, p_{k+1}}{\mathbb{E}[D_n]}, \qquad k \in \mathbb{N}.$$

Consider now two asynchronous independent random walks X, Y with initial position o. We want to compute $\mathbb{E}[\mathbf{P}(\tau_{\text{meet}} \leq t)]$. There is one key observation that makes the above computation possible explicitly: τ_{meet} is a *local* observable depending only on the neighbours of the two walks at each step. This obvious, but crucial, fact brings two consequences:

1. We can exploit the *local weak limit* of the CM.

2. We can make an explicit description of the annealed law in terms of a non-Markovian process

   $$\mathbb{P}^{\text{ann}}(\cdot) = \mathbb{E}[\mathbf{P}(\cdot)]$$

   that at each step simultaneously explores the random graph and lets the walks move. This equivalent description is only based on the fact that in the definition the description of the meeting time is given by a local exploration of the graph structure, not a global exploration. For a precise description, see Bordenave et al. [8, 9], Cai et al. [15] and Avena et al. [4].

Suppose that t is uniformly bounded in n, i.e., $t = t_n = \mathcal{O}(1)$ as $n \to \infty$. In order to compute $\mathbb{P}^{\text{an}}(\tau_{\text{meet}} \leq t)$, we initialise G having n vertices equipped with $\{d_i\}_{i \in [n]}$ half-edges and without any formed edge (empty matching). We set both the walks X and Y such that $X_0 = Y_0 = o \sim \text{Unif}([n])$. Since the walks are asynchronous, we select which of the two moves first with probability (w.p.) $\frac{1}{2}$, say X moves first. Then it selects uniformly at random a half-edge e of $X_0 = o \in [n]$ and selects another half-edge f among the unmatched ones, say that f is incident to some vertex $z \in [n]$. Finally, we create the edge (o, z) by matching the half-edges e and f, and we move the walk X to z, i.e., $X_1 = z$, while $Y_1 = Y_0 = o$. We iterate this procedure until the walks meet at some time t, with the difference that if a walk selects

a half-edge that is already matched, then we do not sample a half-edge uniformly at random and we just let the walk move to the other end of the edge.

Since $t = \mathcal{O}(1)$, we can perform a local exploration of a Galton-Watson tree rooted at $o \sim \mathrm{Unif}([n])$ with offspring distribution μ as above, in place of a local exploration of the whole graph G. In order to exhibit an explicit example of $\mathbb{P}^{\mathrm{an}}(\tau_{\mathrm{meet}} \leq t)$, let us try to compute $\mathbb{P}^{\mathrm{an}}(\tau_{\mathrm{meet}} = 4)$. Suppose that the walks are *non-backtracking*.

- One of the walks, say X, moves to a neighbour z_1 of $X_0 = o$. This event happens w.p. $\frac{1}{2}$.
- Since the first meeting happens at time 4, Y cannot follow immediately X, and so X must move again as in Step 1 to some vertex z_2. This event happens w.p. $\frac{1}{2}$.
- Now Y must follow the unique path connecting X to Y for two steps, from o to z_1 and from z_1 to z_2. This happens with probability

$$\left(\tfrac{1}{2}\right)^2 \frac{1}{d_o} \frac{1}{d_{z_1}}.$$

  The offspring distribution of o is p_k as above, while the degree of z_1 is distributed according to μ. It follows that

$$\mathbb{P}^{\mathrm{an}}(\tau_{\mathrm{meet}} = 4) = \left(\tfrac{1}{2}\right)^4 \sum_{k \in \mathbb{N}} \frac{1}{k}\, p_k \sum_{k \in \mathbb{N}} \frac{1}{k}\, \mu(k).$$

REMARK: It is unclear what to do when the random walk is backtracking, like simple random walk.

3.5 VM on the random regular graph with random rewiring

What happens when the graph itself evolves over time, e.g. the edges are *randomly rewired*?

Suppose that every pair of edges swaps endpoints at rate $\nu/2M$ with $\nu \in (0, \infty)$, where we recall that $M = \frac{dN}{2}$ is the number of edges in the d-regular graph with N vertices. With this choice of parametrisation, the *rate at which a given edge is involved in a rewiring* converges to ν as $N \to \infty$. The voter model on this dynamic random graph evolves as before: at rate 1 opinions are adopted along the edges that are currently present.

In work in progress we show that Theorems 3.2–3.4 carry over with θ_d replaced by $\theta_{d,\nu}$ given by a continued fraction.

Theorem 3.7 (Avena, Hazra, den Hollander, Quattropani [3]). *Let* $\beta_d = \sqrt{d-1}$ *and* $\rho_d = \frac{2}{d}\sqrt{d-1}$. *Then*

$$\theta_{d,\nu} = 1 - \frac{\Delta_{d,\nu}}{\beta_d}$$

with

$$\Delta_{d,\nu} = \frac{1\,|}{|\,\frac{2+\nu}{\rho_d}} - \frac{1\,|}{|\,\frac{2+2\nu}{\rho_d}} - \frac{1\,|}{|\,\frac{2+3\nu}{\rho_d}} - \cdots$$

1. Since the d-regular random graph locally looks like a d-regular tree, the proof proceeds by analysing the meeting time of two random walks on a d-regular tree. On short to modest time scales the two random walks do not notice the difference. Work is needed to show that on longer time scales the approximation is still good.

2. We replace rewiring of edges on the d-regular random graph by disappearance of edges on the d-regular tree. This is a good approximation because, as soon as one random walk moves along a rewired edge in the d-regular random graph, it is thrown far away from the other random walk and meeting becomes difficult.

KEY OBSERVATION: Because $\nu \mapsto \theta_{d,\nu}$ is strictly increasing, the *dynamics speeds up consensus*.

4 LECTURE 4: The Contact Process (CP)

4.1 CP on graphs

Given a connected graph $G = (V, E)$, the contact process is the Markov process $(\xi_t)_{t\geq 0}$ on state space $\{0,1\}^V$ where each vertex is either healthy (0) or infected (1). Each infected vertex becomes healthy at rate 1, independently of the state of the other vertices, while each healthy vertex becomes infected at rate λ times the number of infected neighbours, with $\lambda \in (0,\infty)$ the infection rate.

The configuration at time t is $\xi_t = \{\xi_t(i)\colon\ i \in V\}$, with $\xi_t(i)$ the state at time t of vertex i. In what follows we will analyse the behaviour of the CP on various classes of graphs, both random and deterministic. Our focus will be on understanding how the extinction time

$$\tau_{[0]_N} = \inf\{t \geq 0 :\ \xi_t(i) = 0\ \forall\, i \in V\}$$

behaves as $|V| = N \to \infty$, depending on the value of λ and the properties of the graph. We will mostly zoom in on

$$\mathbb{E}_{\tau_{[1]_N}}(\tau_{[0]_N}),$$

the average extinction time starting from the configuration where every vertex is infected. We will see that it is hard to get control on this quantity, so we will have to content ourselves with rough bounds.

4.2 CP on the complete graph

As a prelude we look at the CP on the complete graph, for which computations can be carried through explicitly. Indeed, the number of infections at time t, given by

$$I_t^N = \sum_{i\in V} \xi_t(i),$$

evolves as a continuous-time nearest-neighbour random walk on the set $\{0,\dots,N\}$ with transition rates

$$\begin{aligned} n &\to n+1 \quad \text{at rate } \lambda n(N-n),\\ n &\to n-1 \quad \text{at rate } n. \end{aligned}$$

Put $\mathcal{I}_t^N = \frac{1}{N} I_t^N$ for the fraction of infected vertices at time t. This process is a continuous-time nearest-neighbour random walk on the set

$$\{0, \tfrac{1}{N}, \dots, \tfrac{N-1}{N}, 1\}$$

with transition rates

$$\begin{array}{ll} x \to x + N^{-1} & \text{at rate } \lambda x(1-x)N^2, \\ x \to x - N^{-1} & \text{at rate } xN. \end{array}$$

This process has a strong drift upward, which becomes zero for $x = 1 - \frac{1}{\lambda N}$, i.e., very close to full infection when $\lambda N \gg 1$.

Theorem 4.1 *Let $\tau_{[0]_N} = \inf\{t \geq 0 \colon \xi_t = [0]_N\}$ be the extinction time. Then*

$$\log \mathbb{E}_{[1]_N}(\tau_{[0]_N}) = N[1 + \log(\lambda N)] + o(N), \quad N \to \infty.$$

Thus, the CP on the complete graph is supercritical for all $\lambda > 0$ as $N \to \infty$.

Theorem 4.2 (Schapira, Valesin [50].)
For every $\lambda \in (0, \infty)$,

$$\lim_{N\to\infty} \mathbb{P}_{[1]_N}\left(\frac{\tau_{[0]_N}}{\mathbb{E}_{[1]_N}(\tau_{[0]_N})} > t\right) = \mathrm{e}^{-t} \quad \forall\, t > 0.$$

In the mean-field setting of the complete graph, the fraction of infected vertices performs a random walk. The latter property *fails* on non-complete graphs, in particular, on random graphs: $(\mathcal{I}_t^N)_{t\geq 0}$ loses the Markov property. The CP is harder than the VM because it *does not have a tractable dual*. In fact, it is self-dual.

4.3 CP on the configuration model

Consider CP on CM with an empirical degree distribution f_N satisfying $\lim_{N\to\infty} \|f_N - f\|_\infty = 0$ with

$$f(k) = k^{-\tau + o(1)}, \qquad k \to \infty,$$

where $\tau \in (1, \infty)$ is the tail exponent.

Theorem 4.3 (Chatterjee, Durrett [20], Mountford, Mourrat, Valesin, Yao [45]).
If $\tau \in (2, \infty)$, then for every $\lambda \in (0, \infty)$ the average time to extinction grows exponentially fast with N whp.

Thus, CP on CM with a power law degree distribution is supercritical regardless of the value of λ. Apparently, hubs easily transmit the infection.

Theorem 4.4 (Can, Schapira [17]).
The same is true for $\tau \in (1, 2]$, even though local convergence breaks down.

For the CP on the CM with a power law degree distribution there is anomalous scaling of the density of infections $\rho(\lambda)$ as $\lambda \downarrow 0$, namely,

$$\rho(\lambda) \asymp \begin{cases} \lambda^{1/(3-\tau)}, & \tau \in (2, \frac{5}{2}], \\ \lambda^{2\tau-3} \left[\log(1/\lambda)\right]^{-(\tau-2)}, & \tau \in (\frac{5}{2}, 3], \\ \lambda^{2\tau-3} \left[\log(1/\lambda)\right]^{-2(\tau-2)}, & \tau \in (3, \infty). \end{cases}$$

(Mountford, Valesin, Yao [46], Linker, Mitsche, Schapira, Valesin [41].) The three regimes reflect different optimal strategies to survive extinction for small infection rates: the infection survives close to hubs.

Sharp estimates of the extinction time have been obtained for the CP on the CM with i.i.d. degrees $(D_i)_{i=1}^N$ taking values in $\mathbb{N}_0$. When $\mathbb{E}[D_1] < \infty$ we expect a strictly positive critical threshold.

Theorem 4.5 (Cator, Don [19]).
Suppose that $\mathbb{E}[D_1] < \infty$ and $\mathbb{E}[2^{-D_1/2}] < \frac{1}{4}$. Then there exists a constant $\alpha \in (0, \mathbb{E}[D_1]]$ such that if $\lambda > 1/\alpha$, then there exists a constant $c > 0$ such that

$$\mathbb{E}_{[1]_N}(\tau_{[0]_N}) \geq \mathrm{e}^{cN} \quad \text{whp.}$$

For the random regular graph with degree $d \geq 3$, the claim holds with $\alpha = d - 2$.

4.4 CP on the Erdős-Rényi random graph

Let $p = p_N$ be the edge retention probability and $\lambda = \lambda_N$ the infection rate.

Theorem 4.6 (Cator, Don [19].)
(a) If $\lim_{N\to\infty} Np_N = \infty$, then

$$\mathbb{E}_{[1]_N}(\tau_{[0]_N}) \geq \mathrm{e}^{c_N N} \quad \text{whp}$$

for $c_N < \log(Np_N\lambda_N) - \frac{1}{Np_N\lambda_N} + 1$ and N large enough.
(b) If $\lim_{N\to\infty} Np_N = \sigma \in (4\log 2, \infty)$, then

$$\mathbb{E}_{[1]_N}(\tau_{[0]_N}) \geq \mathrm{e}^{cN} \quad \text{whp}$$

for $\lambda > 1/\sigma$ and $c < \log(\sigma\lambda) - \frac{1}{\sigma\lambda} + 1$ and N large enough.

4.5 CP on the preferential attachment model

Theorem 4.7 (Berger, Borgs, Chayes, Saberi [6], Can [16]).
There exists a $c > 0$ such that

$$\log \mathbb{E}_{[1]_N}(\tau_{[0]_N}) \geq c \, \frac{\lambda^2 N}{(\log N)^{\frac{1+\gamma}{1-\gamma}}}$$

for $\lambda > 0$ small enough and N large enough, with $\gamma \in [0, 1)$ a parameter that controls the attachment probabilities of newly incoming vertices.

4.6 CP on tree-like random graphs

Many sparse random graphs are locally tree-like, and hence it is interesting to study the extinction time of the CP on regular trees. Let

$$0 < \lambda_d < \infty$$

be the critical threshold for survival on the d-regular tree.

Theorem 4.8 (Mourrat, Valesin [47]).
Let $\mathcal{G}_{d,N}$ be the class of connected graphs with N vertices and maximal degree d. Then, for $\lambda < \lambda_d$,

$$\lim_{N\to\infty} \sup_{G\in\mathcal{G}_{d,N}} \mathbb{P}_{[1]_G}(\tau_{[0]_G} > c \log N) = 0 \quad \textit{for some } c = c(\lambda, d).$$

Theorem 4.9 (Mourrat, Valesin [47]).
On the d-regular random graph with $d \geq 3$, the crossover from logarithmic to exponential extinction time occurs at λ_d.

Theorem 4.10 (Baptista da Silva, Oliveira, Valesin [5], Schapira, Valesin [51]).
On the dynamic d-regular random graph with $d \geq 3$ and rewiring rate $\nu > 0$, the crossover occurs at a strictly smaller value than λ_d.

The rewiring helps the infection spread more easily.

4.7 CP on general finite graphs

What can be said about the extinction time for the CP on general finite graphs?

Theorem 4.11 (Mountford, Mourrat, Valesin, Yao [45]).
For any $\lambda > \lambda_1$, any $D \in \mathbb{N}$ and any connected graph G whose degrees are bounded by D,

$$\mathbb{E}_{[1]_G}(\tau_{[0]_G}) \geq \exp\left[c|V|\right]$$

$$\textit{for somec} = c(\lambda, D) > 0.$$

Theorem 4.12 (Schapira, Valesin [49]).
For any $\lambda > \lambda_1$, any $\epsilon > 0$ and any connected graph G,

$$\mathbb{E}_{[1]_G}(\tau_{[0]_G}) \geq \exp\left[\frac{c|V|}{(\log|V|)^{1+\epsilon}}\right] \quad \textit{for some } c = c(\epsilon) > 0.$$

CONCLUSION: For the CP it is hard to get sharp control on the extinction time, and many questions remain open.

References

[1] D. Aldous, J.A. Fill, *Reversible Markov Chains and Random Walks on Graphs* (unfinished monograph 2002, recompiled 2014) [http://www.stat.berkeley.edu].

[2] L. Avena, R. Baldasso, R.S. Hazra, F. den Hollander, M. Quattropani, Discordant edges for the voter model on regular random graphs, ALEA Lat. Am. J. Probab. Math. Stat. 21 (2024) 431–464.

[3] L. Avena, R. Baldasso, R.S. Hazra, F. den Hollander, M. Quattropani, The voter model on random regular graphs with random rewiring. Work in progress.

[4] L. Avena, F. Capannoli, R.S. Hazra, M. Quattropani, Meeting, coalescence and consensus time on directed random graphs, Ann. Appl. Probab. 34 (2024) 4940–4997.

[5] G.L. Baptista da Silva, R.I. Oliveira, D. Valesin, The contact process over a dynamical d-regular graph, *arXiv preprint* [arXiv:2111.11757].

[6] N. Berger, C. Borgs, J.T. Chayes, A. Saberi, Asymptotic behavior and distributional limits of preferential attachment graphs, Ann. Probab. 42 (2014) 1–40.

[7] B. Bollobás, *Random Graphs* (2nd. ed.), Cambridge Studies in Advanced Mathematics 73, Cambridge University Press, 2001.

[8] C. Bordenave, P. Caputo, J. Salez, Random walk on sparse random digraphs, Probab. Theory Relat. Fields 170 (2018) 933–960.

[9] C. Bordenave, P. Caputo, J. Salez, Cutoff at the "entropic time" for sparse Markov chains, Probab. Theory Relat. Fields 173 (2019) 261–292.

[10] A. Bovier, M. Eckhoff, V. Gayrard, M. Klein, Metastability and low lying spectra in reversible Markov chains, Comm. Math. Phys. (2002) 228:219–255.

[11] A. Bovier, V. Gayrard, The thermodynamics of the Curie-Weiss model with random couplings, J. Stat. Phys. 72 (1993) 643–664.

[12] A. Bovier, F. den Hollander, *Metastability – a Potential-Theoretic Approach*, Grundlehren der mathematische Wissenschaften 351, Springer, Berlin, 2015.

[13] A. Bovier, F. den Hollander, S. Marello, E. Pulvirenti, M. Slowik, Metastability of Glauber dynamics with inhomogeneous coupling disorder, ALEA, Lat. Am. J. Probab. Math. Stat. 21 (2024) 1249–1273.

[14] A. Bovier, S. Marello, E. Pulvirenti, Metastability for the dilute Curie–Weiss model with Glauber dynamics, Electron. J. Probab. 26 (2021) 1–38.

[15] X.S. Cai, P. Caputo, G. Perarnau, M. Quattropani, Rankings in directed configuration models with heavy-tailed in-degrees, Ann. Appl. Probab. 33 (2023) 5613–5667.

[16] V.H. Can, Metastability for the contact process on the preferential attachment graph, arXiv preprint arXiv:1502.05633.

[17] V.H. Can, B. Schapira, Metastability for the contact process on the configuration model with infinite mean degree, Electron. J. Probab. 20 (2015) paper no. 26, 1–22.

[18] F. Capannoli, Evolution of discordant edges in the voter model on random sparse digraphs, arXiv preprint arXiv:2407.06318.

[19] E. Cator, H. Don, Explicit bounds for critical infection rates and expected extinction times of the contact process on finite random graphs, Bernoulli 27 (2021) 1556–1582.

[20] S. Chatterjee, R. Durrett, Contact processes on random graphs with power law degree distributions have critical value 0, Probab. Theory Relat. Fields 37 (2009) 2332–2356.

[21] C. Cooper, A. Frieze, T. Radzik, Multiple random walks in random regular graphs, SIAM J. Discr. Math. 23 (2009) 1738–1761.

[22] J.T. Cox, Coalescing random walks and voter model consensus times on the torus in $\mathbb{Z}^d$, Ann. Probab. 17 (1989) 1333–1366.

[23] J.T. Cox, A. Greven, On the long term behavior of some finite particle systems, Probab. Theory Relat. Fields 85 (190) 195–237.

[24] J.T. Cox, A. Greven, T. Shiga, Finite and infinite systems of interacting diffusions, Probab. Theory Relat. Fields 103 (1995) 165–197.

[25] J.T. Cox, A. Greven, T. Shiga, Finite and infinite systems of interacting diffusions: cluster formation and universality properties, Math. Nachr. 192 (1998) 105–124.

[26] Y.-T. Chen, J. Choi, and J.T. Cox, On the convergence of densities of finite voter models to the Wright-Fisher diffusion, Ann. Inst. Henri Poincaré Probab. Stat. 52 (216) 286–322.

[27] S. Dommers, Metastability of the Ising model on random regular graphs at zero temperature, Probab. Theory Relat. Fields 167 (2017) 305–324.

[28] S. Dommers, F. den Hollander, O. Jovanovski, F.R. Nardi, Metastability for Glauber dynamics on random graphs, Ann. Appl. Probab. 27 (2017) 2130–2158.

[29] H. Duminil-Copin, V. Tassion, A. Teixeira, The box-crossing property for critical two-dimensional oriented percolation. Probab. Theory Relat. Fields 171 (2018) 685–708.

[30] R. Durrett, *Lecture Notes on Particle Systems and Percolation*, Wadsworth & Brooks Cole, Pacific Grove, California, USA, 1988.

[31] R. Durrett, X.-F. Liu, The contact process on a finite set, The contact process on a finite set, Ann. Prob. 16 (1988) 1158-1173.

[32] R. Durrett, R.H. Schonmann, The contact process on a finite set, II. Ann. Prob. 16 (1988) 1570–1583.

[33] S.N. Ethier and T.G. Kurtz (1986), *Markov Processes: Characterization and Convergence*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York.

[34] R. van der Hofstad, *Random Graphs and Complex Networks*, Volume 1, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 2017.

[35] R. van der Hofstad, *Random Graphs and Complex Networks*, Volume 2, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 2014.

[36] F. den Hollander, O. Jovanovski, Glauber dynamics on the Erdős-Rényi random graph, Progress in Probability 77, Birkhäuser, 2021, pp. 519-589.

[37] J.F.C. Kingman, The coalescent, Stoch. Proc. Appl. 13 (1982) 235–248.

[38] T.M. Liggett, *Interacting Particle Systems*, Grundlehren der mathematische Wissenschaften 276, Springer, New York, 1985.

[39] T.M. Liggett, *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*, Grundlehren der mathematische Wissenschaften 324, Springer, Berlin, 1999.

[40] T. Lindvall, *Lectures on the Coupling Method*, Wiley Series in Probability and Statistics, Wiley-Interscience, 1992.

[41] A. Linker, D. Mitsche, B. Schapira, D. Valesin, The contact process on random hyperbolic graphs: metastability and critical exponents, Ann. Probab. 49 (2021) 1480–1512.

[42] F.J. López, G. Sanz, Duality for general interacting particle systems, Markov Proc. Relat. Fields 6 (2000) 305–328.

[43] T. Mountford, A metastable result for the finite multidimensional contact process, Can. Math. Bull. 36 (1993) 216–226.

[44] T. Mountford, Existence of a constant for finite system extinction, J. Stat. Phys. 96 (1999) 1331–1341.

[45] T. Mountford, J-C. Mourrat, D. Valesin, Q. Yao, T. Mountford, Exponential extinction time of the contact process on finite graphs, Stoch. Proc. Appl. 126 (2016)1974–2013.

[46] T. Mountford, D. Valesin, Q. Yao, Metastable densities for the contact process on power law random graphs, Electron. J. Prob. 18 (2013), paper no. 103, 1–36.

[47] J.-C. Mourrat, D. Valesin, Phase transition of the contact process on random regular graphs, Electron. J. Probab. 21 (2016) 1–17.

[48] R.I. Oliveira, Mean field conditions for coalescing random walks, Ann. Probab. 41 (2013) 3420–3461.

[49] B. Schapira, D. Valesin, Extinction time for the contact process on general graphs, Probab. Theory Relat. Fields 169 (2017) 871–899.

[50] B. Schapira, D. Valesin, Exponential rate for the contact process extinction time, Annales de la Faculté des Sciences de Toulouse: Mathématiques, Vol. 30 (2021) 503–526.

[51] B. Schapira, D. Valesin, The contact process on dynamic regular graphs: monotonicity and subcritical phase, arXiv preprint arXiv:2309.17040.

[52] R.H. Schonmann, S.B. Shlosman, Wulff droplets and the metastable relaxation of kinetic Ising models, Comm. Math. Phys. 194 (1998) 389–462.