Università degli studi di Padova

Dipartimento di Matematica "Tullio Levi-Civita"

Corso di Laurea Magistrale in Matematica

Voter model on k-partite graphs

Asymptotic behavior conditional on non-absorption

Supervisor:
Alessandra Bianchi

Graduating:
Federico Capannoli
Badge number:
1207240

25 June 2021

Academic Year 2020-21

Contents

Introduction

The first questions that should be answered when meeting a new subject to study are: is it really interesting? Why and how could it benefit the community? Regarding this thesis, we will deal with the so called interacting particle systems (IPSs). Without any doubt, the interest of such a research field lies in the application of the huge theory of stochastic processes to real life models, such as the macroscopic description of ferromagnetic materials or the voting attitude of a large group of people during an election. Moreover, the modelling component itself witnesses the great contribution that could give to the community. As sake of example we recall that epidemiology, which through all 2020 has unfortunately acquired notoriety all over the word, is all about the study of the spread of a disease; that is, an IPS where the particles set is composed by people and the set of interactions is given by their contacts. As often happens when it comes to mathematical issues, the field of interacting particle systems started to be studied for a specific topic such as statistical mechanics, and soon after it spread to other fields of research. From a mathematical point of view, interacting particle systems represents a natural departure from the established theory of Markov processes. A typical interacting particle system consists of finitely or infinitely many particles which, in absence of interaction, would evolve according to independent finite or countable state Markov chains. If we add some type of interaction, the evolution of an individual particle is no longer Markovian, while the system as a whole still is, though a very complex one. A peculiarity of interacting particle system is the state space, given by a metric space $X = W^V$, where V is a finite or countable set of locations (called *sites*), and W is a compact metric space which will play the role of the phase space of the component located at each site. The whole process is defined by $\{\eta_t(v), t \geq 0, v \in V\}$, where $\eta_t(\cdot) \in W^V$ represents the phase of a general vertex at a fixed time $t \geq 0$, while $\eta \in W^V$ constitutes a state of the system, given by the datum of the phases of all vertices. Another difference between a classical Markov chain and an IPS is the mechanism of the flipping rates from one state η to another. It is described by a set of nonnegative functions $c(x, \eta)$ that are strictly related to both interaction between agents $x \in V$ and the transition probabilities on V. The interaction among sites comes from the dependence of $c(x, \eta)$ on $\eta \in X$.

A relevant example of IPS is the so called voter model. In the previous notations, it consists in considering W as a finite set of opinions, and as V a finite or countable metric space; in the work of Holley and Liggett [15] was treated the case $W = \{0, 1\}$ and $V = \mathbb{Z}^d$, maybe the most relevant one, since it shows directly the connection between the model and the random walks on the lattice $\mathbb{Z}^d$. We will mainly consider the case where there are only two phases: 0 and 1. The voter interpretation which gives this process its name views V as a collection of individuals, each one taking two possible positions, denoted by 0 and 1, on a political issue. The process evolves as follows: starting from any distribution of opinions on the sites, a randomly chosen site waits for an exponential time of parameter one and then possibly changes his opinion adopting the one of a (still randomly chosen) neighbour site.

The main problems which have been treated involve the long-time behaviour of the system. We first want to derive limit theorems, and for this purpose we need to describe the class of invariant measures for the process, since these are the possible asymptotic distribution. We know that the voter model is not ergodic, since there are two trivial invariant measures given by the pointmasses at $\eta \equiv 0$ and $\eta \equiv 1$. We emphasize that the latter states of the system represent the situations in which every site has the same opinion, they will be referred to as *consensus* states; it is clear that consensus states are absorbing for the process, since once the system reaches the agreement it gets trapped and remains in that state forever. If we take into account the case of Holley and Liggett on the d-dimensional lattice, it can be proved that: if $d \leq 2$, then there are no extremal invariant measures except from the two trivial ones, and therefore the system reach consensus states almost surely. On the contrary, if $d \geq 3$, there exist a continuum of extremal invariant measures; it implies that, as time goes by, coexistence of different opinions occurs almost surely. This dichotomy is closely related to the fact that a simple random walk on $\mathbb{Z}^d$ is recurrent if $d \leq 2$ and transient if $d \geq 3$.

In this thesis, we will study the voter model with two opinions on finite graphs G, i.e. the case in which the set of sites V coincides with the vertex set of G. If we suppose that the Markov chain on V is irreducible, then it follows that, by finiteness of the graph, consensus states are reached almost surely. It is equivalent to say that all the stationary distributions are trivial, concentrated on the absorbing sets. With these hypothesis, it is relevant to look for the consensus time: the hitting time to reach one of the consensus states. We will not address this problem, but in this regard we want to cite the remarkable work of Cox [5] on the d-dimensional tourus.

From another point of view, it may be interesting to study the evolution of the process, still on finite graphs, conditioned to never reach consensus states. For this purpose we will restrict the state space, eliminating all the absorbing -consensus- states. Thus, in order to study the long-time behaviour of the conditioned process, we can no longer use the invariant measures of the original process, we must instead compute the quasi-stationary distributions (QSDs). The latter tool will play a fundamental role because it is the analog of the invariant measures for the conditioned process: they are eigenfunctions corresponding to eigenvalues with a modulus that is strictly less than one. A

sufficient condition for the existence and uniqueness of the QSDs is that the corresponding stochastic process must be irreducible and aperiodic before reaching the absorbing states. In that case, we know that the absorbing time, i.e. the hitting time of the absorbing states for the process, starting from the QSD is exponentially distributed with a parameter that strongly depends on the QSD itself. Moreover, it can be proved that QSDs for Markov chains $\{X(t), t \geq 0\}$ conditioned on absorbing states has the same limit behaviour of the invariant measures related to the non-conditioned chain, namely that they are limit distributions for the conditioned process as $t \to \infty$. In this work, therefore, we will deal with the voter model on finite graphs, studying its limit behaviour when the process is conditioned to never reach the consensus states.
From the recent article of Ben-Ari, Panzo, Speegle and VandenBerg [2], we know that several result has been obtained for the voter model on complete bipartite graphs. Following their paper we will discuss about the stated results, trying to generalize them on several aspects. For a matter of clarity, we recall that a complete bipartite graph $K_{n,m}$ is an heterogeneous graph whose vertex set can be partitioned in two disjoint groups, a "large" group L of size n and a "small" one S of size m, where each vertex of L is connected to all of the vertices of S and vice versa, and there are no connections between vertices in the same group. Our aim is to investigate what happens when consensus is conditioned to never occur, as we know that in finite graphs it happens almost surely. More specifically, we will study the quasi-stationary distributions for the voter model with two opinions on $K_{n,m}$ and its limit behaviour under the QSDs when m is fixed and $n \to \infty$. The reasons why we are interested in such a limit are: to find out if the lack of consensus is due to a minority number of dissenters, or if the opinions are relatively balanced, and, taking $n \gg m$, to find out how the distributions of opinions in L differs from S. There are three main tools used to achieve the desired results. The first, fundamental one is the well-known duality (we will consider a discrete-time version) between the voter model and the coalescing random walk. The latter consists in a system of coupled random walks (RWs) which evolves independently through the vertices of the bipartite graph, and once at least two of them meet in the same vertex they start to move as a single one. The primary application will concern the rewriting of the distribution of the hitting time for consensus states as the distribution of the hitting time for which all the RWs coalesced into one. This leads to the next argument: the spectral radius of the sub-stochastic transition matrix. Initially we will deal with the discrete time voter model on complete bipartite graphs, thus the transition function conditioned on non-absorption will be a sub-stochastic matrix. Using the Gelfand's Formula, we can write the spectral radius of such matrix as a limit concerning the distribution of the hitting time for consensus; this fact will be crucial since the latter limit coincides with the eigenvalue of the (unique) QSD for the process. In other words, we are able to determine the eigenvalue of the QSD just computing the spectral radius of the sub-stochastic transition matrix of the process. As a last claim, we will also prove that the whole duality approach, used to

compute the spectral radius, can be restricted to a two-dimensional chain composed by two coalescing random walks only.
Following the results in [2], we wondered if it was possible to extend them to more general contexts. We tried to answer to the following questions: given the result on the bipartite case, what happens if we consider the complete k-partite graphs, $k \geq 2$? Do the same conclusions hold? Moreover, in the bipartite setting, what should we expect if we suppose that both sizes of the sets diverges as the size of the former one goes to infinity? Is there a threshold on the size of the second set such that we observe a phase transition in the dynamic of the asymptotic system? During this thesis we will gradually deal with such questions, and we will try to exhaustively answer them in the last chapter.
Another interesting approach to generalize the problem is given by Fernley and Ortgiese in [12]. They discuss the consensus time for the voter model, without any conditioning, on subcritical inhomogeneous random graphs. The main difference between the latter setting and ours relays on the randomness of the graph itself: there will be an edge between any couple of vertices only with a certain probability distribution.
This thesis is organized as follows. In Chapter 1 we recall all the preliminary results, regarding continuous-time Markov chains and stationary measures, that we need in order to treat the subsequent topics. Then, in Chapter 2 we introduce the concept of quasi-stationary distributions in the more general context of Markov processes, highlighting the general properties from both spectral and asymptotic point of view. In particular, in Section 2.3 we discuss the case where the state space is finite and we state the Perron-Frobenious Theorem, an essential tool which provides us conditions for existence and uniqueness of QSDs. At the end of the chapter we provide an example for QSDs in birth-death processes, where we briefly show cases in which: QSDs are unique, do not exist, and there are an uncountable infinity. In Chapter 3 we entirely change argument, treating the interacting particle systems. To do so, we first introduce IPSs in their generality and then we focus on the special case of the voter model. Besides describing the latter model, the crucial aspects of this chapter regards the connection between invariant measure for the process and the absorbing states, and the duality of the voter model with the coalescing random walk. Our main reference for this chapter is [16].
Through Chapter 4 we study the voter model with two opinions in complete bipartite graphs $K_{n,m}$. Since those are finite graphs, we know that consensus states are reached almost surely by the process. The aim of this chapter is to investigate the asymptotic behaviour of the quasi-stationary distributions for the process conditioned to never reach consensus states, as $n \to \infty$ and m fixed. The leading result shows that the unique QSD on $K_{n,m}$ converges in distribution to a probability measure, as n tends to infinity; we will show with an example that it is not obvious in general. Moreover, it holds that the distribution of opinions under the unique QSD for the voter model (with two opinions) on $K_{n,m}$ as $n \to \infty$ converges weakly to the following: all vertices of S, the "small" set of size m, have opinion $C \sim \mathrm{Bern}(1/2)$, while all but $D \sim \mathrm{Sib}(\gamma_m)$ vertices in L,

the "large" set of size $n \to \infty$, have opinion C. Where, with Bern(1/2) we indicate the Bernoulli distribution of parameter $\frac{1}{2}$, and with $\mathrm{Sib}(\gamma_m)$ we mean the Sibuya distribution, a discrete probability distribution with parameter γ_m which depends on the size m of S. Finally, in Chapter 5 we will discuss two possible generalizations of the results given by Ben-Ari et al. in [2]. In the first generalization we will consider the Voter model evolving in complete k-partite graphs, with $k \geq 2$. Our goal is to provide the limit behaviour of the QSD for the conditional process as the size of one of the k sets goes to infinity, while the other $k-1$ sizes remain fixed. We will find that for the k-partite graphs hold a result that is completely similar to the bipartite case: the dissenters are all located in the set whose size goes to infinity, and are distributed with a Sibuya distribution with parameter $\gamma \in (0, 1)$ which depends on the fixed sizes of the remaining $k-1$ sets, while all vertices in the latter share the same opinion -there is no space for dissenters in the small sets-, which can be 0 or 1 according to a Bernoulli distribution of parameter 1/2. As our second generalization, we will still consider the bipartite case, but this time we will suppose that both sizes of the sets diverges as the size of the first one goes to infinity. We will show that the limit in distribution for the QSD does not converge to a probability measure, leading to the same result that holds for complete graphs.

Chapter 1

Preliminaries

1.1 Continuous time Markov chains

We begin the first of this preliminary sections with an introduction to Markov chains in continuous time. Most of the results, definitions and proofs were taken from [19] and [3]. As first we recall some basics of discrete time Markov chains. Let us consider a stochastic process, that is a collection of random variables $\{X_j : j \in J\}$ defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$; we will always consider the index set J as a discrete or continuous set of time points as $J \subset \mathbb{Z}_+$ or $J \subset \mathbb{R}_+$. Let S be a finite or countable set, the state space of the process.

Definition 1.1. A sequence $(X_n)_{n\in\mathbb{N}}$ of S-valued random variables is a *Markov chain* if for each $n \in \mathbb{N}_{\geq 1}$ and x_0,x_1,... $x_{n+1} \in S$

$$P(X_{n+1} = x_{n+1} \,|\, X_n = x_n, \dots X_1 = x_1, X_0 = x_0) = P(X_{n+1} = x_{n+1} \,|\, X_n = x_n). \quad (1.1)$$

We will refer to this equation as *Markov property*. If the right-hand side of (1.1) does not depend on n, we say that the chain is *time- homogeneous*. In this case is well defined the matrix $\mathbf{P}$ given by

$$\mathbf{P} := (p(x,y))_{x,y\in S}, \quad \text{where} \quad p(x,y) = P(X_{n+1} = y \,|\, X_n = x), \quad \forall x, y \in S,$$

and it is called the *transition matrix* of the (discrete time) Markov chain $(X_n)_n$. Note that since all the entries are probabilities, and since $\{X_n = x\}, x \in S$ is a partition of Ω, it follows that $\mathbf{P}$ is a stochastic matrix, in particular

$$p(x,y) \geq 0 \quad \forall x, y \in S \quad \text{and} \quad \sum_{z\in S} p(x,z) = 1, \quad \forall x \in S.$$

Next, the random variable X_0 is called *initial state* and its probability distribution, say ν, given by $\nu(x) = P(X_0 = x)$, is the initial distribution of the chain. From Bayes's sequential rule and in view of the time-homogeneous Markov property, it follows that

$$P(X_0 = x_0, \dots X_k = x_k) = \nu(x_0)p(x_0,x_1)\cdots p(x_{k-1},x_k), \quad \forall k \geq 0, x_0, x_1, \dots x_{k+1} \in S.$$

In other words the initial distribution and the transition matrix determine the law of the process. Moreover, defining ν_n as the vector representing the distribution of the chain at time n, it's straightforward verify that

$$\nu_n^T = \nu_0^T \mathbf{P}^n,$$

where with ν_n^T we indicate the transpose of the column vector ν_n, and with $\mathbf{P}$ the transition matrix defined above.
Now that we got all the main results of a discrete-time Markov chain, we are able to define its continuous-time analog. Let's consider a stochastic process $\{X(t), t \geq 0\}$ with state space S finite or countable. Given that the process is in a state $i \in S$, the holding time in that state will be exponentially distributed with parameter $\lambda(i)$. The sequence of states visited will follow a discrete time Markov chain. Recall finally that holding times are independent random variables. These are our hypothesis. Consequently, the building blocks to construct the desired process $\{X_t\}_{t\geq 0}$ are: a discrete time Markov chain $\{X_n, n \geq 0\}$ on the state space S with transition matrix $\mathbf{P} = (p(i,j))_{i,j\in S}$, assuming $p(i,i) = 0$ for all $i \in S$; a sequence $\{E_n, n \geq 0\}$ of i.i.d. exponentially distributed random variables with rate 1 and independent of $\{X_n\}$; and finally a sequence of holding times parameters $\{\lambda(i) > 0, i \in S\}$. Given these ingredients, it can be seen that the resulting process is a continuous time Markov chain, in the following sense

Definition 1.2. The S-valued process $\{X_t\}_{t\geq 0}$ is called a *continuous-time Markov chain* if for all $i, j, i_1, \ldots, i_k \in S$ and all $s_1, \ldots, s_k \geq 0$ with $s_l \leq s$ for all $l \in [1, k]$

$$P(X(t+s) = j \,|\, X(s) = i, X(s_1) = i_1, \ldots, X(s_k) = i_k) = P(X(t+s) = j \,|\, X(s) = i), \tag{1.2}$$

whenever both sides are well-defined.

Note that equation (1.1) is the analogous of (1.2). This continuous-time Markov chain is called *homogeneous* if the right-hand side of (1.2) is independent of s. In that case we can define

$$P(t) := \{p_t(i,j)\}_{i,j\in S},$$

where

$$p_t(i,j) := P(X(t+s) = j \,|\, X(s) = i), \quad \forall s \geq 0.$$

The family $\{P(t)\}_{t\geq 0}$ is called the *transition semi-group* of the continuous-time homogeneous Markov chain.
In other words, the construction criterion mentioned above concerning the definition of our new process is all about setting $X(t) = X_n$ for $T_n \leq t < T_{n+1}$, where $\{T_{n+1} - T_n\}_{n\geq 1}$ is a sequence of conditionally independent and exponentially distributed random variables given $\{X_n\}$, i.e. $T_{n+1} - T_n = E_n \backslash \lambda(X_n)$. Note that this kind of construction defines $\{X_t\}$ only up to time $T_\infty := \lim_{n\to\infty} T_n$. If $T_\infty < \infty$, we say an explosion occurs, because an infinite number of transitions have taken place in finite time. When

$$P_i(T_\infty = \infty) := P(T_\infty = \infty \,|\, X(0) = i) = 1 \quad \forall i \in S$$

we say that the process is *regular*.

Notation. Keep in mind the difference between the similar expressions $p_t(i,j)$ and $p(i,j)$: the first one is the general term of the semi-group, while the second one is the general term of the transition matrix of the discrete-time Markov chain $\{X_n\}_n$.

$\{P(t)\}_{t\geq 0}$ is a semi-group thanks to the two following properties:

1. The Chapman-Kolmogorov equation

$$p_{t+s}(i,j) = \sum_{k\in S} p_t(i,k)p_s(k,j),$$

   that is, in compact form,

$$P(t+s) = P(t)P(s). \tag{1.3}$$

2. Also, if we plug $t=0$ in the definition of $\{P(t)\}_{t\geq 0}$, we get

$$P(0) = I,$$

   where I is the identity matrix.

The distribution at time t of $X(t)$ is the vector $\mu(t) = \{\mu_t(i)\}_{i\in S}$, where $\mu_t(i) = P(X(t) = i)$. As before, it is obtained from the initial distribution by the formula

$$\mu(t)^T = \mu(0)^T P(t).$$

At this point we can introduce the second main component - the first one was the transition semi-group - of a continuous time Markov chain: the infinitesimal generator. First of all, let us assume an additional hypothesis: suppose that the semi-group is a continuous function w.r.t. time at the origin, that is

$$\lim_{h\to 0} P(h) = P(0) = I,$$

where the convergence therein is pointwise and for each entry. This implies the continuity at any time $t \geq 0$, i.e.

$$\lim_{h\to 0} p_{t+h}(i,j) = p_t(i,j),$$

for all states $i,j \in S$. We can now enunciate the following

Proposition 1.1. *Let $\{P(t)\}_{t\geq 0}$ be a continuous transition semi-group on the countable state space S. For any state i there exists*

$$q(i) := \lim_{h\downarrow 0} \frac{1-p_h(i,i)}{h} \in [0,\infty], \tag{1.4}$$

and for any pair i,j of different states, there exists

$$q(i,j) := \lim_{h\downarrow 0} \frac{p_h(i,i)}{h} \in [0,\infty). \tag{1.5}$$

If for all states $i \in S$, $q(i) < \infty$, we say that the semi-group $\{P(t)\}$ is *stable*; if for all states $i \in S$, $q(i) = \sum_{i \neq j} q(i,j)$, it is called *conservative*. Unless otherwise specified, we will always assume that our semi-group $\{P(t)\}$ is both stable and conservative.
For each state $i \in S$, we set

$$q(i,i) := -q(i).$$

Definition 1.3. The numbers $q(i,j)$ are called the *local characteristics* of the semi-group $\{P(t)\}_{t \geq 0}$. The matrix

$$A := \{q(i,j)\}_{i,j \in S}$$

is called the *infinitesimal generator* of the semi-group.

With our notation, it can be proved (see [19], chapter 5) that

$$q(i,j) = \begin{cases} -\lambda(i), & \text{if } i = j \\ \lambda(i)\, p(i,j), & \text{if } i \neq j \end{cases} \tag{1.6}$$

where $\lambda(i)$, $i \in S$, are the holding time parameters, and $p(i,j)$ are the components of the transition matrix $\mathbf{P}$ of the discrete-time Markov chain used to construct $\{X(t)\}_{t \geq 0}$. Furthermore, for any i,

$$\sum_{j} q(i,j) = -\lambda(i) + \sum_{j \neq i} \lambda(i)\, p(i,j) = -\lambda(i) + \lambda(i) = 0$$

so that row sums are always equal 0.
Note that $\lambda(i)$, $i \in S$, and $\mathbf{P}$ determine A and vice versa. Given A, we obtain $\lambda(i)$ by denying the main diagonal entries and, for $i \neq j$,

$$p(i,j) = -\frac{q(i,j)}{q(i,i)}.$$

Observe now that in view of the definition of the infinitesimal generator and the expressions (1.4) and (1.5), one can write

$$A = \lim_{h \downarrow 0} \frac{P(h) - P(0)}{h},$$

which, together with the following definition

$$P'(t) := \{\frac{d}{dt}\, p_t(i,j)\}_{i,j \in S},$$

leads us to

Proposition 1.2. *For all $i, j \in S$ we have $p(i, j)(t)$ differentiable and the derivative is continuos. At $t = 0$, the derivative is*

$$P'(0) = A. \tag{1.7}$$

Also the backward differential equation *holds:*

$$P'(t) = AP(t), \tag{1.8}$$

i.e.

$$\frac{d}{dt} p_t(i, j) = \sum_k q(i, k) p_t(k, j).$$

Remark 1. Another, maybe less intuitive, way to get (1.8) is from the alternative expression of the transition semi-group given by

$$p_t(i, j) = \delta_{ij} e^{-\lambda(i)t} + \int_0^t \lambda(i) e^{-\lambda(i)s} \sum_{k \neq i} p(i, k)\, p_{t-s}(k, j) ds, \quad i, j \in S \tag{1.9}$$

with the same notation as above and, as usual, $\delta_{ij} = 1$ if $i = j$ and 0 otherwise.
If we consider the integral in (1.9), we see that the integrand is not only bounded on finite intervals, but it is also continuous, and hence the integral is a continuously differentiable function of t. This shows that $p_t(i, j)$ is absolutely continuous and continuously differentiable. Differentiating (1.9), after some calculations, one arrive at the desired result.

Remark 2. We can interpret (1.7) in terms of *flow rates of probability.* In fact, using (1.6), we have

$$-\lambda(i) = p'_0(i, i) = \lim_{t \downarrow 0} \frac{p_t(i, i) - p_0(i, i)}{t} = \lim_{t \downarrow 0} \frac{p_t(i, i) - 1}{t},$$

therefore

$$1 - p_t(i, i) = \lambda(i)\, t + o(t).$$

Since $1 - p_t(i, i)$ is almost (modulo the term $o(t)$) a linear function of t, we have

$$\lambda(i)\, t \approx \text{probability the system leaves } i \text{ before } t,$$

and hence we can interpret

$$\lambda(i) \approx \text{flow rate for the probability the system leaves } i \text{ before } t.$$

The same conclusion holds for $i \neq j$ and $\lambda(i) p(i, j)$ instead of $\lambda(i)$.

Remark 3. Note that just as $\{X_n\}$,$\{E_n\}$ carry the same information as $\{X_t\}_{t\geq 0}$, we have that $\{P(t)\}_{t\geq 0}$ carries the same analytic information as $\mathbf{P}$,$\{\lambda(i), i \in S\}$. In fact, given $\{P(t)\}_{t\geq 0}$, we have $A = P'(0)$; while, given $\mathbf{P}$,$\{\lambda(i), i \in S\}$, we compute A and then solve the backward differential equation to get $P(t)$.

Finally, there is a companion equation to the (1.8), called forward differential equation, which is obtained by conditioning on the last jump before time t. For regular processes such a last jump exists, but if explosions are possible the last jump may fail to exist. The forward equation is obtained formally by using the Chapman-Kolmogorov equation (1.3) with $t > 0$, $s > 0$, differentiating with respect to s to get

$$P'(t+s) = P(t)P'(s);$$

setting $s = 0$ yields finally

$$P'(t) = P(t)A. \tag{1.10}$$

When the state space S is finite, both the backward and forward equations have the formal solution

$$P(t) = e^{At}, \tag{1.11}$$

with the usual notation for the matrix exponential function given by

$$e^{At} = \sum_{n=0}^{\infty} \frac{t^n A^n}{n!}.$$

Now that we have the basics of Markov chains in continuous times, we can deal with one of the main concepts related to Markov chains: stationary measures.

1.2 Stationary measures

In this section we are going to approach stationary distributions, invariant measures and their main properties. As we will see, the main reason why these quantities are so important is that they control the long run behavior of the chain.
From now on we will use the notations of the previous section, and in particular we will always refer to $\{X(t)\}_{t\geq 0}$ as the continuous-time homogeneous Markov chain. Let $B \subset S$, we denote with

$$T_B = \inf\{t > 0 \,:\, X_t \in B\} \tag{1.12}$$

the hitting time of B of the Markov chain $\{X(t)\}_{t\geq 0}$, where we put $\inf \emptyset = \infty$. Abusing a bit the notation, we set $T_j = T_{\{j\}}$. Let us now list a series of useful definitions:

1. For $i, j \in S$, we say that j is *accessible* from i if $P_i(T_j < \infty) > 0$.

2. States i and j *communicate* if i is accessible from j and vice versa.

3. A Markov chain is *irreducible* if for any $i, j \in S$ we have that i is accessible from j.

4. The period d_i of a state $i \in S$ is, by definition,

$$d_i = \gcd\{n \geq 1 : p^{(n)}(i,i) > 0\}$$

with the convention $d_i = +\infty$ if there is no $n \geq 1$ with $p^{(n)}(i,i) > 0$. If $d_i = 1$ we say that the state i is *aperiodic*. Note that here we use the notation $p^{(n)}(i,j)$, $i, j \in S$, $n \in \mathbb{N}$, denoting the probability in the embedded chain $\{X_n\}$ to pass from i to j in n steps.

5. A state $i \in S$ is called *recurrent* if the chain returns to i with probability 1 in a finite number of steps, i.e. if $P_i(T_i < \infty) = 1$. In the contrary case, where $P_i(T_i = \infty) < 1$, the state i is called *transient*; in that case there is a positive probability of never returning on state i. A chain is called recurrent (transient) if all its states are recurrent (transient).

6. A recurrent state $i \in S$ is called *positive* recurrent if $E_i[T_i] < \infty$, otherwise is called *null* recurrent.

7. A regular homogeneous Markov chain is called *ergodic* if it is irreducible and positive recurrent.

Given the state space S finite or countable and the semi-group $P(t)$ of $\{X(t)\}_{t\geq 0}$, a measure $\pi = \{\pi(i), i \in S\}$ on S is called *invariant* if for any $t > 0$

$$\pi^T P(t) = \pi^T. \tag{1.13}$$

If this measure is a probability distribution, then it is called a *stationary distribution*.

Remark 4. Note that since S is discrete, the term measure just denotes a family of non-negative numbers indexed by S. Moreover, for the same reason, the semi-group $P(t)$ is an (at most infinite) matrix, thus the expression in (1.13) says that π is a left eigenvector of eigenvalue 1 and can be rewritten as

$$\sum_{i\in S} \pi(i) P_i(X(t) = j) = \pi(j), \quad \forall j \in S,$$

or, better yet

$$P_\pi(X(t) = j) = \pi(j).$$

The latter notation will be crucial later.

We are now going to present the main results regarding stationary distributions.

Proposition 1.3. *If the initial distribution of the Markov chain $\{X(t)\}$ is π, i.e.*

$$P(X(0) = j) = \pi(j), \quad j \in S,$$

and π is a stationary distribution, then $\{X(t)\}_{t \geq 0}$ is a stationary process. Thus for any $k \geq 0$ and $s > 0$, $0 < t_1 < \cdots < t_k$

$$(X(t_i), 1 \leq i \leq k) \overset{d}{=} (X(t_i + s), 1 \leq i \leq k).$$

In particular, for any $t \geq 0$, $j \in S$,

$$P(X(t) = j) = \pi(j)$$

is independent of t.

We are ready to introduce a concept that we will carry with us throughout the paper

Definition 1.4. Let $\{X(t)\}$ a Markov chain with an irreducible embedded (discrete-time) Markov chain $\{X_n\}$. A probability distribution $\{L = L_i, i \in S\}$ on S is called a *limit distribution* for the process $\{X(t)\}$ if for all $i, j \in S$

$$\lim_{t \to \infty} P(X(t) = j \mid X(0) = i) = \lim_{t \to \infty} p_t(i, j) = L_j.$$

And its main interest is due to the following:

Proposition 1.4. *A limit distribution is a stationary distribution.*

Proof. We have, for any $s > 0$

$$\pi(j) = \lim_{t \to \infty} p_{t+s}(i, j) = \lim_{t \to \infty} \sum_{k \in S} p_t(i, k) p_s(k, j).$$

A priori we cannnot interchange limit and sum because the dominated convergence does not apply when S are infinite space since $\{p_{kj}(s), k \in S\}$ is not a probability distribution in k. If $S = \{0, 1, 2, \dots\}$, we proceed as follows:

$$\begin{aligned}
\pi(j) &\geq \lim_{t \to \infty} \sum_{k=0}^{M} p_t(i, k) p_s(k, j) \\
&= \sum_{k=0}^{M} \lim_{t \to \infty} p_t(i, k) p_s(k, j) \\
&= \sum_{k=0}^{M} \pi(k) p_s(k, j).
\end{aligned}$$

Since for all $M > 0$ and $j \in S$,

$$\pi(j) \geq \sum_{k=0}^{M} \pi(k) p_s(k, j),$$

we let $M \to \infty$ to conclude

$$\pi(j) \geq \sum_{k=0}^{\infty} \pi(k) p_s(k, j), \tag{1.14}$$

If for some j_0 we had the strict inequality, then summing (1.14) over $j \in S$ yields

$$\begin{aligned} \sum_{j \in S} \pi(j) &> \sum_{j \in S} \sum_{k \in S} \pi(k) p_s(k, j) \\ &= \sum_{k} \pi(k) \sum_{j} p_s(k, j) = \sum_{k} \pi(k) 1 = 1, \end{aligned}$$

a contradiction. Hence the strict inequality in (1.14) can happen for no j_0, thus π is a stationary distribution. □

Thus, as long as the limit distribution exists, we have existence of stationary distribution. Let us now study in a more general framework the existence and uniqueness of stationary distributions. We propose the following

Theorem 1.5. *Let $\{X(t)\}$ a continuous-time Markov chain such that $\{X_n\}$, the embedded chain, is irreducible and recurrent. Then $\{X(t)\}$ has an invariant measure π which is unique up to multiplicative factors and can be found as the unique (up to multiplicative factors) solution to the equation*

$$\pi^T A = 0. \tag{1.15}$$

Also, π satisfies

$$0 < \pi(j) < \infty, \quad \forall j \in S.$$

Moreover, a stationary distribution exists for $\{X(t)\}$ iff

$$\sum_{i \in S} \pi(i) < \infty,$$

in which case

$$\{\frac{\pi(i)}{\sum_{i \in S} \pi(i)}, i \in S\}$$

is the stationary distribution.

Furthermore, let us recall two main results regarding ergodic chains and stationary distributions: the first one is about the limiting behavior of homogeneous continuous-time Markov chains

Theorem 1.6. *Let $\{X(t)\}_{t\geq 0}$ be a regular, ergodic Markov chain with state space S and transition semi-group $\{P(t)\}_{t\geq 0}$. Then, for all $i, j \in S$,*

$$\lim_{t\to\infty} p_t(i,j) = \pi(j),$$

where π is the (unique) stationary distribution.

Note that the only difference between this theorem and Proposition 1.4 is the uniqueness of the stationary distribution, given from the ergodicity assumption. The second result is a generalization of the LLN to Markov chains: the Ergodic Theorem

Theorem 1.7. *Let $\{X(t)\}_{t\geq 0}$ be ergodic and let π be its stationary distribution. Then*

$$\lim_{t\uparrow\infty} \int_0^t f(X(s))\, ds = \sum_{i\in S} f(i)\pi(i), \quad P_\mu - a.s. \tag{1.16}$$

for all initial distributions μ an all $f : S \to \mathbb{R}$ such that $\sum_{i\in S} |f(i)|\pi(i) < \infty$.

At this point we have enough knowledge to begin developing the core of this paper. First of all, we are interested in a particular type of state space: those that have at least one absorbing state. We need some other definitions:

1. A set of states $C \subset S$ is *closed* if for any $i \in C$ we have $P_i(T_{C^c} = \infty) = 1$. So, if the chain starts in C, it never escapes outside C.
2. If $\{j\}$ is closed, we call the state j *absorbing.*

Speaking in terms of our Markov chain $\{X(t)\}$, this means that, taking j as absorbing state, we have that if $X(s) = j$ for some s, then $X(t) = j$ for any $t \geq s$. Since all the properties concerning the states are peculiar to the subordinate chain $\{X_n\}$ of $\{X(t)\}$, one can state these two equivalent criteria:

- C is closed iff
$$\text{for all } i \in C, j \in C^c : p(i,j) = 0.$$
- $j \in S$ is absorbing iff
$$p(j,j) = 0,$$

where we recall that $p(i,j)$ are the elements of the transition matrix $\mathbf{P}$.
So far, we only listed most of the basic property concerning Markov chains in continuous time and stationary distributions. Excluding the mathematical beauty that surrounds them, without any kind of interpretation they are just meaningless tools. We want to give them a proper environment, but before that let us develop a last instrument that bridges the gap between the behavior of a Markov chain, its stationary distribution and the absorbing states. We are talking about the quasi-stationary distributions.

Chapter 2

Quasi-stationary distributions

In this chapter we are going to introduce the quasi-stationary distributions (QSDs) for Markov processes, giving some intuition and reporting the main properties. We refer largely to the work of Collet, Martínez and San Martín [4], for a general introduction to the theme; while Méléard and Villemonais [17], for a precise treatment of QSDs focused on models derived from population dynamics. We will adopt a general approach and, in order to give an idea of what we are going to cover next, we will make a continuous connection between the general theory and its application to population dynamics.
Considering the scenario explained in the previous chapter, let us introduce the basic idea of QSDs. Suppose to interpret the absorbing states as states to be avoided, just because they would make our chain trivial after the first meeting. Therefore we should study the process conditioned by the non-achievement of the absorbing states. In our discrete-space-continuous-time context, this is equivalent to "delete" from the (at most infinite) matrix of the transition semi-group $\{p_t(i,j)\}_{i,j \in S}$ and of the infinitesimal generator A the rows and columns corresponding to the absorbing states. If now we consider the new semi-group and generator, adopting the theory discussed in section 1.2, we will find a kind of stationary measure for the process restricted to non absorbing states, with the difference that the eigenvalue corresponding to the left eigenvector is strictly less than 1. Such new left eigenvector is the QSD for the conditioned process and, using a powerful result of linear algebra (Perron-Frobenius Theorem) we will show existence and uniqueness of the QSD related to Markov processes.
Before starting to be more formal, we need to make a premise about the notation that we will use during the following sections. Until now we have always supposed to work with a (discrete or continuous-time) Markov chain in a discrete state space, instead, throughout all this chapter we will use a more general setting, assuming S to be a generic metric space. Thus we are going to talk about Markov processes, no more chains, although all results that we are going to show obviously hold also for the sub-case of discrete spaces.

2.1 Definitions and general properties

In the framework of this theory, we consider a Markov process evolving in a domain where there is a set of forbidden -absorbing- states that constitutes a trap. The process

is said to be killed when it hits the trap, and it is assumed that this happens almost surely. We investigate the behavior of the process before being killed, more precisely we study what happens when the process is conditioned to survive for a long time.
Assume S to be Polish, that is a metric, separable and complete space, and consider its Borel σ-algebra $\mathcal{B}(S)$. Let Ω be the set of right continuous trajectories on S, indexed on $\mathbb{R}_+$. So $\omega = (\omega_s : s \in [0, \infty))$ means that $\omega_s \in S$ and $\lim_{h \to 0^+} \omega_{s+h} = \omega_s$ for all $s \geq 0$. Let $(\Omega, \mathcal{F})$ be a measurable space where $\mathcal{F}$ contains all the projections $pr_s : \Omega \to S$, $pr_s(\omega) = \omega_s$ for all $s \geq 0$. We consider a process $Z = (Z_t : t \geq 0)$ on S such that its trajectories belong to Ω. We denote by $\mathbb{F} = (\mathcal{F}_t : t \geq 0)$ a filtration of σ-algebras such that Z is adapted to $\mathbb{F}$. We assume that $Z = (Z_t : t \geq 0)$ is a Markov process taking values in S, i.e.

1. $\mathbb{P}_x(Z_0 = x) = 1$ for all $x \in S$;
2. For all $A \in \mathcal{F}$ the function $x \to \mathbb{P}_x(A)$ is $\mathcal{B}(S)$-measurable;
3. $\mathbb{P}_x(Z_{t+s} \in A|\mathcal{F}_t) = \mathbb{P}_{Z_t}(A)$ $\mathbb{P}_x$-a.s. $\forall x \in S$ and $\forall A \in \mathcal{F}$,

where $(\mathbb{P}_x, x \in S)$ is a family of probability measures defined in $(\Omega, \mathcal{F})$.
In this theory there is a set of forbidden states for the process, denoted by ∂S. We assume that $\partial S \in \mathcal{B}(S)$ and $\emptyset \neq \partial S \neq S$, this last condition is made in order to avoid trivial situations. Its complement $S^* := S \setminus \partial S$ is called the set of allowed states.

Notation. For any probability measure μ on S^*, we denote by $\mathbb{P}_\mu$ (resp. $\mathbb{E}_\mu$) the probability (resp. expectation) associated with the process Z initially distributed with respect to μ. For any $x \in S^*$, we set $\mathbb{P}_x = \mathbb{P}_{\delta_x}$ and $\mathbb{E}_x = \mathbb{E}_{\delta_x}$. More precisely:

$$\mathbb{P}_\mu = \int_{S^*} \mathbb{P}_x \, d\mu(x) \quad \text{for any } \mu \in \mathcal{P}(S^*)$$

where $\mathcal{P}(S^*)$ denotes the set of all probability measures on $(S^*, \mathcal{B}(S^*))$.

Recall from (1.12) the definition of the Markov chain's hitting time T_B, $B \subset S$. In a very similar way we let

$$T := T_{\partial S} = \inf\{t > 0 \, : \, Z_t \in \partial S\}$$

be the hitting time of ∂S, it is called the *killing time* (or extinction time in population models), where, as in the previous section, with T_B we denote the hitting time of $B \in \mathcal{B}(S)$ of the Markov process Z. Regarding the population framework, the Markov process Z represent the population size at any time (continuous or discrete) and the state space S is a subset of $\mathbb{N}$ or $\mathbb{R}_+$. We will consider then an isolated population, namely without immigration, so that the state 0 is the -only- trap because it describes the extinction of the population. Indeed, if there are no more individuals, no reproduction can occur and

the population disappears. Therefore we have $\partial S = \{0\}$ and $T = T_0$.
Going back to the general context, we will assume that there is sure killing at ∂S, i.e.

$$\forall x \in S \quad \mathbb{P}_x(T < \infty) = 1, \tag{2.1}$$

it follows that there is no explosion.

Remark 5. Note that there cannot exist stationary distributions supported on S^*. Indeed, condition (2.1) is equivalent to

$$\mathbb{P}_\mu(T < \infty) = 1 \quad \forall \mu \in \mathcal{P}(S^*),$$

and it clearly implies that

$$\mathbb{P}_x(T < \infty) > 0 \quad \forall x \in S^*.$$

Thus

$$\forall \mu \in \mathcal{P}(S^*)\, \exists\, t(\mu) : \forall t \geq t(\mu) \quad \mathbb{P}_\mu(T > t) < 1, \tag{2.2}$$

therefore, if μ was a stationary distribution with support on S^*, from (2.2) one would have that $\mathbb{P}_\mu(Z_t \in S^*) < 1$ for $t \geq t(\mu)$, which contradicts the stationarity because $\mathbb{P}_\mu(Z_0 \in S^*) = 1$. The same observation can be done in the discrete-time-discrete-space setting of Markov chains, using the Perron-Frobenius theorem and the fact that the semi-group matrix restricted to non absorbing states is sub-stochastic. We will explore this topic further below.

Since we are considering a Markov process, we need to generalize, then redefine, some fundamental tools that we introduced in the previous chapter.

Notation. We denote by $(P_t)_{t\geq 0}$ the semi-group of the process Z killed at ∂S. More precisely, for any $z \in S^*$ and f a measurable and bounded function on S^*, one defines

$$P_t f(z) := \mathbb{E}_z(f(Z_t)\mathbb{1}_{t<T}).$$

Moreover, for any finite measure μ and any bounded measurable function f, we set

$$\mu(f) = \int_{S^*} f(x)\mu(dx),$$

and we also define the finite measure μP_t by

$$\mu P_t(f) := \mu(P_t f) = \int_{S^*} \mathbb{E}_z(f(Z_t)\mathbb{1}_{t<T})\mu(dz) = \mathbb{E}_\mu(f(Z_t)\mathbb{1}_{t<T}). \tag{2.3}$$

One last note. We have made, and we will always make from now on, an abuse of notation: at the beginning of this chapter we supposed to work in a general space $(S, \mathcal{B}(S))$, therefore if we write $\mu(dx)$ and, further on, $\mathbb{P}_\mu(Z_t \in dx)$, we are tacitly saying that our space is $\mathbb{R}_+$. Also, we should use the notation of the transition kernels, but since the focus of this paper is not about these matters, we prefer to adopt this more intuitive notation.

At this point we have everything we need to define some new interesting objects. The first natural question that arises when we talk about a process Z conditioned to not be killed (up to time t) is what its distribution is. We can rewrite it, using our terminology, as follows

$$\mathbb{P}_\mu(Z_t \in A | T > t) = \frac{\mathbb{P}_\mu(Z_t \in A, T > t)}{\mathbb{P}_\mu(T > t)} \overset{(2.3)}{=} \frac{\mu P_t(\mathbb{1}_A)}{\mu P_t(\mathbb{1}_{S^*})}, \tag{2.4}$$

for any Borel subset $A \subset S^*$, where μ is the initial distribution of Z_0.
At this point we want to bring the reader's attention to what we have discussed in section 1.2: we referred to stationary distributions, limiting distributions and how to link them. We will see that in this "conditioned" context we can make a sort of parallelism. For this purpose we need to study the asymptotic behaviour of the conditional probability in (2.4), when t tends to infinity.

Definition 2.1. Let α be a probability measure on S^*. We say that α is a **quasi-limiting distribution** (QLD) for Z, if there exists a probability measure μ on S^* such that, for any measurable set $A \subset S^*$,

$$\lim_{t\to\infty} \mathbb{P}_\mu(Z_t \in A | T > t) = \alpha(A).$$

In some cases the long time behavior of the conditioned distribution can be proved to be initial state independent. This leads to the following definition.

Definition 2.2. We say that Z has an **Yaglom limit** if there exists a probability measure α on S^* such that, for any $x \in S^*$ and any measurable set $A \subset S^*$,

$$\lim_{t\to\infty} \mathbb{P}_x(Z_t \in A | T > t) = \alpha(A). \tag{2.5}$$

When it exists, the Yaglom limit is a QLD. The reverse is not true in general and (2.5) will actually not imply the same property for any initial distribution.
One might ask why, in population models, there is an interest in knowing such asymptotic behaviour. There are two main reasons: the first one is related to the fact that a Markov process with extinction which possesses a quasi-stationary distribution has a mortality plateau; while the second one is due to the fact that, if the time scale of absorption is substantially larger than the one of the quasi-limiting distribution, the process relaxes to the quasi-limiting regime after a relatively short time, and then, after a much longer period, absorption will eventually occur. Thus the quasi-limiting distribution bridges the gap between the known behavior (extinction) and the unknown time-dependent behavior of the process. See [17] for more details.
We have found in definition 2.1 that the analog of the limiting distribution, in the process conditioned on non-killing, is called quasi-limiting distribution. So we shouldn't be surprised by the following results.

Definition 2.3. Let α be a probability measure on S^*. We say that α is a **quasi-stationary distribution** (QSD) for the process killed at ∂S, if for all $t \geq 0$ and any $A \in \mathcal{B}(S^*)$,

$$\alpha(A) = \mathbb{P}_\alpha(Z_t \in A | T > t). \tag{2.6}$$

Recall that, as we saw in Remark 5, condition (2.1) ensures that a QSD cannot be stationary.
We are going to start with the main results that connect the newly defined objects. It is immediate to note that any QSD and any Yaglom limit is also a QLD; the following statement generalize what has just been stated.

Proposition 2.1. *Let $\alpha \in \mathcal{P}(S^*)$. The probability distribution α is a QLD for the process Z if and only if it is a QSD for Z.*

Proof. "$\Leftarrow$" If α is a QSD then, passing to the limit of $t \to \infty$ on both sides of the definition of QSDs, it is a QLD for Z starting with distribution α.
"$\Rightarrow$" Assume now that α is a QLD for Z and for an initial probability measure μ on S^*. Thus, for any bounded and measurable function f on S^*,

$$\begin{aligned}
\alpha(f) &= \int_{S^*} f(x)\alpha(dx) = \int_{S^*} f(x) \lim_{t\to\infty} \mathbb{P}_\mu(Z_t \in dx | T > t) \\
&= \int_{S^*} f(x) \lim_{t\to\infty} \int_{S^*} \mathbb{P}_y(Z_t \in dx | T > t)\mu(dy) = \lim_{t\to\infty} \int_{S^*} \mathbb{E}_y(f(Z_t) | T > t)\mu(dy) \\
&= \lim_{t\to\infty} \mathbb{E}_\mu(f(Z_t) | T > t) = \lim_{t\to\infty} \frac{\mathbb{E}_\mu(f(Z_t), T > t)}{\mathbb{P}_\mu(T > t)}.
\end{aligned}$$

where clearly we could interchange limit and integrals, and the two integrals because everything is finite and bounded. If we apply the latter with $f(z) = \mathbb{P}_z(T > s)$, $\forall s \geq 0$ fixed, we have

$$\mathbb{P}_\alpha(T > s) = \lim_{t\to\infty} \frac{\mathbb{E}_\mu(\mathbb{P}_{Z_t}(T > s), T > t)}{\mathbb{P}_\mu(T > t)}. \tag{2.7}$$

Developing the numerator inside the limit and using the Markov property, we get

$$\begin{aligned}
\mathbb{P}_\mu(T > t + s) &= \mathbb{E}_\mu(\mathbb{1}_{T>t+s}) = \mathbb{E}_\mu(\mathbb{E}(\mathbb{1}_{T>t+s} | \mathcal{F}_t)) = \mathbb{E}_\mu(\mathbb{E}(\mathbb{1}_{T>t+s}\mathbb{1}_{T>t} | \mathcal{F}_t)) \\
&= \mathbb{E}_\mu(\mathbb{1}_{T>t}\mathbb{E}(\mathbb{1}_{T>t+s} | \mathcal{F}_t)) = \mathbb{E}_\mu(\mathbb{1}_{T>t}\mathbb{E}_{Z_t}(\mathbb{1}_{T>t+s})) \\
&= \mathbb{E}_\mu(\mathbb{P}_{Z_t}(T > s), T > t).
\end{aligned}$$

Plugging the last equality in (2.7) one gets

$$\mathbb{P}_\alpha(T > s) = \lim_{t\to\infty} \frac{\mathbb{P}_\mu(T > t + s)}{\mathbb{P}_\mu(T > t)}. \tag{2.8}$$

Let us now consider $f(z) = \mathbb{P}_z(Z_s \in A, T > s)$, for every $s > 0$ and $A \in \mathcal{B}(S^*)$. With a calculation similar to the previous one, using again the Markov property, one obtains

$$\begin{aligned}\mathbb{P}_\alpha(Z_s \in A, T > s) &= \lim_{t\to\infty} \frac{\mathbb{P}_\mu(Z_{t+s} \in A, T > t+s)}{\mathbb{P}_\mu(T > t)} \\ &= \lim_{t\to\infty} \frac{\mathbb{P}_\mu(Z_{t+s} \in A, T > t+s)}{\mathbb{P}_\mu(T > t+s)} \frac{\mathbb{P}_\mu(T > t+s)}{\mathbb{P}_\mu(T > t)}.\end{aligned}$$

Finally, by definition of the QLD α we have that $\frac{\mathbb{P}_\mu(Z_{t+s}\in A, T>t+s)}{\mathbb{P}_\mu(T>t+s)} = \mathbb{P}_\mu(Z_{t+s} \in A | T > t+s)$ converges to $\alpha(A)$ and, by (2.8), $\frac{\mathbb{P}_\mu(T>t+s)}{\mathbb{P}_\mu(T>t)}$ converges to $\mathbb{P}_\alpha(T > s)$, when t goes to infinity. Thus, for any $s > 0$ and $A \in \mathcal{B}(S^*)$,

$$\alpha(A) = \mathbb{P}_\alpha(Z_s \in A).$$

We conclude that the probability measure α is then a QSD. □

Summing up we proved that, if there exists, the Yaglom limit is uniquely defined and it is always a QLD. Moreover, by the last proposition, we showed up that any QLD is indeed a QSD. One would like to have also the last piece of the puzzle, the one which links QSDs to Yaglom limits, but unfortunately there are processes with an infinity -a continuum!- of QSDs (see the Section 2.2 for the birth and death case), thus we deduce that there exists QSDs which aren't a Yaglom limit.
Next, one of the main facts related to QSDs is the distribution of the killing time T:

Proposition 2.2. *Let us consider a Markov process Z with absorbing states ∂S satisfying the sure-killing condition (2.1). Let α be a quasi-stationary distribution for the process. Then there exists a positive real number $\theta(\alpha)$ depending on the QSD such that*

$$\mathbb{P}_\alpha(T > t) = e^{-\theta(\alpha)t}. \tag{2.9}$$

Proof. Using the Markov property in the same way that we used just after (2.7), for all $s > 0$ we get

$$\begin{aligned}\mathbb{P}_\alpha(T > t+s) &= \mathbb{E}_\alpha(\mathbb{P}_{Z_t}(T > s)\mathbb{1}_{T>t}) \\ &= \mathbb{P}_\alpha(T > t)\mathbb{E}_\alpha(\mathbb{P}_{Z_t}(T > s)|T > t).\end{aligned}$$

By definition of QSD, we have

$$\begin{aligned}&\mathbb{E}_\alpha(\mathbb{P}_{Z_t}(T > s)|T > t) = \int_{S^*} \mathbb{P}_z(T > s)\,\mathbb{P}_\alpha(Z_t \in dz|T > t) \\ &= \int_{S^*} \mathbb{P}_z(T > s)\,\alpha(dz) = \mathbb{P}_\alpha(T > s).\end{aligned}$$

Hence we obtain that for all $s, t > 0$,

$$\mathbb{P}_\alpha(T > t + s) = \mathbb{P}_\alpha(T > t)\mathbb{P}_\alpha(T > s). \tag{2.10}$$

If we consider the function $g(t) = \mathbb{P}_\alpha(T > t)$, we note that $g(0) = 1$, $0 \leq g(t) \leq 1$ and, because of (2.1), $g(t)$ tends to 0 when t goes to infinity. It follows immediately that a function with these characteristics and with the property (2.10) can only be a negative exponential, i.e. there exists a strictly positive real number $\theta(\alpha)$ such that

$$\mathbb{P}_\alpha(T > t) = e^{-\theta(\alpha)t}.$$

□

This proposition shows us that starting from a QSD, the killing time has an exponential distribution with parameter $\theta(\alpha)$ independent of $t > 0$, given by

$$\theta(\alpha) = -\frac{\ln \mathbb{P}_\alpha(T > t)}{t}.$$

Remark 6. For discrete time, the argument in the proof of the latter proposition shows that when starting from a QSD α, the killing time at ∂S is geometrically distributed, so $\mathbb{P}_\alpha(T > n) = \kappa(\alpha)^n$ for all $n \in \mathbb{Z}_+$, where $\kappa(\alpha) = \mathbb{P}_\alpha(T > 1)$. Indeed, taking $\kappa(\alpha) \in (0, 1)$ to avoid trivial situations, if we denote with $\theta(\alpha) := -\ln \kappa(\alpha) \in (0, \infty)$ the exponential rate of survival of α, it verifies $\mathbb{P}_\alpha(T > t) = e^{-\theta(\alpha)t}$ for all $t \in \mathbb{Z}_+$.

2.2 Existence and uniqueness

Until this moment, from (2.9) and (2.6), we get that $\alpha \in \mathcal{P}(S^*)$ is a QSD for the Markov process $Z = (Z_t, t \in \mathbb{R}_+)$ if and only if $\exists\, \theta(\alpha) > 0$ such that for every $A \in \mathcal{B}(S^*)$ and $t \geq 0$

$$\mathbb{P}_\alpha(Z_t \in A, T > t) = \alpha(A)\, e^{-\theta(\alpha)t}.$$

The following statement gives a necessary condition for the existence of QSDs in terms of exponential moments of the hitting time T.

Proposition 2.3. *Assume that α is a QSD. Then, for any $0 < \theta < \theta(\alpha)$,*

$$\mathbb{E}_\alpha(e^{\theta T}) < +\infty. \tag{2.11}$$

In particular, there exists a $z \in S^*$ such that $\mathbb{E}_z(e^{\theta T}) < +\infty$.

Proof. We just compute the exponential moments in the continuous and discrete time setting and we show that it is finite under the above hypothesis.

In the continuous time setting, by (2.9), the killing time T under $\mathbb{P}_\alpha$ has an exponential distribution with parameter $\theta(\alpha)$. Thus, for any $\theta < \theta(\alpha)$, the mean

$$\mathbb{E}_\alpha(e^{\theta T}) = \frac{\theta(\alpha)}{\theta(\alpha) - \theta}$$

is finite and well defined. Same argument in the discrete time case where, as seen in Remark 6, T under $\mathbb{P}_\alpha$ has a geometric distribution with parameter $e^{-\theta(\alpha)}$. So we deduce that, for any $\theta < \theta(\alpha)$, the mean

$$\mathbb{E}_\alpha(e^{\theta T}) = \frac{1 - e^{-\theta(\alpha)}}{e^{-\theta} - e^{-\theta(\alpha)}}$$

is finite and well defined. Finally, since $\mathbb{E}_\alpha(e^{\theta T}) = \int_{S^*} \mathbb{E}_z(e^{\theta T})\alpha(dz)$, the finiteness of the integral implies that at least for a $z \in S^*$ we have $\mathbb{E}_z(e^{\theta T}) < \infty$. $\square$

The exponential moment condition can be written in the following way.

Proposition 2.4. *We have the equality*

$$\theta_z^* := \sup\{\theta : \mathbb{E}_z(e^{\theta T}) < +\infty\} = \liminf_{t\to\infty} -\frac{1}{t} \ln \mathbb{P}_z(T > t),$$

and a necessary condition for the existence of a QSD is the existence of a positive exponential rate of survival:

$$\exists\, z \in S^* : \theta_z^* > 0.$$

The necessary condition for the existence of a QSD is given by the definition of θ_z^* with the $z \in S^*$ from Proposition 2.3. For a complete proof see Proposition 2.4 of [4].

Remark 7. Note that if α is a QSD then $\mathbb{E}_\alpha(e^{\theta(\alpha)T}) = \infty$. Then, if $\theta > 0$ satisfies the condition

$$\sup\{\mathbb{E}_z(e^{\theta T}) : z \in S^*\} < \infty,$$

there can not be a QSD α with $\theta = \theta(\alpha)$, because otherwise we should have

$$\infty = \mathbb{E}_\alpha(e^{\theta(\alpha)T}) \le \sup\{\mathbb{E}_z(e^{\theta T}) : z \in S^*\} < \infty.$$

It is useful to mention that the article of Ferrari, Kesten, Martìnez and Picco [13]. In Theorem 1.1, they proved that in the case of an irreducible continuous time Markov chain with state space $\mathbb{N}$ such that $\lim_{x\to\infty} \mathbb{P}_x(T \le t) = 0$, $\forall t \ge 0$, the existence of the moment (2.11) for some $x \in \mathbb{N}$ and some $\lambda > 0$ is also a sufficient condition for the existence of a quasi-stationary distribution. Here with T we mean the hitting time of the state 0. It is actually not true in any case, as shown by the following counter-example. Let Z be a continuous time random walk on $\mathbb{N}$ reflected on 1 and killed at rate 1, i.e.

$\mathbb{P}(T > t) = e^{-t}$, $\forall t \geq 0$. Thus for any $\lambda > 0$ and any probability measure μ on $\mathbb{N}$, $\mathbb{E}_\mu(e^{\lambda T})$ is finite because $\mathbb{E}_\mu(e^{\lambda T}) = \int_0^{+\infty} -e^{\lambda t}e^{-t}\,dt = \frac{1}{1-\lambda}$. Nevertheless the conditional distribution $\mathbb{P}_x(Z_t \in \cdot | T > t)$ is the distribution of a standard continuous time random walk reflected on 1, which converges to 0 as t tends to infinity. In particular Z has no QLD and thus no QSD.
Another point of view relating the characterization of the QSDs is the one which involves the spectral properties of the semi-group (P_t) and the associated infinitesimal generator A. First of all recall that we are working in a metric space S, therefore A will no longer be a matrix, but an operator with a certain domain $\mathcal{D}(A)$. The next proposition links the existence of QSDs for the process Z and the spectral properties of the dual of A.

Proposition 2.5. *Let α be a probability measure on S^*. We assume that there exists a set $D \subset \mathcal{D}(A)$ such that, for any measurable subset $B \subset S^*$, there exists a uniformly bounded sequence (f_n) in D converging pint-wisely to $\mathbb{1}_B$.*
Then α is a quasi-stationary distribution if and only if there exists $\theta(\alpha) > 0$ such that

$$\alpha(Af) = -\theta(\alpha)\alpha(f), \quad \forall f \in D. \tag{2.12}$$

Proof. "$\Rightarrow$" Let α be a QSD for Z. By definition of a QSD, and recalling (2.4), we have, for every $B \in \mathcal{B}(S^*)$,

$$\alpha(B) = \frac{\alpha P_t(\mathbb{1}_B)}{\alpha P_t(\mathbb{1}_{S^*})}. \tag{2.13}$$

By (2.9), there exists $\theta(\alpha) > 0$ such that for each $t > 0$,

$$\alpha P_t(\mathbb{1}_{S^*}) = \mathbb{P}_\alpha(T > t) = e^{-\theta(\alpha)t}.$$

We deduce that, for any measurable set $B \subset S^*$, $\alpha P_t(\mathbb{1}_B) = e^{-\theta(\alpha)t}\alpha(B)$, which is equivalent to $\alpha P_t = e^{-\theta(\alpha)t}\alpha$. By Kolmogorov's forward equation (1.10) and by assumption on D, we have

$$\left|\frac{\partial P_t f}{\partial t}(x)\right| = |P_t Af(x)| \overset{(1)}{\leq} \|Af\|_\infty < +\infty, \quad \forall f \in D, \tag{2.14}$$

where in (1) we just used that $|P_t Af(x)| = |\mathbb{E}_x(Af(Z_t))\mathbb{1}_{T>t}| \leq \|Af\|_\infty$,and the finiteness of the latter comes from the fact that $f \in D \subset \mathcal{D}(A)$. In particular, one can differentiate $\alpha P_t f = \int_{S^*} P_t f \alpha(dx)$ under the integral sign, which implies that

$$\frac{\partial}{\partial t}\alpha P_t f = \int_{S^*} \frac{\partial}{\partial t} P_t f \alpha(dx) = \int_{S^*} P_t Af(x)\alpha(dx) = \alpha P_t(Af) = e^{-\theta(\alpha)t}\alpha(Af),$$

on the other hand we have

$$\frac{\partial}{\partial t}\alpha P_t f = \frac{\partial}{\partial t} e^{-\theta(\alpha)t}\alpha(f) = -\theta(\alpha)e^{-\theta(\alpha)t}\alpha(f).$$

Putting all together, we get the result

$$\alpha(Af) = -\theta(\alpha)\alpha(f), \quad \forall f \in D.$$

"$\Leftarrow$" Assume now that $\alpha(Af) = -\theta(\alpha)\alpha(f)$, for all $f \in D$. By Kolmogorov's backward equation (1.8) and a similar argument as (2.14), we have

$$\frac{\partial \alpha(P_t f)}{\partial t} = \alpha(AP_t f) \overset{(2)}{=} -\theta(\alpha)\alpha P_t(f), \quad \forall f \in D,$$

where in (2) we used the relation (2.12) with $P_t f$ as f. Reasoning in the same way as before, we get

$$\frac{\partial \alpha(P_t f)}{\partial t} = \frac{\partial}{\partial t} e^{-\theta(\alpha)t}\alpha(f) = -\theta(\alpha)e^{-\theta(\alpha)t}\alpha(f).$$

Thus we deduce that

$$\alpha(P_t f) = e^{-\theta(\alpha)t}\alpha(f), \quad \forall f \in D.$$

By assumption, there exists, for any measurable subset $B \subset S^*$, an uniformly bounded sequence (f_n) in D which converges point-wisely to $\mathbb{1}_B$. Finally, from the uniform boundedness of (f_n) and $\mathbb{P}(S^*) = 1$, we deduce by dominated convergence that

$$\alpha P_t(\mathbb{1}_B) = e^{-\theta(\alpha)t}\alpha(f).$$

This implies immediately that α is a QSD for the process Z, because the relation (2.13) is fulfilled. □

Long time limit of the extinction rate

Here we propose an interesting application of the latter proposition. We want to study a quantity of interest in population's dynamics: the long time behaviour of the killing or extinction rate. In the demography setting, the process Z models the vitality of some individual and t its physical age. Thus T is the death time of this individual, corresponding to the hitting time of the age zero.
The definition of the extinction rate depends on the time setting:

- In the discrete time setting, the extinction rate of Z starting from μ at time $t \geq 0$ is defined by
$$r_\mu(t) = \mathbb{P}_\mu(T = t+1 | T > t).$$
- In the continuous time setting, the extinction rate of Z starting from μ at time $t \geq 0$ is defined by
$$r_\mu(t) = -\frac{\frac{\partial}{\partial t}\mathbb{P}_\mu(T > t)}{\mathbb{P}_\mu(T > t)},$$
when the derivative exists and is integrable with respect to μ.

The QSDs play a main role in this framework. Indeed, recall by Proposition 2.2 that if α is a QSD, then $\mathbb{P}_\alpha(T > t) = e^{-\theta(\alpha)t}$. Moreover the extinction rate $r_\alpha(t)$ is constant and given by

$$r_\alpha(t) = \begin{cases} 1 - e^{-\theta(\alpha)} & \text{in the discrete time setting} \\ \theta(\alpha) & \text{in the continuous time setting} \end{cases}, \quad \forall t \geq 0.$$

In the next proposition, we prove that the existence of a QLD for Z started from μ implies the existence of a long term mortality plateau.

Proposition 2.6. *Let α be a QLD for Z, initially distributed with respect to a probability measure μ on S^*. In the continuous time setting, we assume moreover that there exists $h > 0$ such that $A(P_h\mathbb{1}_{S^*})$ is well defined and bounded. In both time settings, the rate of extinction converges in the long term:*

$$\lim_{t\to\infty} r_\mu(t) = r_\alpha(0).$$

We remark the fact that if we had had a QSD instead of a QLD, the result would have been trivial because we would have had $r_\mu(t)$ constant in both time settings. We refer to the introduction of Steinsaltz-Evans [11] for a discussion of the notion of QSD in relationship with mortality plateaus.

Proof. In the discrete time setting, by the semi-group property and the definition of a QLD, we have

$$\begin{aligned} r_\mu(t) &= \mathbb{P}_\mu(T = t+1 | T > t) = 1 - \mathbb{P}_\mu(T > t+1 | T > t) \\ &= 1 - \frac{\mu P_t(P_1\mathbb{1}_{S^*})}{\mu P_t(\mathbb{1}_{S^*})} \xrightarrow[t\to\infty]{} 1 - \alpha(P_1\mathbb{1}_{S^*}) = \mathbb{P}_\alpha(T = 1 | T > 0) = r_\alpha(0). \end{aligned}$$

The limit is by definition the extinction rate at time 0 of Z starting from α. In the continuous time setting, by the Kolmogorov forward equation (1.10), we have

$$\frac{\partial}{\partial t} P_{t+h}\mathbb{1}_{S^*}(x) = P_t A(P_h\mathbb{1}_{S^*})(x), \quad \forall x \in S^*.$$

Since by hypothesis $A(P_h\mathbb{1}_{S^*})$ is assumed to be bounded, we can differentiate under the integral sign, and we deduce that

$$\frac{\partial}{\partial t} \mu P_{t+h}(\mathbb{1}_{S^*}) = \mu P_t A(P_h\mathbb{1}_{S^*}).$$

Then

$$\frac{\frac{\partial}{\partial t}\mu P_{t+h}(\mathbb{1}_{S^*})}{\mu P_t(\mathbb{1}_{S^*})} = \frac{\mu P_t A(P_h\mathbb{1}_{S^*})}{\mu P_t(\mathbb{1}_{S^*})} \xrightarrow[t\to\infty]{} \alpha(AP_h\mathbb{1}_{S^*}) = -\theta(\alpha)\alpha(P_h\mathbb{1}_{S^*}),$$

by the definition of QLD and by Proposition 2.5. We also have

$$\frac{\mu(P_{t+h}\mathbb{1}_{S^*})}{\mu(P_t\mathbb{1}_{S^*})} \xrightarrow[t\to\infty]{} \alpha(P_h\mathbb{1}_{S^*}).$$

Finally, by multiplying and dividing by $\mu(P_t\mathbb{1}_{S^*})$, we get

$$r_\mu(t+h) = -\frac{\frac{\partial}{\partial t}\mu P_{t+h}(\mathbb{1}_{S^*})}{\mu P_{t+h}(\mathbb{1}_{S^*})} \xrightarrow[t\to\infty]{} \theta(\alpha),$$

which allows us to conclude the proof. □

2.3 The finite case

From now on we will consider a Markov process $(Z_t)_{t\geq 0}$ evolving in continuous time in a discrete state space $S = \{0, 1, \ldots, N\}$, $N \geq 1$, and we will assume that 0 is its unique absorbing state, that is $\partial S = \{0\}$. The semi-group $(P_t)_{t\geq 0}$ is the sub-markovian[1] semi-group of the killed process and we still denote by A the associated infinitesimal generator. In this finite state space case, the operators A and P_t are matrices, and a probability measure on the finite space S^* is a vector of non-negative entries whose sum is equal to 1. The results of this section have been originally proved by Darroch and Seneta in [6] and [7].

2.3.1 Perron-Frobenius Theorem

In this section, we will talk about one of the core tools used in literature regarding QSDs: the Perron-Frobenius Theorem. We will not propose the original theorem, but an adapted version taking as positive matrix the sub-markovian semi-group (P_t). For the proof of the Perron-Frobenius Theorem, we refer to Gantmacher [14].

Theorem 2.7 (Perron-Frobenius Theorem). *Let (P_t) be a sub-markovian semi-group on $\{1, \ldots, N\}$ such that the entries of P_{t_0} are positive for $t_0 > 0$. Thus, there exists a unique positive eigenvalue ρ, which is the maximum of the modulus of the eigenvalues, and there exists a unique left-eigenvector α such that $\alpha(i) > 0$ for every $i = 1, \ldots, N$ and $\sum_{i=1}^N \alpha(i) = 1$, and there exists a unique right-eigenvector π such that $\pi(i) > 0$ for every $i = 1, \ldots, N$ and $\sum_{i=1}^N \alpha(i)\pi(i) = 1$, satisfying*

$$\alpha P_{t_0} = \rho\alpha \quad \textit{and} \quad P_{t_0}\pi = \rho\pi. \tag{2.15}$$

[1] Is related to the fact that we are considering the process conditioned to non-absorbing states. This means that $\|P_t\| < 1$ or, better, the is at least a (since in our case P_t is a matrix) row whose sum is strictly less than 1.

In addition, since (P_t) is a sub-markovian semi-group, $\rho < 1$ and there exists $\theta > 0$ such that $\rho = e^{-\theta}$. Therefore

$$P_t = e^{-\theta t} B + \Theta(e^{-\chi t}), \tag{2.16}$$

where B is the matrix defined by $B_{ij} = \pi(i)\alpha(j)$, $\chi > \theta$ and $\Theta(e^{-\chi t})$ denotes a matrix such that none of the entries exceeds $Ce^{-\chi t}$, for some constant $C > 0$.

Therefore Theorem 2.7 gives us a complete description of the spectral properties of P_t and A. The following result links the latter theorem with the theory of QSDs.

Theorem 2.8. *Assume that Z is an irreducible and aperiodic process before extinction, which means that there exists $t_0 > 0$ such that the matrix P_{t_0} has only positive entries (in particular, it implies that P_t has positive entries for $t > t_0$)*[2]*. Then the Yaglom limit α exists and is the unique QSD of the process Z_t.*
Moreover, denoting by $\theta(\alpha)$ the extinction rate associated to α (see Proposition 2.2), there exists a probability measure π on S^* such that, for any $i, j \in S^*$,

$$\lim_{t\to\infty} e^{\theta(\alpha)t}\mathbb{P}_i(Z_t = j) = \pi(i)\alpha(j),$$

and

$$\lim_{t\to\infty} \frac{\mathbb{P}_i(T > t+s)}{\mathbb{P}_j(T > t)} = \frac{\pi(i)}{\pi(j)} e^{-\theta(\alpha)s}.$$

Proof. Applying Perron-Frobenius Theorem to the sub-markovian semi-group $(P_t)_{t\geq 0}$, it is immediate from (2.16) that htere exists $\theta > 0$ and a probability measure α on S^* such that, for any $i, j \in S^*$,

$$e^{\theta t}\mathbb{P}_i(Z_t = j) = e^{\theta t}[P_t]_{ij} = \pi(i)\alpha(j) + \Theta(e^{-(\chi-\theta)t}). \tag{2.17}$$

We now observe that

$$\sum_{j\in S^*} \mathbb{P}_i(Z_t = j) = \sum_{j\in S^*} (\mathbb{P}_i(Z_t = j, T > t) + \underbrace{\mathbb{P}_i(Z_t = j, T < t)}_{=\delta_{j,0}}) = \mathbb{P}_i(T > t),$$

since Z_t is defined on $S^* = S \setminus \partial S$. Thus, using the fact that $\sum_{i\in S^*}\alpha(i) = 1$, summing over $j \in S^*$ in (2.17), we deduce that

$$e^{\theta t}\mathbb{P}_i(T > t) = \pi(i) + \Theta(e^{-(\chi-\theta)t}). \tag{2.18}$$

[2] Recall from Section 1.2 that Z is irreducible if for every $i, j \in S$, $\mathbb{P}_i(T_j < \infty) > 0$. Thus $\forall i, j \in S\ \exists t_0 > 0, k \in \mathbb{N}$, $\exists s_1, \ldots, s_k \in \mathbb{R}_+$ with $t_0 < s_1 < \cdots < s_k$, and $\exists i_1, \ldots, i_k \in S$ such that $p_{t_0}(i, i_1)p_{s_1}(i_1, i_2)\cdots p_{s_k}(i_k, j) > 0$. Moreover, by aperiodicity there is no cyclic structure, thus the latter inequality does not depend on the choice of $i_1, \ldots, i_k$.

It follows that, for any $i, j \in S^*$,

$$\begin{aligned}\mathbb{P}_i(Z_t = j | T > t) &= \frac{\mathbb{P}_i(Z_t = j)}{\mathbb{P}_i(T > t)} = \frac{e^{\theta t}\mathbb{P}_i(Z_t = j)}{e^{\theta t}\mathbb{P}_i(T > t)} \\ &= \frac{\pi(i)\alpha(j) + \Theta(e^{-(\chi-\theta)t})}{\pi(i) + \Theta(e^{-(\chi-\theta)t})} \xrightarrow[t\to\infty]{} \alpha(j).\end{aligned}$$

Thus the Yaglom limit exists and is equal to α. Since S is finite, in the following we can interchange the sum with the limit, therefore we have, for any initial distribution ν on S^*

$$\lim_{t\to\infty} \mathbb{P}_\nu(Z_t = j|T > t) = \sum_{i\in S^*} \nu(i) \lim_{t\to\infty} \mathbb{P}_i(Z_t = j|T > t) = \sum_{i\in S^*} \nu(i)\alpha(j) = \alpha(j).$$

For the arbitrariness of the choice of ν, we deduce that the Yaglom limit is the unique QLD of Z, and thus it is its unique QSD. By Proposition 2.2, we have $\alpha P_1(\mathbb{1}_{S^*}) = e^{-\theta(\alpha)}$. If we suppose that $t_0 < 1$, by (2.15) we have

$$\alpha P_1 = \rho\alpha \Rightarrow e^{-\theta(\alpha)} = \alpha P_1(\mathbb{1}_{S^*}) = \rho\alpha(\mathbb{1}_{S^*}) = e^{-\theta},$$

so that $\theta = \theta(\alpha)$. Finally, using both (2.17) and (2.18), we get

$$\frac{\mathbb{P}_i(T > t+s)}{\mathbb{P}_j(T > t)} = \frac{e^{-\theta(t+s)}\mathbb{P}_i(T > t+s)}{e^{-\theta t}\mathbb{P}_j(T > t)} \cdot e^{-\theta s} \xrightarrow[t\to\infty]{} \frac{\pi(i)}{\pi(j)} e^{-\theta s},$$

which allows us to conclude the proof. □

Remark 8. One can deduce from (2.17) and (2.18) that there exists a positive constant C_A such that

$$\sup_{j\in S^*, i\in S^*} |\mathbb{P}_i(Z_t = j|T > t) - \alpha(j)| \le C_A e^{-(\chi-\theta(\alpha))t},$$

where the quantity $\chi - \theta(\alpha)$ is the spectral gap of A, i.e. the distance between the first and second eigenvalue of A. Note that if $e^{-\theta}$ is an eigenvalue of P_t, then $-\theta$ is an eigenvalue of A because we are working with S finite, so (1.11) is satisfied. We want to compare the last inequality with (2.9): if the time scale $\chi - \theta(\alpha)$ of the convergence to the quasi-limiting distribution is substantially bigger than the time scale of absorption ($\chi - \theta(\alpha) \gg \theta(\alpha)$), the process will relax to the QSD after a relatively short time, and after a much longer period, extinction will occur. On the contrary, if $\chi - \theta(\alpha) \ll \theta(\alpha)$, then the extinction happens before the process had time to relax to the quasi-limiting distribution. In intermediate cases, where $\chi - \theta(\alpha) \approx \theta(\alpha)$, the constant C_A, which depends on the whole set of eigenvalues and eigenfunctions of A, plays a main role which needs further investigations.
In [17] one can find some interesting examples, together with numerical simulations.

2.3.2 The Q-process

Let us now study the marginal distributions of the trajectories that survive forever, i.e. conditioned to never be extinct.

Theorem 2.9. *Assume that we are in the hypothesis of Theorem 2.8. For any* $i_0, i_1, \dots, i_k \in S^*$*, any* $0 < s_1 < \dots < s_k < t$*, the limiting value* $\lim_{t\to\infty} \mathbb{P}_{i_0}(Z_{s_1} = i_1, \dots, Z_{s_k} = i_k | T > t)$ *exists.*
Let $(Y_t)_{t\geq 0}$ *be the process starting from* $i_0 \in S^*$ *and defined by its finite dimensional distributions*

$$\mathbb{P}_{i_0}(Y_{s_1} = i_1, \dots, Y_{s_k} = i_k) = \lim_{t\to\infty} \mathbb{P}_{i_0}(Z_{s_1} = i_1, \dots, Z_{s_k} = i_k | T > t). \tag{2.19}$$

Then Y *is a Markov process with values in* S^* *and transition probabilities given by*

$$\mathbb{P}_i(Y_t = j) = e^{-\theta(\alpha)t} \frac{\pi(i)}{\pi(j)} p_t(i,j),$$

for all $i, j \in S^*$ *and* $t \geq 0$*. Moreover,* Y *is conservative, and has the unique stationary distribution* $(\alpha(j)\pi(j))_{j\in S^*}$*.*

Proof. Let us denote $\theta(\alpha)$ by θ for simplicity. Let $i_0, \dots, i_k \in S^*$ and $0 < s_1 < \dots < s_k < t$. We recall that, in our notation, $\mathcal{F}_s$ is the natural filtration of the process, equal to $\sigma(Z_u, u \leq s)$, $\forall s \geq 0$. Then we have

$$\begin{aligned}
\mathbb{P}_{i_0}(Z_{s_1} = i_1, \dots, Z_{s_k} = i_k; T > t) &= \mathbb{E}_{i_0}(\mathbb{1}_{Z_{s_1}=i_1,\dots,Z_{s_k}=i_k} \mathbb{E}_{i_0}(\mathbb{1}_{T>t} | \mathcal{F}_{s_k})) \\
&\overset{(1)}{=} \mathbb{E}_{i_0}(\mathbb{1}_{Z_{s_1}=i_1,\dots,Z_{s_k}=i_k} \mathbb{E}_{i_k}(\mathbb{1}_{T>t-s_k})) \\
&= \mathbb{P}_{i_0}(Z_{s_1} = i_1, \dots, Z_{s_k} = i_k) \mathbb{P}_{i_k}(T > t - s_k),
\end{aligned}$$

where in (1) we used the Markov property. By Theorem 2.8,

$$\lim_{t\to\infty} \frac{\mathbb{P}_{i_k}(T > t - s_k)}{\mathbb{P}_{i_0}(T > t)} = \frac{\pi(i_k)}{\pi(i_0)} e^{\theta(\alpha)s_k}.$$

Thus

$$\lim_{t\to\infty} \mathbb{P}_{i_0}(Z_{s_1} = i_1, \dots, Z_{s_k} = i_k | T > t) = \mathbb{P}_{i_0}(Z_{s_1} = i_1, \dots, Z_{s_k} = i_k) \frac{\pi(i_k)}{\pi(i_0)} e^{\theta(\alpha)s_k}, \tag{2.20}$$

which gives us the existence of the limiting value. Let us now show that Y is a Markov process. We have

$$\begin{aligned}
\mathbb{P}_{i_0}(Y_{s_1} = i_1, \dots, Y_{s_k} = i_k, Y_t = j) &\overset{(2.20)}{=} e^{\theta t} \frac{\pi(j)}{\pi(i_0)} \mathbb{P}_{i_0}(Z_{s_1} = i_1, \dots, Z_{s_k} = i_k, Z_t = j) \\
\text{(By Markov property of } Z) &= e^{\theta(t-s_k)} e^{\theta s_k} \frac{\pi(j)}{\pi(i_k)} \frac{\pi(i_k)}{\pi(i_0)} \mathbb{P}_{i_0}(Z_{s_1} = i_1, \dots, Z_{s_k} = i_k) \mathbb{P}_{i_k}(Z_t = j) \\
&= \mathbb{P}_{i_0}(Y_{s_1} = i_1, \dots, Y_{s_k} = i_k) \mathbb{P}_{i_k}(Y_t = j),
\end{aligned}$$

and thus $\mathbb{P}_{i_0}(Y_t = j | Y_{s_1} = i_1, \ldots, Y_{s_k} = i_k) = \mathbb{P}_{i_k}(Y_{t-s_k} = j)$.
By (2.20) and Theorem 2.8, we have

$$\mathbb{P}_i(Y_t = j) = \frac{\pi(j)}{\pi(i)} \mathbb{P}_i(Z_t = j) e^{\theta t} \xrightarrow[t\to\infty]{} \frac{\pi(j)}{\pi(i)} \alpha(j)\pi(i) = \alpha(j)\pi(j),$$

therefore, by Theorem 1.6, one gets that $(\alpha(j)\pi(j))_{j\in S^*}$ is the unique stationary distribution of Y_t. Moreover let us compute the infinitesimal generator $\widetilde{A} = \{\widetilde{q}(i,j)\}$ of Y from the infinitesimal generator A of Z. Denoting also with $\widetilde{P}(t) = \{\widetilde{p}_t(i,j)\}$ the semi-group of Y, since

$$\widetilde{p}_s(i,j) = \mathbb{P}_i(Y_s = j) = \frac{\pi(j)}{\pi(i)} e^{\theta s} p_s(i,j) \xrightarrow[s\to 0]{} \frac{\pi(j)}{\pi(i)} q(i,j),$$

we have for $j \neq i$,

$$\widetilde{q}(i,j) = \lim_{s\to 0} \widetilde{p}_s(i,j) = \frac{\pi(j)}{\pi(i)} q(i,j).$$

For $j = i$,

$$\begin{aligned}\widetilde{q}(i,i) &= -\lim_{s\to 0} \frac{1-\widetilde{p}_s(i,i)}{s} = -\lim_{s\to 0} \frac{1-e^{\theta s} p_s(i,i)}{s} \\ &= -\lim_{s\to 0} \frac{1-e^{\theta s} + e^{\theta s}(1-p_s(i,i))}{s} \stackrel{(2)}{=} \theta + q(i,i),\end{aligned}$$

where in (2) we used the first order Taylor develop of the exponential. We thus check that

$$\sum_{j\in S^*} \widetilde{q}(i,j) = \sum_{j\in S^*} \frac{\pi(j)}{\pi(i)} q(i,j) + \theta.$$

Finally, from (2.15) we know that $A\pi = -\theta\pi$, then $\sum_{j\in S^*} \pi(j) q(i,j) = -\theta\pi(i)$ and thus $\sum_{j\in S^*} \widetilde{q}(i,j) = 0$, showing that Y is conservative. $\square$

2.3.3 An example: QSDs for birth and death processes

We are describing here the dynamics of isolated asexual populations, as for example populations of bacteria with cell binary division, in continuous time. Individuals may reproduce or die, and there is only one child per birth. The population size dynamics will be modeled by a birth and death process in continuous time.
We consider a birth and death processes with rates (λ_i) and (μ_i), that is a $\mathbb{N}$-valued regular Markov processes, whose jumps are +1 or -1, with transitions

$$\begin{aligned} i &\to i+1 \quad \text{with rate} \quad \lambda_i, \\ i &\to i-1 \quad \text{with rate} \quad \mu_i, \end{aligned}$$

where λ_i and μ_i, $i \in \mathbb{N}$, are non-negative real numbers.
Knowing that the process is at state i at a certain time, it will wait for an exponential time of parameter λ_i before jumping to $i+1$ or, independently, will wait for an exponential time of parameter μ_i before jumping to $i-1$. The total jump rate from state i is therefore $\lambda_i + \mu_i$. We will assume in what follows that $\lambda_0 = \mu_0 = 0$. This condition ensures that 0 is an absorbing point modeling the extinction of the population.
We consider a birth and death process $Z = (Z_t)_{t\geq 0}$ with almost sure extinction. We will show first a result concerning a necessary and sufficient condition for a sequence $(\alpha(i))$ in $\mathbb{N}^*$ to be a QSD for Z, then we will see cases under which QSDs are not unique.

Theorem 2.10. *The sequence $(\alpha(j))_{j\in\mathbb{N}^*}$ is a QSD for Z if and only if*

1. $\alpha(j) \geq 0$, $\forall j \geq 1$ *and* $\sum_{j\geq 1} \alpha(j) = 1$.
2. $\forall j \geq 2$,

$$\begin{aligned} \lambda_{j-1}\alpha(j-1) - (\lambda_j + \mu_j)\alpha(j) + \mu_{j+1}\alpha(j+1) &= -\mu_1\alpha(1)\alpha(j); \\ -(\lambda_1 + \mu_1)\alpha(1) + \mu_2\alpha(2) &= -\mu_1\alpha(1)^2. \end{aligned}$$

Proof. By Proposition 2.5 and for a QSD α, there exists $\theta > 0$ such that

$$\alpha A = -\theta\alpha,$$

where A is the infinitesimal generator of Z restricted to $\mathbb{N}^*$. Taking the j^{th} component of this equation, we get

$$\begin{aligned} \lambda_{j-1}\alpha(j-1) - (\lambda_j + \mu_j)\alpha(j) + \mu_{j+1}\alpha(j+1) &= -\theta\alpha(j); \\ -(\lambda_1 + \mu_1)\alpha(1) + \mu_2\alpha(2) &= -\theta\alpha(1). \end{aligned}$$

Summing over $j \geq 1$, we get after re-indexing

$$0 = \sum_{j\geq 1}(\lambda_j\alpha(j) - (\lambda_j + \mu_j)\alpha(j) + \mu_j\alpha(j)) = -\theta\overbrace{\sum_{j\geq 1}\alpha(j)}^{=1} + \mu_1\alpha(1).$$

Thus we deduce that $\theta = \mu_1\alpha(1)$, which concludes the proof. □

The next result follows immediately.

Corollary 2.11. *Let us define inductively the sequence of polynomials $(H_n(x))_{n\in\mathbb{N}}$ as follows: $H_1(x) = 1$ for all $x \in \mathbb{R}$ and for $n \geq 2$,*

$$\begin{aligned} \lambda_n H_{n+1}(x) &= (\lambda_n + \mu_n - x)H_n(x) - \mu_{n-1}H_{n-1}(x); \\ \lambda_1 H_2(x) &= \lambda_1 + \mu_1 - x. \end{aligned}$$

Then, any QSD $(\alpha(j))_j$ *satisfies for all* $j \geq 1$,

$$\alpha(j) = \alpha(1)\eta_j H_j(\mu_1 \alpha(1)), \tag{2.21}$$

where

$$\eta_1 = 1; \quad \eta_n = \frac{\lambda_1 \dots \lambda_n}{\mu_1 \dots \mu_n}.$$

For the study of polynomials and their implication in birth and death processes we refer to Van Doorn [20]. In particular, it is shown that there exists a non-negative number ξ_1 such that

$$x \leq \xi_1 \iff H_n(x) > 0, \quad \forall n \geq 1$$

By Corollary 2.11 then $\alpha(j) = \alpha(1)\eta_j H_j(\mu_1\alpha(1))$. Since for any j, $\alpha(j) > 0$, we have $H_j(\mu_1\alpha(1)) > 0$ for all $j \geq 1$ and then $0 < \mu_1\alpha(1) \leq \xi_1$. We can immediately deduce from this property that if $\xi_1 = 0$, then there is no quasi- stationary distribution. To go further, define the series (S) with general term

$$S_n = \frac{1}{\lambda_n \eta_n} \sum_{i=n+1}^{\infty} \eta_i.$$

It can be shown that (see [20], Theorems 3.2 and 4.1)

$$\lim_{t\to\infty} \mathbb{P}_i(Z_t = j | T > t) = \frac{1}{\mu_1}\eta_j \xi_1 H_j(\xi_1), \quad \text{thus} \quad \xi_1 = \lim_{t\to\infty} \mu_1 \mathbb{P}_i(Z_t = 1 | T > t), \quad \forall i \in S^*,$$

also:

1. If $\xi_1 = 0$, there is *no QSD*.
2. If (S) converges, then $\xi_1 > 0$ and the Yaglom limit is *the unique QSD*.
3. If (S) diverges and $\xi_1 \neq 0$, then there is *a continuum of QSD*, given by the one parameter family $(\widetilde{\alpha}_j(x))_{0<x\leq\xi_1}$:

$$\widetilde{\alpha}_j(x) = \frac{1}{\mu_1}\eta_j x H_j(x).$$

With the latter result we conclude the chapter related to quasi-stationary distributions. The next step will concern an introduction to interacting particle systems, in particular to a model that we will develop in its entirety and will be the core of this paper: the voter model.

Chapter 3

Interacting particle systems

Throughout the following chapter we will make an overview of what interacting particle system are, summarizing the general ideas of this field of research, in order to introduce the model on which this thesis is focused on. We will mainly refer to the work of Liggett [16] for the general theory, while we cite the papers of Cox [5], Aldous and Fill [1] and Oliveira [18] for their approach to the duality of the problem and the study of the limit behaviour of the system.

The field of interacting particle systems began as a branch of probability theory in the late 1960's. The original motivation for this field came from statistical mechanics, but, as time passed, it became clear that models with a very mathematical structure could be naturally formulated in other contexts: neural networks, spread of infection, behavioral systems, etc. From a mathematical point of view, interacting particle systems represents a natural departure from the established theory of Markov processes. A typical interacting particle system consists of finitely or infinitely many particles which, in absence of interaction, would evolve according to independent finite or countable state Markov chains. If we add some type of interaction, the evolution of an individual particle is no longer Markovian, while the system as a whole still is, though a very complex one.

The main problems which have been treated involve the long-time behaviour of the system. The first step to derive limit theorems is to describe the class of invariant measures for the process, since these are the possible asymptotic distributions. The next step is to characterize the domain of attraction of each invariant measure, i.e. the class of all initial distributions for which the distribution at time t of the process converges to the invariant measures as $t \to \infty$.

To be more precise, we will consider a compact metric space X with measurable structure given by the σ-algebra of Borel sets. In our case, we will take $X = W^V$, where V will denote a finite or countable set of spacial locations (called *sites*) and W will be a compact metric space which will play the role of the phase space of the random variable located at each site; further on, we will consider as V a more specific set such as (the vertex set of) graphs or lattices. Moreover, in the first part of this chapter we will use all the notions described in Section 2.1 regarding Markov processes. In the view of interacting particle systems, every -at most countable- $\eta_t(\cdot) \in X = W^V$, $t \geq 0$, corresponds to the evolution in time of stochastic processes (called *coordinates*) on the state space W, while η_t corresponds to the evolution of the entire system, seen as a Markov process on the uncountable totally disconnected space W^V. This relation motivate the need to

construct a new theory instead of using the more standard one related to Markov processes.
In order to be more concrete, let us introduce informally some examples of interacting particle systems on special state spaces. We remark the fact that we will deal only with systems in which only one coordinate of η_t changes at a time. In general, however, infinitely many coordinates may change in any interval of time.

- *The Stochastic Ising Model.* This is a model for magnetism which was introduced by Glauber (1963). It is a Markov process with $W = \{-1, +1\}$ and $V = \mathbb{Z}^d$, thus $X = \{-1, +1\}^{\mathbb{Z}^d}$. The sites represent iron atoms, which are laid out on the d-dimensional integer lattice $\mathbb{Z}^d$, while the value of ± 1 at a site represent the spin of the atom at that site. A configuration of spins η is then a point in $\{-1, +1\}^{\mathbb{Z}^d}$. The dynamics of the evolution are specified by the requirement that a spin $\eta(x)$ at $x \in \mathbb{Z}^d$ flips to $-\eta(x)$ at rate

$$\exp\left[-\beta \sum_{y:|y-x|=1} \eta(x)\eta(y)\right],$$

  where β is a nonnegative parameter which represents the reciprocal of the temperature of the system. Of course when $\beta = 0$, the coordinates $\eta_t(x)$ are independent two-state Markov chains (there is no more interaction), so the system has a unique invariant measure, which is the the product of the stationary distributions for the individual two-state chains of every site $x \in \mathbb{Z}^d$, that is the Bernoulli product measure ν on $\{-1, +1\}^{\mathbb{Z}^d}$ with parameter $\frac{1}{2}$. Furthermore, for any initial distribution, the distribution at time t converges weakly to ν as $t \to \infty$ by the convergence theorem for finite-state irreducible Markov chains. Such a system, which has a unique invariant measure to which convergence occurs for any initial distribution, will be called *ergodic*. Nonergodicity corresponds to the occurrence of phase transitions, with distinct invariant measures corresponding to distinct phases. The main problem to be resolved for the stochastic Ising model is to determine for which values of β and d the process is ergodic. A complete answer is given in [16], Chapter IV.

- *The Contact Process.* This process was introduced and first studied by Harris (1974). It is a Markov process with $W = \{0, 1\}$ and $V = \mathbb{Z}^d$, thus $X = \{0, 1\}^{\mathbb{Z}^d}$. The dynamics are given by the following transition rates: at site $x \in \mathbb{Z}^d$:

$$1 \to 0 \quad \text{at rate } 1,$$

  and

$$0 \to 1 \quad \text{at rate } \lambda \sum_{y:|y-x|=1} \eta(y),$$

where λ is a positive parameter which is interpreted as the infection rate. With such interpretation, sites at which $\eta(x) = 1$ are regarded as infected, while sites at which $\eta(x) = 0$ are regarded as healthy. Infected individuals became healthy after an exponential time with parameter one, independently of the configuration, while healthy individuals become infected at a rate that is proportional to the number of infected neighbors. The contact process has a trivial invariant measure: the pointmass at $\eta \equiv 0$, i.e. $\mu \in \mathcal{P}(X)$ such that $\mu(\eta(x) = 0) = 1$ for all $x \in X$. As might be expected, depending on the choice of λ and d, there could be others or the process could be ergodic. Again, a complete treatment of the argument is given in [16], Chapter VI.

- *The Voter Model.* The voter model was introduced and mainly studied by Holley and Liggett (1975). Here again $W = \{0, 1\}$ and $V = \mathbb{Z}^d$, thus the state space is $X = \{0, 1\}^{\mathbb{Z}^d}$, and the evolution mechanism is described by saying that $\eta(x)$ changes to $1 - \eta(x)$ at rate

$$\frac{1}{2d} \sum_{y:|y-x|=1} \mathbb{1}_{\eta(x) \neq \eta(y)}.$$

  In the voter interpretation of Holley and Liggett, sites in $\mathbb{Z}^d$ represents voters who can hold either of two political positions, which are denoted by zero and one. A voter waits an exponential time of parameter one, and then adopts the position of a neighbor chosen at random. The voter model has two trivial invariant measures: the pointmasses at $\eta \equiv 0$ and $\eta \equiv 1$ respectively. Thus the voter model is not ergodic. The first main question is whether or not are any other extremal invariant measures. As we will see in the following sections, there are no others if $d \leq 2$. On the other hand, if $d \geq 3$, there is a one-parameter family $\{\mu_\rho, 0 \leq \rho \leq 1\}$ of extremal invariant measures, where μ_ρ is translation invariant and ergodic, and $\mu_\rho(\eta : \eta(x) = 1) = \rho$ for all $x \in X$. This dichotomy is closely related to the fact that a simple random walk on $\mathbb{Z}^d$ is recurrent if $d \leq 2$ and transient if $d \geq 3$. In terms of the voter interpretation, one can describe the result by saying that a consensus is approached as $t \to \infty$ if $d \leq 2$, but that disagreements persists indefinitely if $d \geq 3$.

Most of the latter introduction was taken from the Introduction Chapter of Liggett [16]. The core of this thesis is exactly the study of the *Voter Model* in a certain site space V, given by a graph structure that we will explain later, in particular we will look for quasi-stationary distributions for the process conditioned not to be absorbed.
This chapter is organized as follows: in Section 3.1 we will completely describe the model, and we will recall some basics definitions and results which we will exploit in this context. Then, in Section 3.2 we will analyze, both in general and specific site spaces, such as $V = \mathbb{Z}^d$ or $V = \Lambda \subset \mathbb{Z}^d$ finite, the set of invariant measures and their

domain of attraction. Furthermore, we will see in which cases there will be almost surely absorption (consensus, in the voter interpretation), and, in those situations, we will cite some bounds for the time to consensus. Eventually, in Section 3.3 we will motivate the coalescing duality assumption, developing the voter model in a regular graph.

3.1 The Voter Model

Let us first introduce a more general family of models: the spin systems. They are interacting particle systems in which each coordinate has two possible values, and only one coordinate changes in each transition. Throughout all this chapter, the state space of the system will be taken to be $X = \{0,1\}^V$, where V is a finite or countable set. The transition mechanism is specified by a nonnegative function $c(x,\eta)$ defined for $x \in V$ and $\eta \in X = \{0,1\}^V$. It represent the rate at which the coordinate $\eta(x)$ flips from 0 to 1, or vice versa, when the system is in state η. Therefore the process η_t with state space X will satisfy

$$\mathbb{P}_\eta(\eta_t(x) \neq \eta(x)) = c(x,\eta)t + o(t),$$

as $t \downarrow 0$, for each $x \in V$ and $\eta \in X$. Recall that $\mathbb{P}_\eta(\eta_t = \cdot) = \mathbb{P}(\eta_t = \cdot|\eta_0 = \eta)$, for all $t \geq 0$ and $\eta \in X$. The requirement that only one coordinate change in each transition can be described by saying that

$$\mathbb{P}_\eta(\eta_t(x) \neq \eta(x), \eta_t(y) \neq \eta(y)) = o(t),$$

as $t \downarrow 0$ for each $x, y \in V$ and $\eta \in X$. The interaction among sites comes from the dependence of $c(x,\eta)$ on η: the transition rates are influenced not only by the value $\eta(x)$ but also by the entire configuration.

The voter model is the spin system with rates $c(x,\eta)$ given by

$$c(x,\eta) = \begin{cases} \sum_y p(x,y)\eta(y) & \text{if } \eta(x) = 0, \\ \sum_y p(x,y)[1-\eta(y)] & \text{if } \eta(x) = 1, \end{cases} \tag{3.1}$$

where $p(x,y) \geq 0$ for all $x, y \in V$, and $\sum_y p(x,y) = 1$ for all $x \in V$. Moreover, we will assume that $p(x,y)$ is such that the Markov chain with this transition probabilities is irreducible. Here $p(x,y)$ represent the transition probability of the embedded chain which describes the probability to pass from a site x to a site y. In fact, an equivalent way of describing the rates is to say that a site x waits an exponential time of parameter one, at which it flips to the value it sees at that time at a site y, chosen with probability $p(x,y)$. The voter interpretation which gives this process its name views V as a collection of individuals, each of them taking two possible positions, denoted by 0 and 1, on a political issue. After independent exponential times, an individual reassesses his opinion by looking to the opinion of an acquaintance chosen at random.

One of the main tools used in literature to study the voter model, which we will strongly adopt from now on, is its duality with coalescing random walks. Such duality property allows us to recast problems involving the voter model in terms of the dual system of coalescing Markov chains.
Before we start to be more formal, we need to give a better understanding of the Markov processes tools we will use to examine the model.

3.1.1 General Results

Here we are going to transpose some notions introduced for quasi-stationary distributions in our new interacting particle context. Let X, V and W be as in the introduction of this chapter. First of all, let us denote as a Markov process the collection $\{\mathbb{P}_\eta, \eta \in X\}$ of probability measures on $D[0,\infty)$ that satisfies the conditions given in Section 2.1, where $D[0,\infty)$ is the set of all functions $\eta.$ on $[0,\infty)$ with values in X which are right continuous and have left limits. The corresponding expectation is $\mathbb{E}_\eta$, moreover, if we consider the Banach space $C(X)$ of the continuous functions on X endowed with the sup-norm, for $f \in C(X)$, we define

$$P(t)f(\eta) = \mathbb{E}_\eta[f(\eta_t)].$$

Definition 3.1. A Markov process $\{\mathbb{P}_\eta, \eta \in X\}$ is said to be a *Feller process* if $P(t)f \in C(X)$ for every $t \geq 0$ and $f \in C(X)$.

Theorem 3.1. *Let $\{\mathbb{P}_\eta, \eta \in X\}$ be a Feller process on X. Then the family of linear operators $\{P(t), t \geq 0\}$ on $C(X)$ has the following properties:*

1. *$P(0) = I$, the identity operator on $C(X)$.*
2. *The mapping $t \to P(t)$ from $[0,\infty)$ to $C(X)$ is right continuous for every $f \in C(X)$.*
3. *$P(t+s) = P(t)P(s)$ for all $f \in C(X)$ and all $s, t \geq 0$.*
4. *$P(t)1 = 1$ for all $t \geq 0$.*
5. *$P(t)f \geq 0$ for all nonnegative $f \in C(X)$.*

Such a family is called a Markov semi-group *for the Markov process $\{\mathbb{P}_\eta, \eta \in X\}$.*
Moreover, given a Markov semi-group $\{P(t), t \geq 0\}$ on $C(X)$, there exists a unique Markov process $\{\mathbb{P}_\eta, \eta \in X\}$ such that $P(t)f(\eta) = \mathbb{E}_\eta[f(\eta_t)]$.

Similarly to what we have seen in (2.3), if we take $\mu \in \mathcal{P}(X)$, the set of all probability measures on X, and a Markov process $\{\mathbb{P}_\eta, \eta \in X\}$, then the corresponding Markov process with initial distribution μ is a stochastic process η_t whose distribution is given by

$$\mathbb{P}_\mu = \int_X \mathbb{P}_\eta \mu(d\eta).$$

In view of this,

$$\mathbb{E}_\mu[f(\eta_t)] = \int_X P(t)f\, d\mu, \quad \forall f \in C(X).$$

This suggest the following definition.

Definition 3.2. Suppose $\{P(t), t \geq 0\}$ is a Markov semi-group on $C(X)$. Given $\mu \in \mathcal{P}(X)$, $\mu P(t) \in \mathcal{P}(X)$ is defined by the relation

$$\int_X f\, d[\mu P(t)] = \int_X P(t)f\, d\mu,$$

for all $f \in C(X)$.

The probability measure $\mu P(t)$ is interpreted as the distribution at time t of the process when the initial distribution is μ. We now want to study the limit behaviour of $\mu P(t)$ as $t \to \infty$. To do that we need the following

Definition 3.3. A measure $\mu \in \mathcal{P}(X)$ is said to be *invariant* for the process with Markov semi-group $\{P(t), t \geq 0\}$ if $\mu P(t) = \mu$ for all $t \geq 0$. The class of invariant $\mu \in \mathcal{P}(X)$ will be denoted with $\mathcal{I}$. Moreover, define $\mathcal{I}_e$ as the set of extreme invariant measures , i.e. the elements of $\mathcal{I}$ which can not be written as a nontrivial convex combinations of elements of $\mathcal{I}$.

Proposition 3.2. *The following properties holds:*

1. *$\mu \in \mathcal{I}$ if and only if*
$$\int_X P(t)f\, d\mu = \int_X f\, d\mu,$$
for all $f \in C(X)$ and all $t \geq 0$.

2. *$\mathcal{I}$ is a compact convex subset of $\mathcal{P}(X)$.*

3. *$\mathcal{I}$ is the closed convex hull of $\mathcal{I}_e$.*

4. *If $\nu = \lim_{t\to\infty} \mu P(t)$ exists for some $\mu \in \mathcal{P}(X)$, then $\nu \in \mathcal{I}$.*

5. *$\mathcal{I}$ is not empty.*

Let us conclude this introductory section with a definition that describes the nicest situation which one can have relative to convergence of $\mu P(t)$ as $t \to \infty$.

Definition 3.4. The Markov process with semi-group $\{P(t), t \geq 0\}$ is said to be *ergodic* if

i. $\mathcal{I} = \{\nu\}$ is a singleton, and

ii. $\lim_{t\to\infty} \mu P(t) = \nu$ for all $\mu \in \mathcal{P}(X)$.

Now we are ready to study the ergodic theory related to the voter model.

3.2 Invariant measures

In this section, we will find all the extremal invariant measures of the voter model and their domain of attraction. Our aim is just to provide the main ideas and general results, hence we will not give any proof. For a detailed discussion of the topics, we suggest the references [16], [5] and [18].
Recall that the flip rates of this interacting particle system are given by (3.1), and $p(x, y)$ are the transition probability of an irreducible Markov chain on the site space V. Let us mention the definition of the set of bounded harmonic functions for a Markov chain with transitions $p(x, y)$. Thus, let

$$\mathcal{H} = \Big\{\alpha : V \to [0, 1] \text{ such that } \sum_y p(x, y)\alpha(y) = \alpha(x),\ \forall x \in V\Big\}. \tag{3.2}$$

In particular, we will need a certain subset $\mathcal{H}^*$ of $\mathcal{H}$. In order to describe it, let $X(t)$ and $Y(t)$ be two independent copies of the continuous time Markov chain, having $\mathbf{P} = \{p(x, y)\}_{x,y\in V}$ as the embedded chain transition matrix, i.e. with transition probabilities (we are assuming the transition rate equal to one)

$$p_t(x, y) = e^{-t} \sum_{n=0}^{\infty} \frac{t^n}{n!} p^{(n)}(x, y),$$

where $p^{(n)}(x, y)$ are the n-step transition probabilities associated to $p(x, y)$. It can be shown that, for $\alpha \in \mathcal{H}$, $\alpha(X(t))$ is a bounded martingale. Therefore, by martingale convergence theorem, $\lim_{t\to\infty} \alpha(X(t))$ exists with probability one and the following subset of $\mathcal{H}$ is well defined

$$\mathcal{H}^* = \Big\{\alpha \in \mathcal{H} : \lim_{t\to\infty} \alpha(X(t)) = 0 \text{ or } 1 \text{ a.s. on the event } \mathcal{E}\Big\},$$

where $\mathcal{E} = \{\text{there exists } t_n \uparrow \infty \text{ such that } X(t_n) = Y(t_n)\}$. Using the Markov property on

$$g(x, y) := \mathbb{P}^{(x,y)}(X(t) = Y(t) \text{ for some } t \geq 0), \tag{3.3}$$

$x, y \in V$ and $(X(0), Y(0)) = (x, y)$, it can be seen that $g(x, y)$ is a nonnegative supermartingale, thus $\lim_{t\to\infty} g(X(t), Y(t))$ exists a.s., so $\mathcal{E}$ can be nonempty. For $\alpha \in \mathcal{H}$, define ν_α to be the product measure on X with marginals

$$\nu_\alpha\{\eta : \eta(x) = 1\} = \alpha(x).$$

We can now state

Theorem 3.3. *The following holds:*

(a) $\mu_\alpha = \lim_{t\to\infty} \nu_\alpha P(t)$ *exists for every* $\alpha \in \mathcal{H}$*, and* $\mu_\alpha \in \mathcal{I}$*.*

(b) $\mu_\alpha\{\eta : \eta(x) = 1\} = \alpha(x)$ *for all* $x \in V$*.*

(c) $\mathcal{I}_e = \{\mu_\alpha : \alpha \in \mathcal{H}^*\}$*.*

Therefore we have a complete description of the set $\mathcal{I}$ of the invariant measures for the voter model, and the set $\mathcal{I}_e$ of the extremal measures of $\mathcal{I}$. Before stating the next theorem concerning the domain of convergence of invariant measures, we define

$$\hat{\mu}(A) := \mu\{\eta : \eta(x) = 1 \text{ for all } x \in A\},$$

where $\mu \in \mathcal{P}(X)$ and $A \in V_F$, the class of finite subset of V.

Theorem 3.4. *The following holds*

(a) Suppose $\mu \in \mathcal{P}(X)$ *and* $\alpha \in \mathcal{H}^*$*. Then* $\lim_{t\to\infty} \mu P(t) = \mu_\alpha$ *if and only if*

$$\lim_{t\to\infty} \sum_{y\in V} p_t(x, y)\hat{\mu}(\{y\}) = \alpha(x), \quad \textit{and}$$
$$\lim_{t\to\infty} \sum_{u,v\in V} p_t(x, u)p_t(x, v)\hat{\mu}(\{u, v\}) = \alpha^2(x),$$

for all $x \in V$*.*

(b) Suppose $g(x, y) = 1$ *for all* $x, y \in V$*, so that* $\mathcal{H}^* = \{0, 1\}$*. If* $\theta \in [0, 1]$ *and* $\mu \in \mathcal{P}(X)$*, then* $\lim_{t\to\infty} \mu P(t) = \theta\delta_1 + (1 - \theta)\delta_0$ *if and only if*

$$\lim_{t\to\infty} \sum_{y\in V} p_t(x, y)\hat{\mu}(\{y\}) = \theta, \quad \textit{for all } x \in V,$$

where δ_x *is the Dirac measure with mass in* $x \in V$*.*

Weaker conditions still hold if we take $\alpha \in \mathcal{H}$ instead of $\mathcal{H}^*$, see [16] Chapter V.
At this point we are able to characterize the set $\mathcal{I}$ of all invariant measures for the voter model and their domain of attraction. Since we are interested in the case in which the elements of V represent the vertices of a finite graph, first we will briefly discuss the difference between the cases V finite and V infinite, then we will approach the finite graph case.

3.2.1 Absorbing states

As we are studying the voter model with two opinions, i.e. $W = \{0, 1\}$, it is obvious that the states $\{\eta : \eta(x) = 1 \; \forall x \in V\}$ and $\{\eta : \eta(x) = 0 \; \forall x \in V\}$ are absorbing for the process, since once the system is in one of those it gets trapped and remains in that state forever. We want to connect this concept with that of the invariant measure. To do this, we first observe that if V is finite (and the $p(x, y)$ chain irreducible) then all stationary distributions are trivial, i.e., concentrated on absorbing states, since no matter the initial state, η_t will get trapped at all 0's or 1's with probability one. In such a situation, the first question that might be asked is how can one estimate the hitting time τ_c to reach consensus (or consensus time), where

$$\tau_c := \inf\{t > 0 : \eta_t(x) = \eta_t(y) \quad \forall x, y \in V\}. \tag{3.4}$$

In particular we cite the work of Cox [5], where $V = \mathbb{Z}^d \cap [-N/2, N/2)^d$, $N = 2, 4, \dots,$ is the d-dimensional torus in $\mathbb{Z}^d$ and $p(x, y)$ are given by a symmetric random walk in V. In such setting, several estimates of $\mathbb{E}[\tau_c]$ are reported, mainly using the coalescing duality of the voter model which we will explain in the next section. Another relevant observation is that we can study the behaviour of the process conditioned on non absorption: it will be our focus for the next chapters.
If we now consider the case in which V is not finite, the preceding argument that links the invariant measures to the absorbing states fails. Indeed it is possible to have a nontrivial stationary distributions. For a matter of simplicity, let us take $V = \mathbb{Z}^d$, and let $p(x, y)$ be the transition functions of a symmetric random walk on V, $p(x, y) = \frac{1}{2d}\mathbb{1}_{|x-y|=1}$. With such assumptions $\mathcal{H}$ is characterized by all the constants functions $\theta \in [0, 1]$ and $\mathcal{H}^* = \{0, 1\}$. Similar to the previous section, let ν_θ be the product measure with density θ on $X = \{0, 1\}^{\mathbb{Z}^d}$, i.e.

$$\nu_\theta\{\eta : \eta(x) = 1\} = \theta \quad \forall x \in \mathbb{Z}^d.$$

We can then state

Theorem 3.5. *Let $\{\eta_t, t \geq 0\}$ be the Markov process representing the voter model on the state space $X = \{0, 1\}^{\mathbb{Z}^d}$. The following holds*

(i) If $d \leq 2$, then $\mathcal{I}_e = \{\nu_0, \nu_1\}$ and $\lim_{t\to\infty} \mu P(t) = \theta\nu_0 + (1-\theta)\nu_1$, where $\mu \in \mathcal{T}$[1] and $\theta = \mu\{\eta : \eta(x) = 1\}$ for all $x \in \mathbb{Z}^d$.

(ii) If $d \geq 3$, then $\mathcal{I}_e = \{\nu_\theta, \; \theta \in [0, 1]\}$, and $\lim_{t\to\infty} \mu P(t) = \nu_\theta$, where $\mu \in \mathcal{T}$ and $\theta = \mu\{\eta : \eta(x) = 1\}$ for all $x \in \mathbb{Z}^d$.

Remark 9. The latter theorem is a consequence of a result of Holley and Liggett [15], which suppose in *(i)* that $X(t) - Y(t)$ (an irreducible symmetric random walk on $\mathbb{Z}^d$, with

[1] The set of translation invariant probability measures of X

$X(t)$ and $Y(t)$ as in (3.3)) is recurrent, instead of $d \leq 2$, and in *(ii)* that $X(t) - Y(t)$ is transient, instead of $d \geq 3$. Clearly the cited hypothesis are equivalent, since an irreducible symmetric random walk in $\mathbb{Z}^d$ is recurrent if and only if $d \leq 2$, and transient if and only if $d \geq 3$.

Thus, if $d \leq 2$ and letting $t \to \infty$, we have *consensus* almost surely. More precisely, there is a probability θ that clustering occurs with all states 0's, and a probability $1 - \theta$ that clustering occurs with all states 1's, where $\theta = \mu\{\eta : \eta(x) = 1\}$ and $\mu \in \mathcal{P}(X)$ is the initial distribution of η_t. This implies that

$$\mathbb{P}_{\eta_0}(\eta_t(x) \neq \eta_t(y)) \underset{t\to\infty}{\longrightarrow} 0,$$

for every $x, y \in \mathbb{Z}^d$ and all initial state η_0.
On the contrary, if $d \geq 3$ and letting $t \to \infty$, *coexistence* of opinions occurs almost surely, i.e. there exists a translation invariant stationary distribution in which each of the possible states in $\{0, 1\}$ has positive density. In fact, for all possible invariant measures $\nu_\theta \in \mathcal{I}_e$, $\theta \in [0, 1]$, we have that $\theta = \nu_\theta\{\eta : \eta(x) = 1\} = \mu\{\eta : \eta(x) = 1\}$ for all $x \in \mathbb{Z}^d$ and every initial distribution $\mu \in \mathcal{T}$.

3.3 Duality with coalescing random walk

The aim of this section is to explain and motivate the duality relation between the voter model and the coalescing random walks. To do so, in view of the next chapters, we will consider the voter model on a r-regular[2]n-vertex graph $\mathbf{G}$, i.e. the corresponding Markov process $\{\eta_t, t \geq 0\}$ will be defined as follows

$$\eta_t : V(\mathbf{G}) \longrightarrow \{0, 1\}, \quad \forall t \geq 0,$$

where $V(\mathbf{G})$ is the vertex set of $\mathbf{G}$ and $\{0, 1\}$ are the usual possible opinions of each site at time t. Using the above notation, $V = V(\mathbf{G})$ is the site space, furthermore, the edge set of $\mathbf{G}$ is given by $E = \{(x, y) : x, y \in V(\mathbf{G}),\ p(x, y) > 0\}$. Let us now briefly recall the evolution of the model. In the continuous time voter model we envisage a person at each vertex, which initially has an opinion corresponding to 0 or 1. As time passes, opinions change according to the following rule. For each person x and each time interval $[t, t+dt]$, with chance dt the person chooses uniformly at random a neighbor, say y, and, if they are of the same opinion nothing happens, otherwise x changes his opinion to the current opinion of person y. Since the graph is finite, the consensus time τ_c described in (3.4) is well defined.
We now introduce the coalescing random walk process on the graph $\mathbf{G}$. At time $t = 0$

[2]A regular graph is a graph where each vertex has the same number of neighbors or, equivalently, the same degree. A regular graph with vertices of degree r is called a r-regular graph.

there is one particle at each vertex of the graph. These particles perform independent continuous-time random walks on the graph, but when particles meet they coalesce into clusters and the cluster thereafter sticks together and moves as a single random walk. Therefore if at time t there are clusters at distinct vertices, composed by one or more particles, then during $[t, t+dt]$ each cluster has chance dt to move to a random neighbor and (if that neighbor is occupied by another cluster) to coalesce with that other cluster. Note that, since the graph is finite, the total number of clusters can only decrease over time, and at some random time σ_c the particles will have all coalesced into a single cluster.
In order to connect these two processes, we are going to show an example of their evolution in the 8-cycle graph, where the i'th vertex is connected to the vertices $i+1$ and $i-1$, furthermore the vertices 0 and 8 are considered equal. Regarding the following construction, we adapted the one used into the work of Aldous and Fill [1], Chapter 14.

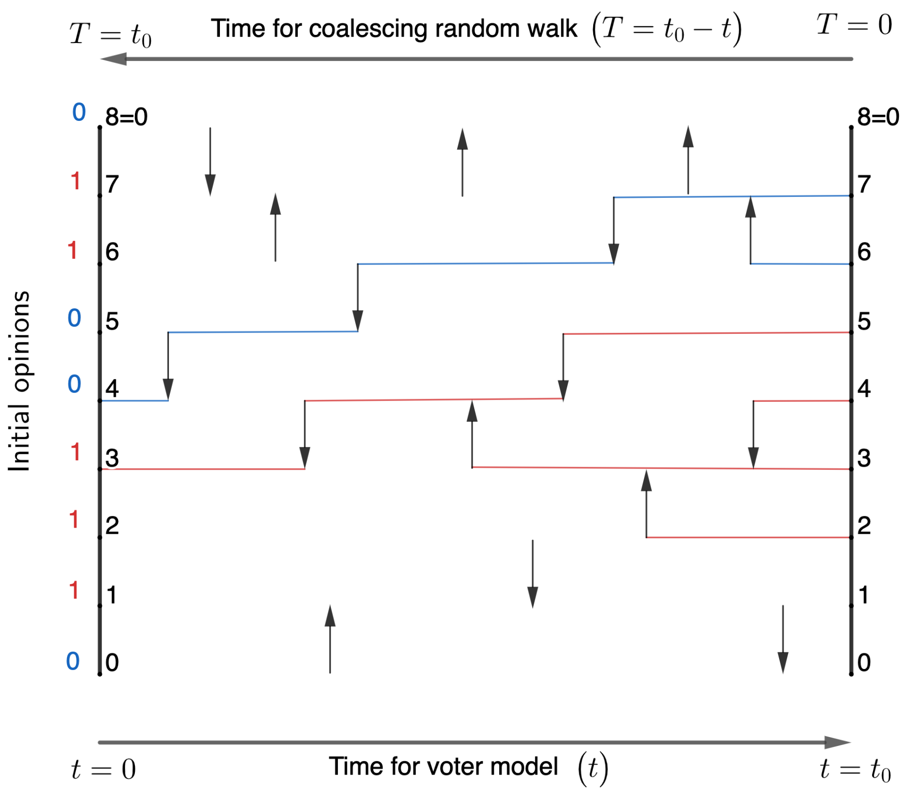

Figure 3.1: Comparing voter model and coalescing random walk

Here in Figure 3.1, for each edge e and each direction on e, was created a Poisson

process of rate $\frac{1}{r}$. Fixing $t_0 > 0$, time evolution is horizontal and an event of the Poisson process for edge (i, j) at time t is indicated by a vertical black arrow $i \to j$ at time t. Note that we introduced two time parameters: t for the voter model, which evolves foreward, and T for the coalescing RW, which evolves backward.
In the voter model, we interpret time (t) as increasing left-to-right from 0 to t_0, and we interpret an arrow $j \to i$ at time t as meaning that person j adopts i's opinion a time t. While in the coalescing random walk model, we interpret time (T) as increasing right-to-left from 0 to t_0, and we interpret an arrow $j \to i$ at time t as meaning that the cluster (if any) at state j at time t jumps to state i, and coalesces with the cluster at i (if any). The blue and red horizontal lines in the figure indicate part of the trajectories. Moreover, we have shown, by way of example, the evolution of initial opinions 0 and 1 relating to the 4th and 3rd vertex ($t = 0$), respectively, which will influence the opinions of vertices from 2 to 7 ($t = t_0$). If we observe such evolution back in time, putting a random walker on every vertex from 2 to 7 and considering the coalescing process, it turns out that starting from six walkers ($T = 0$) they coalesced into two ($T = t_0$). Generalizing what we have just observed, we can state that: for any vertices i, j, k the event (for the voter model)

"The opinions of persons i and j at time $t = t_0$ are both the opinion initially held by k"

is exactly the same as the event (for the coalescing random walk process)

"The particles starting at i and at j have coalesced before time $T = t_0$ and their cluster is at vertex k at time $T = t_0$.".

Therefore, the event (for the voter model)

"By time $t = t_0$ everyone's opinion is the opinion initially held by person k."

is exactly the same as the event (for the coalescing random walk process)

"All particles have coalesced by time $T = t_0$, and the cluster is at k at time $T = t_0$.".

As an immediate consequence, the hitting times τ_c and σ_c have the same distribution until time t_0, i.e.

$$\mathbb{P}(\tau_c \leq t) = \mathbb{P}(\sigma_c \leq t), \quad \forall t \leq t_0. \tag{3.5}$$

However, as we will see further on, it holds $\mathbb{P}(\tau_c \leq t) \leq \mathbb{P}(\sigma_c \leq t)$, for all $t \geq 0$. In particular, the inequality strongly depends on the choice of the initial distributions of opinions. Turning back to the setting of (3.5), let us denote with C the distribution of consensus and coalescing time. In [1] several bounds for $\mathbb{E}[C]$ are given, mainly using the structure of the finite graph we are working with, moreover, the problem of finding an universal bound was posed. In early 2012, Oliveira [18] proposed the following solution

Theorem 3.6. *There exists a universal constant $K > 0$ such that, for any graph $\mathbf{G}$ and any set W of opinions, the expected value of the consensus time of the voter model defined in terms of $\mathbf{G}$ and W, started from an arbitrary initial state, is bounded by $K\,T_{hit}^{\mathbf{G}}$.*

Where

$$T_{hit}^{\mathbf{G}} := \max_{v,w \in V(\mathbf{G})} \mathbb{E}_w[\tau_v] = \text{ largest expected hitting time for A,}$$

τ_v is the usual hitting time for $v \in V(\mathbf{G})$, and A is the infinitesimal generator of the Markov chain on $V(\mathbf{G})$, which we consider to be irreducible and reversible[3].
In order to conclude this chapter, let us summarize some of the key concepts in what follows. The crucial differences between a classical Markov chain and an interacting particle system are: the state space, that is no longer countable since is given by $X = \{0,1\}^V$, and the mechanism of the flipping rates $c(x,\eta)$, which is strictly related to both interaction between agents $x \in V$ and the transition probabilities $p(x,y)$, $x, y \in V$. Let us now focus the attention on the voter model. The approach to invariant measures revolves around the concept of absorbing states. If we are in a site state V where there is no sure absorption (as in $\mathbb{Z}^d$, with $d \geq 3$), then the properties of the invariant measures depends on the characteristics of the lattice; otherwise, if we consider the model in a context where absorption happens almost surely, then the invariant measures are only the trivial ones related to the absorbing state. Thus, an interesting treatment could be the study of the process conditioned not to be absorbed, in particular we can look for quasi-stationary distributions.
In the next chapter we will study the voter model in a complete bipartite graph $K_{n,m}$, $n, m \in \mathbb{N}$, following the very recent paper of Ben-Ari, Panzo, Speegle and VandenBerg [2]. We will look for QSDs for the voter model conditioned on non absorption, and we will see how they can be used to determine a particular distribution which will characterize the size of the "dissenters" in our voter context.

[3]Since we are assuming irreducibility, the chain has a unique stationary distribution π. We say that the infinitesimal generator $A = \{q(i,j)\}_{i,j \in V(\mathbf{G})}$ (and thus the associated Markov chain) is reversible w.r.t. π if $\pi(x)q(x,y) = \pi(y)q(y,x)$, for all distinct $x, y \in V(\mathbf{G})$.

Chapter 4

Voter model on Complete Bipartite Graphs

Through this chapter we will connect all the notions and results encountered so far. In particular, we will study the discrete-time voter model on the *complete bipartite graphs* $K_{n,m}$. These are heterogeneous graphs whose vertex set can be partitioned in two disjoint groups, a "large" group L of size n and a "small" one S of size m, where each vertex of L is connected to all of the vertices of S and vice versa, and there are no connections between vertices in the same group. The aim of this chapter is to investigate what happens when consensus is conditioned to never occur, as we have seen that in finite graphs it happens almost surely. More specifically, we study the quasi-stationary distribution (QSD) for the voter model on $K_{n,m}$ and its limit behaviour under the QSD when m is fixed and $n \to \infty$. The reasons why we are interested in such a limit are: to find out if the lack of consensus is due to a minority number of dissenters, or if the opinions are relatively balanced, and, taking $n \gg m$, to find out how the distributions of opinions in L differs from S. We will follow closely [2] during all this chapter, it will be our main reference.
In Section 4.1 we summarize the properties of both the QSD and voter model -still on a general finite graph- that we will use in order to obtain the results, emphasizing the connection between the eigenvalue of the QSD (we say "the" QSD because we will show that under our conditions it exists and is unique) and the consensus time. Then, in Section 4.2 we continue the duality discussion introduced in Section 3.3. More precisely, we point out in which cases the knowledge of the distribution of the coalescing time is a necessary and sufficient condition for determine the eigenvalue of the QSD for the voter model conditioned on nonabsorbing. Finally, in Section 4.3 we completely develop the core result of this chapter, which regards the distribution of opinions under the QSD for the voter model on $K_{n,m}$, letting $n \to \infty$ and m fixed.

4.1 Introduction

For a matter of completeness, in this section we will give a quick summary of the theory concerning the voter model on a general finite graph, and the theory of QSDs. Let us begin with the latter. For more details and all the references, see Chapter 2.
During all this chapter, differently from what we saw previously, we will consider the

voter model with *discrete* time setting and on a *discrete* state space. Thus, we will study the QSDs on that context. Suppose that $\mathbf{Y} = \{Y_t : t \in \mathbb{Z}_+\}$ is a Markov chain on a finite state space $\bar{S}$ and transition function $\bar{P}$. Note that here $\bar{P}$ is a $|\bar{S}| \times |\bar{S}|$ matrix. Recall that the state $i \in \bar{S}$ is absorbing if and only if

$$\mathbb{P}_i(Y_1 = i) = 1,$$

where as usual $\mathbb{P}_i(Y_t = \cdot) = \mathbb{P}(Y_t = \cdot | Y_0 = i)$, for all $t \geq 0$. Denote with $\Delta \subset \bar{S}$ the set of the absorbing states. We will assume that

(1) Δ is nontrivial, i.e. it is not empty and its complement Δ^c is not empty; and

(2) Δ is accessible from every state. That is, for every state $i \in \bar{S}$, there exists $t \in \mathbb{Z}_+$ such that
$$\mathbb{P}_i(Y_t \in \Delta) > 0.$$

An immediate consequence of these assumptions is that the hitting time τ of the absorbing set Δ is finite almost surely. That is, for any initial distribution $\mu \in \mathcal{P}(\bar{S})$ we have $\mathbb{P}_\mu(\tau < \infty) = 1$, where

$$\tau = \inf\{t \in \mathbb{Z}_+ : Y_t \in \Delta\}.$$

Let P be the sub-stochastic transition function obtained by restricting $\bar{P}$ to the complement of Δ in $\bar{S}$. We denote this complement by S. In other words, as we anticipated in the previous chapters, P is the principal sub-matrix obtained from $\bar{P}$ by removing all rows and columns corresponding to states in Δ. Therefore, taking any initial distribution μ for $\mathbf{Y}$ whose support is contained in S, we have

$$\mathbb{P}_\mu(Y_t = j, \tau > t) = \mu P^t(j), \quad \forall t \in \mathbb{Z}_+, \ \forall j \in S, \tag{4.1}$$

where, with P^t we denote the t-power of the sub-stochastic matrix P, and with $\mu P^t(j)$ we indicate the j'th component of the resulting product between the vector μ and the matrix P^t. In particular,

$$\mathbb{P}_\mu(\tau > t) = \sum_{j \in S} \mu P^t(j), \quad \forall t \in \mathbb{Z}_+. \tag{4.2}$$

Furthermore,
$$\lim_{t\to\infty} \mathbb{P}_\mu(\tau > t)^{\frac{1}{t}} \tag{4.3}$$

exists. Moreover, if one considers the restriction of S to the states accessible from the support of μ, then the limit above coincides with spectral radius of the resulting principal sub-matrix. In fact, it is a direct consequence of the Gelfand's Formula, using $||\cdot||_1$ as matrix norm while considering the rewrites (4.1) and (4.2).

Remark 10. There is a specific case in which the above statements are immediate. If in (4.3) we take as initial distribution μ a quasi-stationary distribution (assume it exists), from (2.9) and, more precisely, Remark 6 we know that

$$\mathbb{P}_\mu(\tau > t)^{\frac{1}{t}} = \left[\lambda^t\right]^{\frac{1}{t}} = \lambda,$$

where $\lambda \in (0,1)$ is the spectral radius of P restricted to the support of μ, and such that $\mu P = \lambda\mu$. Thus it follows that $\lim_{t\to\infty} \mathbb{P}_\mu(\tau > t)^{1/t} = \lambda$.

Recall now that a probability measure ν on $\bar{S}$ is called a quasi-stationary distribution for the Markov chain $\mathbf{Y}$ if

$$\mathbb{P}_\nu(Y_t = j|\tau > t) = \nu(j), \quad \forall j \in S,\ \forall t \in \mathbb{Z}_+.$$

Note that ν is supported on S, therefore will be viewed as a probability measure on S. Moreover, we can rewrite the latter equation using (4.1) and (4.2) to have

$$\nu P^t = C_\nu(t)\nu,$$

where $C_\nu(t) = \sum_{j\in S} \nu P^t(j)$. If we plug $t = 1$ in the above equation, we get the following well-known result, which we have already seen with the Perron-Frobenius Theorem (2.7) and Proposition 2.2.

Proposition 4.1. *The following holds*

(i) A probability vector ν on S is a QSD if and only if ν is a left eigenvector for P, with strictly positive eigenvalue λ. That is, if

$$\nu P = \lambda\nu,$$

with λ being the spectral radius of P restricted to the linear space spanned by the indicators of the support of ν.

(ii) If ν is a QSD, then the distribution of τ under P_ν is geometric with parameter $1 - \lambda$.

Furthermore we recall also the following theorem, that gives us sufficient conditions for the existence and uniqueness of QSDs, and provides the domain of attraction for these QSDs.

Theorem 4.2. *The following holds*

(i) If P is irreducible then there exists a unique QSD.

(ii) If P is primitive, i.e. irreducible and aperiodic, then for any initial distribution μ,

$$\lim_{t\to\infty} \mathbb{P}_\mu(Y_t \in \cdot | \tau > t) = \nu,$$

where ν is the unique QSD.

As last part of this introduction section, we want to describe more formally what stated in the previous chapter. Let $G = (V, E)$ be a finite, connected graph with vertex set V and edge set E. We will study on this thesis a discrete-time version of the voter model on G. To this aim, recall that the state space of the voter model is W^V, where W is the set of phases (or opinions) and V is the vertex set of G; if not differently specified, we will consider $W = \{0, 1\}$. Moreover, the whole process is described by the coordinate functions $\eta_t(\cdot) : V \to W$, for every $t \in \mathbb{Z}_+$, where $\eta_t(v)$ represents the opinion of the vertex v at time t. A state of the system is instead a function $\eta : V \to W$, it represents a sort of snapshot of all the opinions on a fixed time. The evolution of the opinions proceeds as follows: a vertex is picked uniformly, this vertex samples a neighbor uniformly, then the former vertex adopts the opinion of the latter. We call *consensus states* all the states η so that all the opinions are the same. Note that the set of absorbing states Δ coincides with the set of consensus state. Since G is connected, the model reaches a consensus almost surely. In fact, by construction, the hypothesis (2) of Δ is fulfilled.
We write $\boldsymbol{\eta} = \{\eta_t, t \in \mathbb{Z}_+\}$ for the discrete-time voter model on G. The probability of a transition from η to η' is positive if and only if there exists $(v, v') \in V \times V$ such that

(1) $\{v, v'\} \in E$, the edge set of G.

(2) $\eta'(v) = \eta(v')$. That is, the new opinion of the first vertex is taken from the old opinion of the second one.

(3) $\eta'(u) = \eta(u)$ for all $u \neq v$. That is, only one opinion at time step can change.

Now if the pair η and η' satisfy the above conditions, then a transition is obtained by first uniformly sampling the vertex v among all those for which a matching v' exists, and then adopting the opinion of v'. This leads to the following transition function:

$$p(\eta, \eta') = \frac{1}{|V|} \sum_{v\in V} \sum_{v' : \{v,v'\}\in E} \frac{\mathbb{1}_{\eta'(v)}(\eta(v'))}{\deg(v)} \prod_{v\neq v'} \mathbb{1}_{\eta'(u)}(\eta(u)).$$

All other transitions are not allowed.
We can now rewrite the absorption time in the context of the voter model, as we have seen in (3.4), i.e.

$$\tau = \inf\{t \geq 0 : \eta_t(v) = \eta_t(v') \quad \forall v, v' \in V\}.$$

Eventually, using both (4.3) and Proposition 4.1, we write $\lambda_{\mathrm{V}}(G,\mu)$ for the spectral radius associated to the voter model on the graph[1] G, with initial distribution μ. It thus holds

$$\lambda_{\mathrm{V}}(G,\mu) = \lim_{t\to\infty} \mathbb{P}_{\mu}(\tau > t)^{\frac{1}{t}}.$$

We will see that the search for $\lambda_{\mathrm{V}}(G,\mu)$ will be crucial for our problem. At this exact point the duality of the model comes to our aid. In the next section we will prove a result which, under some reasonable conditions, will allow us to compute $\lambda_{\mathrm{V}}(G,\mu)$ in a very clever way.

4.2 Duality approach

The foundation of our analysis is based on the well-known duality between the continuous-time voter model and coalescing random walks, described partially in Section 3.2. In this section, we develop a discrete-time analogue of this duality. In order not to lose generality, we will consider the voter model on a general finite connected graph $G = (V, E)$. Moreover, we assume that the opinion set W can be as large as the number of vertices, i.e., at most we could consider that initially each vertex has a different opinion. For a discussion (and application) of the bipartite graph case with opinions $\{0, 1\}$, we refer to Section 4.3.

As anticipated before, the first step towards finding a QSD is identifying $\lambda_{\mathrm{V}}(G,\mu)$. The connection with coalescing random walks that will be described in this section simplifies the analysis of the time until consensus, by identifying the distribution of τ with the distribution of the time until *two* random walks on the graph first meet. The idea is to describe the propagation of opinions back in time, tracing whose opinion each vertex inherited from previous steps, going all the way back to time zero.

For the following, almost algorithmic, description of the coalescing context, we completely refer to [2]. In passing from time $t-1$ to t in the voter model, we first uniformly select a vertex v, then uniformly select a neighbor u and assign $\eta_t(v) = \eta_{t-1}(u)$. For each $t \in \mathbb{Z}_{>1}$, the sampling of vertex v and its neighbor are independent of and identically distributed as the respective sampling for other times. Furthermore, this sampling is also independent of the actual opinions up to time $t-1$. Fix some time $T \in \mathbb{N}$. We want to trace whose of the original opinions, the ones at $t = 0$, each vertex holds at time T. To this aim, we construct a random directed graph $\mathcal{G}_T$ on $V \times \{0, \dots, T\}$ which evolves exactly the same as the voter model, but back in time. Below we indicate the evolution foreward in time of the voter model with $t \in \{0, \dots, T\}$, while the backward evolution of $\mathcal{G}_T$ is given by $n = T - t$, $t \in \{0, \dots, T\}$. Equivalently, we can define t in terms of n

[1] We mean the spectral radius of the adjacency matrix associated to the finite graph $G = (V, E)$, i.e. the matrix $Ad = \{a_{i,j}\}$ such that $a_{i,j} = 1$ if $\{i, j\} \in E$ and 0 otherwise.

as $t = T - n$.
We now describe the three-steps construction of the random walks on $\mathcal{G}_T$.

1. *Adopting other's opinions.* If at time $t = T - n$ the vertex v is selected to adopt the opinion of vertex u at time $t - 1 = T - (n+1)$, we will draw a directed arrow $(v, n) \to (u, n+1)$, where $(v, n), (u, n+1) \in V \times \{0, \ldots, T\}$. We begin from $n = 0$ up to $n = T$ or, equivalently, from $t = T$ up to $t = 0$. See Figure 4.1 below for an illustration of this procedure on a star graph with $T = 6$.

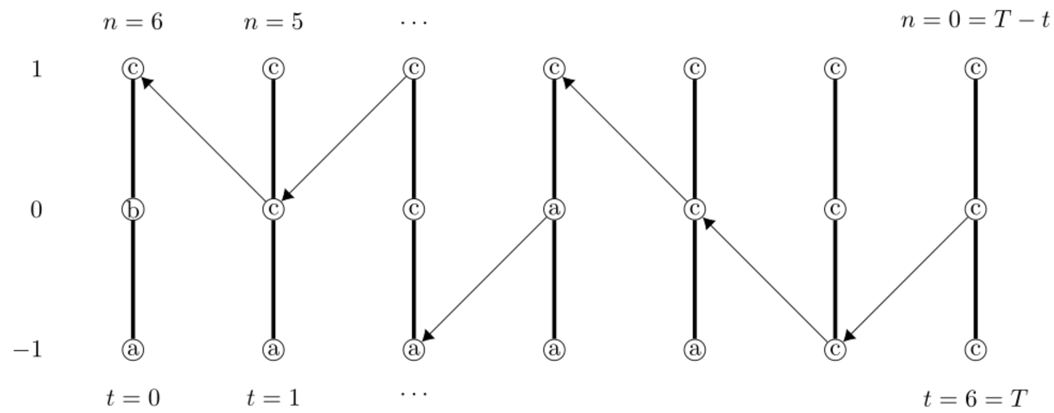

Figure 4.1: Realization of the voter model on a star graph of order 3, where the vertices are the circles, labeled $0, \pm 1$, and the edges are the vertical line segments. Here we consider $W = \{a, b, c\}$ as set of opinions . Note that the t-time of the voter model runs from left to right and appears at the bottom, while the n-time of the random directed graph runs from right to left and appears at the top. An arrow from (v, n) to $(u, n+1)$ represents vertex v adopting at time $t = T - n$ the opinion of vertex u at time $t - 1 = T - (n+1)$. For example, the first arrow from $(1, 6)$ to $(0, 5)$ means that vertex 0 (which has initial opinion b) adopts the opinion c of the near vertex 1 at time $t = 1$.

2. *Keeping one's opinion* Since only one vertex can possibly change opinion in a single time step, in the previous figure we add dashed vectors to represent it. More precisely, let $h(n)$ be the unique vertex (v, n) with an arrow to some $(u, n+1)$, as obtained in Step 1. For all n and $v \in V \setminus \{h(n)\}$, we draw an arrow from (n, n) to $(v, n+1)$. See Figure 4.2 for an illustration of this step.

3. *Construction of paths.* Considering both Step 1 and Step 2, for every $v \in V$ there exists a unique path from $(v, 0)$ to $(\cdot, T)$. This corresponds to the backward tracing of the opinion of $(\cdot, T)$, which is clearly the same of the one of $(v, 0)$. Given $v \in V$, the unique path is a sequence $(v_0 = v, 0) \to \cdots \to (v_T, T)$ where, given $v_n \in V$, $v_{n+1} \in V$ is the unique vertex such that $(v_n, n) \to (v_{n+1}, n+1)$. Since the path

is determined by the choice of v and T, we denote it by $\mathbf{X}^T(v) := \{X_n^T(v) : n = 0, \ldots, T\}$, where $X_n^T(v) = v_n$. Finally, in order to complete our construction, we remove from the graph $\mathcal{G}_T$ all the arrows (dashed or not) that are not used for any path $(v_0, 0) \to \cdots \to (v_T, T)$. This final step is illustrated in Figure 4.3.

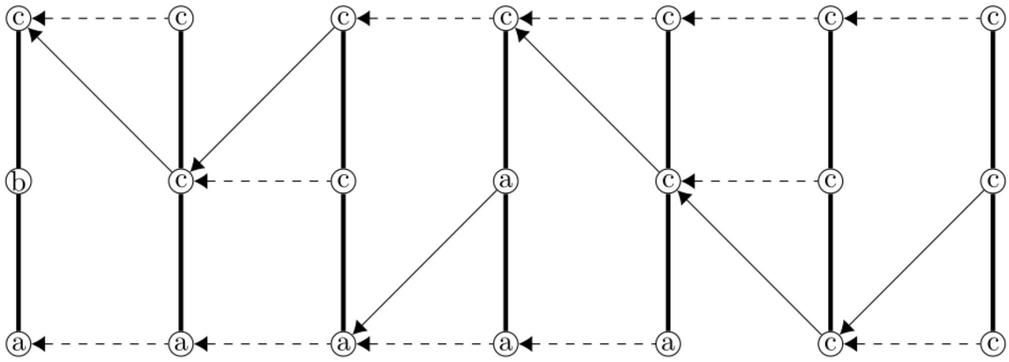

Figure 4.2: We added dashed horizontal arrows to Figure 4.1, which represent all the vertices that keep their opinion in a time step.

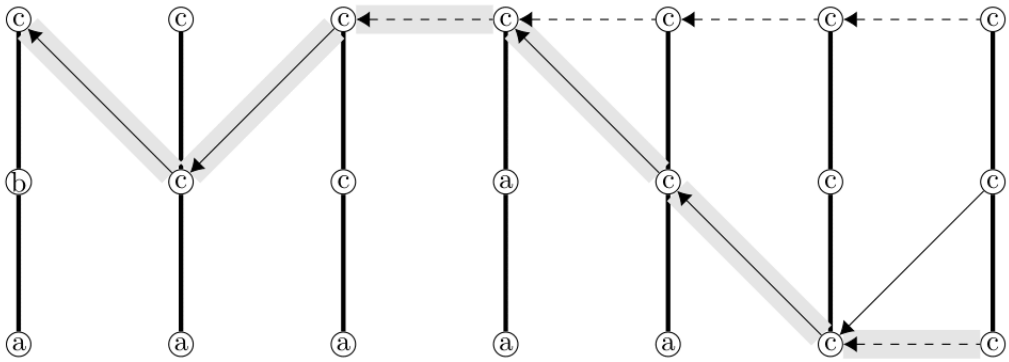

Figure 4.3: The random graph $\mathcal{G}_6$, obtained after removing useless arrows from Figure 4.2. The path of the random walk $\mathbf{X}^6(-1)$ has a shadow on it. Moreover, it is useful to note that the path of the random walk $\mathbf{X}^6(0)$ coincides (or better, coalesces) with the one of $\mathbf{X}^6(-1)$ from time $n = 1$ onwards; same thing happens to the path of $\mathbf{X}^6(1)$, from time $n = 3$ onwards. Observe that it all happens because at time $t = T$ all the vertex have the same opinion.

From our construction, for each $v \in V$, $\mathbf{X}^T(v)$ is a Markov chain on G whose initial distribution is the point-mass at v, i.e. δ_v, and with transition functions given by

$$p(u, u) = \frac{|V| - 1}{|V|}, \quad p(u, w) = \frac{1}{|V|} \frac{1}{\deg(u)} \mathbb{1}_E(u, w), \quad \forall u, w \in V.$$

Equivalently, $\mathbf{X}^T(v)$ is a lazy random walk on G with probability $1 - \frac{1}{|V|}$ of staying put at each step.
Getting inspired by the Markov chain $\mathbf{X}^T(v)$, we are now ready to build a proper system of coalescing random walks (CRWs). Define a system of coupled random walks $\mathbf{Y} = \{Y_t(v) : v \in V, t \in \mathbb{Z}_+\}$ as follows. Set first $Y_0(v) = v$ for each $v \in V$; if we were talking about the voter model, we would have said that each initial opinion is different. Assume that $Y_s(\cdot)$ is defined for $s \leq t$, we uniformly and independently sample $v \in V$ and a neighbour u of v. If $Y_t(\cdot) \neq v$, then we set $Y_{t+1}(\cdot) = Y_t(\cdot)$. Otherwise we set $Y_{t+1}(\cdot) = u$. We will refer to $\mathbf{Y}$ as the *coalescing random walks* on the graph G. The idea is the same of the one mentioned in Section 3.3: we start with $|V|$ random walks on a graph, one for each vertex, they evolve independently, and once two of them meet in a common vertex they became a single new random walk. Note that if we look at the distribution of $\mathbf{Y}$ until a time $T \in \mathbb{N}$ fixed, it coincides with the joint distribution of $\mathbf{X}^T(v)$, $v \in V$. In fact, for any two distinct vertices $v, v' \in V$, let

$$\sigma^T_{v,v'} := \inf\{n \in \{0, \dots, T\} : X^T_n(v) = X^T_n(v')\}, \quad \text{with } \inf \emptyset = \infty,$$

and let

$$\sigma_{v,v'} := \inf\{t \in \mathbb{Z}_+ : Y_t(v) = Y_t(v')\}.$$

Then

$$\mathbb{P}(\sigma^T_{v,v'} = t) = \mathbb{P}(\sigma_{v,v'} = t), \quad \forall t \leq T. \tag{4.4}$$

Also, let

$$\sigma := \max_{v,v'} \sigma_{v,v'}.$$

Before we continue, we would like to recall a fundamental but useful fact. Suppose that η is a $\mathbb{Z}_+$- valued random variable with the property that for some $\lambda \in [0, 1]$,

$$\lambda = \lim_{t\to\infty} \mathbb{P}(\eta > t)^{\frac{1}{t}}.$$

If we take $\rho \geq 1$, we can write

$$\begin{aligned} \mathbb{E}[\rho^\eta] &= \sum_{k=1}^{\infty} \rho^k(\mathbb{P}(\eta > k-1) - \mathbb{P}(\eta > k)) = \rho\mathbb{P}(\eta > 0) + \sum_{k=2}^{\infty} \rho^k \mathbb{P}(\eta > k-1) \\ &\quad - \sum_{k=1}^{\infty} \rho^k \mathbb{P}(\eta > k) = \rho + (\rho - 1) \sum_{k=1}^{\infty} \rho^k \mathbb{P}(\eta > k). \end{aligned} \tag{4.5}$$

It follows that the radius of convergence of $\mathbb{E}[\rho^\eta]$ coincides with the one of $\sum_{k=1}^{\infty} \rho^k \mathbb{P}(\eta > k)$. By the Cauchy-Hadamard theorem, it is equal to $1/\lambda$, and in particular

$$\frac{1}{\lambda} = \sup\{\rho : \mathbb{E}[\rho^\eta] < \infty\}. \tag{4.6}$$

Now define

$$\lambda_{\mathrm{CRW}}(G) := \lim_{t\to\infty} \mathbb{P}(\sigma > t)^{\frac{1}{t}},$$

and recall that

$$\lambda_{\mathrm{V}}(G,\mu) = \lim_{t\to\infty} \mathbb{P}_\mu(\tau > t)^{\frac{1}{t}}.$$

Both limits exists as σ and τ are hitting times of finite-state Markov chains and decay geometrically (possibly with a polynomial correction). Thus, from (4.6) it follows that

$$\frac{1}{\lambda_{\mathrm{CRW}}(G)} = \sup\{\rho : \mathbb{E}[\rho^\sigma] < \infty\}, \tag{4.7}$$

as well as

$$\frac{1}{\lambda_{\mathrm{V}}(G,\mu)} = \sup\{\rho : \mathbb{E}_\mu[\rho^\tau] < \infty\}. \tag{4.8}$$

We are finally ready to state the first result of this section.

Proposition 4.3. *Let $G = (V, E)$ be a finite connected graph. Then, for any initial opinion distribution μ for the voter model on G, it holds*

$$\lambda_V(G,\mu) \le \lambda_{CRW}(G). \tag{4.9}$$

Proof. We want to compare the distributions of the consensus time τ and the coalescing time σ. Let us denote with $\mathbf{i}$ the state of the voter model where all initial opinions are distinct, i.e. where $\eta_0(v) = v$, $v \in V$. We observe first that under $\mathbb{P}_{\mathbf{i}}$ the distribution of τ and σ coincide, that is

$$\mathbb{P}_{\mathbf{i}}(\tau > t) = \mathbb{P}(\sigma > t), \quad \forall t \in \mathbb{Z}_+.$$

This happens because, as noted in the construction of $\mathbf{Y}$, the system of coalescing random walks can be seen as a voter model back in time where all initial opinions are distinct, since at time $t = 0$ we have that $Y_0(v) = v$, for all $v \in V$. Thus, under $\mathbb{P}_{\mathbf{i}}$ the time to reach consensus will be equal to the time for which all random walks coalesce into a single one. Let now μ be any initial distribution for the voter model on G. For a reason similar to the above, the distribution of τ under $\mathbb{P}_\mu$ is stochastically dominated by its distribution under $\mathbb{P}_{\mathbf{i}}$, i.e.

$$\mathbb{P}_\mu(\tau > t) \le \mathbb{P}_{\mathbf{i}}(\tau > t), \quad \forall t \in \mathbb{Z}_+.$$

It follows that

$$\mathbb{P}_\mu(\tau > t) \le \mathbb{P}(\sigma > t), \quad \forall t \in \mathbb{Z}_+.$$

Finally, using the definitions of $\lambda_{\mathrm{V}}(G,\mu)$ and $\lambda_{\mathrm{CRW}}(G)$, we can conclude that

$$\lambda_{\mathrm{V}}(G,\mu) \le \lambda_{\mathrm{CRW}}(G),$$

and $\mu = \mathbf{i}$ is a sufficient condition for equality. □

We will now examine others, more general, sufficient conditions for the equality in (4.9). For this purpose, we state the following

Proposition 4.4. *Let $G = (V, E)$ a finite connected graph and consider the voter model $\{\eta_t, t \in \mathbb{Z}_+\}$ on it. If one of the following two options holds:*

(i) With positive probability, all initial opinions are distinct;

(ii) There exists $t \in \mathbb{Z}_+$, such that for every distinct $v, v' \in V$, $\mathbb{P}_\mu(\eta_t(v) \neq \eta_t(v')) > 0$,

then

$$\lambda_V(G, \mu) = \lambda_{CRW}(G).$$

Proof. In view of Proposition 4.3, it is enough to prove the reverse inequality $\lambda_{\mathrm{V}}(G, \mu) \geq \lambda_{\mathrm{CRW}}(G)$. Suppose first μ to be such that $\mathbb{P}_\mu(\eta_0(v) \neq \eta_0(v'), \forall v, v' \in V) > 0$. For $\rho \geq 1$, it holds

$$\max_{v,v' \in V} \mathbb{E}[\rho^{\sigma_{v,v'}}] \overset{(1)}{\leq} \mathbb{E}[\rho^{\sigma}] \leq \sum_{v,v' \in V} \mathbb{E}[\rho^{\sigma_{v,v'}}], \tag{4.10}$$

where in (1) we used Jensen's inequality. It therefore follows that $\mathbb{E}[\rho^\sigma] < \infty$ if and only if[2] $\max_{v,v' \in V} \mathbb{E}[\rho^{\sigma_{v,v'}}] < \infty$. Now we can write

$$\begin{aligned} \lambda_{\mathrm{CRW}}(G) &\overset{(4.7)}{=} \frac{1}{\sup\{\rho : \mathbb{E}[\rho^\sigma] < \infty\}} \overset{(4.10)}{=} \frac{1}{\sup\{\rho : \max_{v,v'} \mathbb{E}[\rho^{\sigma_{v,v'}}] < \infty\}} \\ &= \max_{v,v'} \frac{1}{\sup\{\rho : \mathbb{E}[\rho^{\sigma_{v,v'}}] < \infty\}} \overset{(4.6)}{=} \max_{v,v'} \lim_{t \to \infty} \mathbb{P}(\sigma_{v,v'} > t)^{\frac{1}{t}} \\ &= \lim_{t \to \infty} \left(\max_{v,v'} \mathbb{P}(\sigma_{v,v'} > t) \right)^{\frac{1}{t}}. \end{aligned} \tag{4.11}$$

Let $v, v' \in V$ be different and fix $\bar{t} \in \mathbb{Z}_+$, which later will be a general time replaced by $t \in \mathbb{Z}_+$. We make this distinction in order to exploit the properties of $\mathbf{X}^{\bar{t}}(v)$. If we suppose that $X_t^{\bar{t}}(v) = u \neq u' = X_t^{\bar{t}}(v')$ and $\eta_0(u) \neq \eta_0(u')$, then necessarily $\eta_t(v) = \eta_0(u)$ and $\eta_t(v') = \eta_0(u')$ since: $\eta_0(u) \neq \eta_0(u')$ implies that we are dealing with two different CRW, and $u \neq u'$ means that until time $t \leq \bar{t}$ they haven't met yet. Hence, at time t there are at least two vertices with two distinct opinions, i.e. $\tau > t$. Summing, we have

$$\mathbb{P}_\mu(\tau > t) \geq \sum_{u \neq u'} \mathbb{P}_\mu(X_t^{\bar{t}}(v) = u, X_t^{\bar{t}}(v') = u', \eta_0(u) \neq \eta_0(u')),$$

and we can sum over all u, u' because $\eta_0(u) \neq \eta_0(u')$ implies $u \neq u'$. Moreover, note that we put $\geq$ instead of $=$ because that is only a sufficient condition for $\tau > t$. Since

[2] The "if" part comes from (4.10), while the "only if" follows from the fact that $\max_{v,v' \in V} \mathbb{E}[\rho^{\sigma_{v,v'}}] \geq \mathbb{E}[\rho^{\sigma_{v,v'}}]$, for all $v, v' \in V$, so in particular it applies to the couple v, v' that fulfills the max in σ.

η_0 is independent of the random walks, we can decouple the condition on the RWs from the condition on the initial opinions, and limit the summation only to pairs u, u' where $u \neq u'$. We can now use (4.4) in order to get

$$\begin{aligned} \mathbb{P}_\mu(\tau > t) &\geq \sum_{u \neq u'} \mathbb{P}_\mu(X_t^{\bar{t}}(v) = u, X_t^{\bar{t}}(v') = u')\, \mathbb{P}_\mu(\eta_0(u) \neq \eta_0(u')) \\ &\geq \sum_{u \neq u'} \mathbb{P}_\mu(X_t^{\bar{t}}(v) = u, X_t^{\bar{t}}(v') = u')\, c \\ &\overset{(2)}{=} \mathbb{P}(\sigma_{v,v'} > t)\, c, \end{aligned} \tag{4.12}$$

where

$$c := \min_{u \neq u'} \mathbb{P}_\mu(\eta_0(u) \neq \eta_0(u')) \overset{(3)}{\geq} \mathbb{P}_\mu(\eta_0(v) \neq \eta_0(v'), \forall v, v' \in V). \tag{4.13}$$

The inequality in (3) holds because the event $\{\text{all initial conditions are distinct}\} = \bigcap_{u \neq u'} \{\eta_0(u) \neq \eta_0(u')\}$ is contained in the event $\{\eta_0(u) \neq \eta_0(u')\}$ for all $u \neq u'$, therefore the containment remains if we take the pair u, u' that realize the minimum value of c. While in (2) we rewrote, using (4.4) and the independence of the CRWs, the event $\{\sigma_{v,v'} > t\}$ decomposing it in the union of all the possible cases in which it occurs.
As long as $c > 0$, this implies that the geometric decay of τ starting from μ is at least as slow as that of $\sigma_{v,v'}$. In our hypothesis, by (i) and (4.13), c is always strictly positive. Thus, by arbitrariness of $v, v' \in V$, and using both (4.12) and (4.11), we get

$$\lambda_{\text{CRW}}(G) \leq \lambda_{\text{V}}(G, \mu).$$

Hence, the sufficient condition (i) for the equality has been proved.
Next, we relax the condition for equality a little further using (ii). Suppose that μ_0 is an initial distribution on any number of opinions such that for some $t_0 \in \mathbb{Z}_+$, we have $\mathbb{P}_{\mu_0}(\eta_{t_0}(u) \neq \eta_{t_0}(u')) > 0$ for every pair of vertices $u \neq u'$. Denote the distribution of η_{t_0} with μ and note that $\lambda_{\text{V}}(G, \mu) = \lambda_{\text{CRW}}(G)$ follows from the previous case, since c in (4.13) is strictly positive. Thus, we are only left to determine the connection between $\lambda_{\text{V}}(G, \mu_0)$ and $\lambda_{\text{V}}(G, \mu)$; we remark the fact that the initial distribution is μ_0, not μ. To this aim, for any $\rho \geq 1$ we can use the Markov property to write

$$\begin{aligned} \mathbb{E}_{\mu_0}[\rho^\tau] \geq \mathbb{E}[\rho^\tau, \tau > t] &= \mathbb{E}_{\mu_0}\left[(1 - \mathbb{1}_{\tau \leq t_0})\rho^{t_0}\mathbb{E}_\mu[\rho^\tau]\right] \\ &= \rho^{t_0}\Big(\mathbb{E}_{\mu_0}\left[\mathbb{E}_\mu[\rho^\tau]\right] - \mathbb{E}_{\mu_0}\left[\mathbb{1}_{\tau \leq t_0}\mathbb{E}_\mu[\rho^\tau]\right]\Big) \\ &\geq \rho^{t_0}(\mathbb{E}_\mu[\rho^\tau] - 1). \end{aligned}$$

Hence if $\mathbb{E}_\mu[\rho^\tau]$ if infinite, then so is $\mathbb{E}_{\mu_0}[\rho^\tau]$. Therefore, it follows from (4.8) that $\lambda_{\text{V}}(G, \mu_0) \geq \lambda_{\text{V}}(G, \mu)$. Thus $\lambda_{\text{V}}(G, \mu_0) \geq \lambda_{\text{CRW}}(G)$ and the equality follows by (4.9), which allows us to conclude the proof. $\square$

An important consequence is the following

Corollary 4.5. *Let $G = (V, E)$ be a finite connected graph and suppose that μ is the initial distribution on the opinions $\{0, 1\}$ where all vertices have opinion 0 except for one uniformly chosen vertex which has opinion 1. Then*

$$\lambda_V(G, \mu) = \lambda_{CRW}(G).$$

As for the proof, we exploit Proposition 4.4, in particular the sufficient condition (ii), since at time $t = 0$ there is a positive probability that every pair of different vertices have distinct opinions.

We conclude this section exposing an example where we apply all the result that we discussed. We consider the complete graph of grade n, K_n, and we look for the limit distribution of the (unique) QSD as $n \to \infty$. We want to evidence the fact that even the existence of a limiting distribution is not obvious. The same example is reported in [2], Section 4.

Example 1. Let K_n be the complete graph[3] with n vertices and consider the voter model on K_n, for $n \geq 3$, with two opinions: "yes" and "no". This Markov chain, when restricted to the nonabsorbing states, is irreducible and aperiodic. Hence it follows from Theorem 4.2 that the chain conditioned on nonabsorbtion converges to the unique QSD. Writing down the eigenvalue equation from Proposition 4.1, we have

$$\begin{aligned}\lambda\,\nu_n(k) &= \frac{k(k-1) + (n-k)(n-k+1)}{n(n-1)}\nu_n(k) \\ &\quad + \frac{(k+1)(n-k-1)}{n(n-1)}\nu_n(k+1) + \frac{(n-k+1)(k-1)}{n(n-1)}\nu_n(k-1),\end{aligned}$$

where k represents the number of "yes" opinions, and $\nu_n(0) = \nu_n(n) = 0$. The equation is solved by taking $\nu_n(k) = \frac{1}{n-1}$ and $\lambda = 1 - \frac{2}{n(n-1)}$. If we now choose an initial distribution μ such that it is not supported on the consensus states $\{\eta : \eta(x) = 1,\ \forall x \in V\}$ and $\{\eta : \eta(x) = 0,\ \forall x \in V\}$, by irriducibility of the chain we have that (ii) of Proposition 4.4 holds, and therefore $\lambda_{\text{CRW}}(G) = \lambda_{\text{V}}(G, \mu)$. By (4.11), we have to look only at two random walks on K_n at different vertices. Following the evolution of $\mathbf{Y}$ for two walkers, the probability that they meet in the next step is $\frac{2}{n(n-1)}$. Thus, again by (4.11), we have that $\lambda_{\text{CRW}}(K_n) = 1 - \frac{2}{n(n-1)}$, as established above through a direct calculation of the QSD.

Summarizing, the QSD for the voter model on the complete graph K_n is uniform on the nonabsorbing states. In particular, the QSDs do not converge to a probability distribution as $n \to \infty$, since $\nu_n(k) \to 0$ for all $k \in \{0, \ldots, n\}$, as $n \to \infty$. In the next

[3] A complete graph is a simple undirected graph in which every pair of distinct vertices is connected by a unique edge.

section, we will study the behaviour of the same limit on the bipartite graph $K_{n,m}$. We will see that, this time, we will have convergence of the QSD to a non-trivial probability distribution, as $n \to \infty$ and $m \in \mathbb{N}$ fixed.

4.3 Main results on Complete Bipartite Graphs

Recall that a complete bipartite graph $K_{n,m} = (V, E)$ is a graph whose vertex set V is the disjoint union of L and S, where $|L| = n$ and $|S| = m$, $m \leq n$, and its edge set is $E = \{\{l, s\} : l \in L, s \in S\}$. We will study the QSDs for the voter model on $K_{n,m}$ with two opinions, "1" and "0", also referred as "yes" and "no", respectively. As noted before, since in the voter interpretation L represent a very large group of people compared to those of S, we assume $m \leq n$. In Figure 4.4 above we present an example where $n = 10$, $m = 5$, and the two opinions "1" and "0" are represented with colors red and green. We will also assume the following additional conditions which we need in order to guarantee irreducibility:

$$\begin{aligned} &m \geq 2 \text{ or} \\ &m = 1 \text{ and } n \geq 3. \end{aligned} \tag{4.14}$$

The set Δ of absorbing states for the voter model on $K_{n,m}$ is given by

$$\Delta = \{\eta : \eta(x) = 1 \quad \forall x \in V, \text{or } \ \eta(x) = 0 \quad \forall x \in V\}.$$

In addiction, there is a set of states

$$BP := \{\eta : \eta(l) = 1 - \eta(s),\ l \in L, s \in S\},$$

which are not accessible from any other state not in BP. This is due to the fact that the number of opinions is equal to the number of partitions of V. In the next chapter, where we will try to generalize some of the results given in the bipartite graphs to the k-partite graphs ($k \geq 3$), we will see that this set of boundary points will no more be a problem. Thus, we will eliminate the subsets Δ and BP from our state space of the model. Since under (4.14) the voter model on $K_{n,m}$ with two opinions is aperiodic and irreducible, by Theorem 4.2 and Proposition 4.1 it follows that, starting from a distribution μ supported on $(\Delta \cup BP)^c$ and conditioning on not reaching consensus, the model converges to the unique QSD which is also supported on $(\Delta \cup BP)^c$. We denote this QSD with $\pi_{n,m}$ and note that it is a left eigenvector corresponding to the eigenvalue $\lambda_{n,m} = \lambda_{\mathrm{V}}(K_{n,m}, \mu) = \lambda_{\mathrm{CRW}}(K_{n,m})$, for the restriction of the transition function of the voter model to $(\Delta \cup BP)^c$. The main idea behind this construction is the following: we want to find the QSD for our model and to do this we need its eigenvalue $\lambda_{n,m} = \lambda_{\mathrm{V}}(K_{n,m}, \mu)$, but this is difficult in general, thus we exploit the duality of the model in order to obtain the equality $\lambda_{\mathrm{V}}(K_{n,m}, \mu) = \lambda_{\mathrm{CRW}}(K_{n,m})$. We therefore reduced the

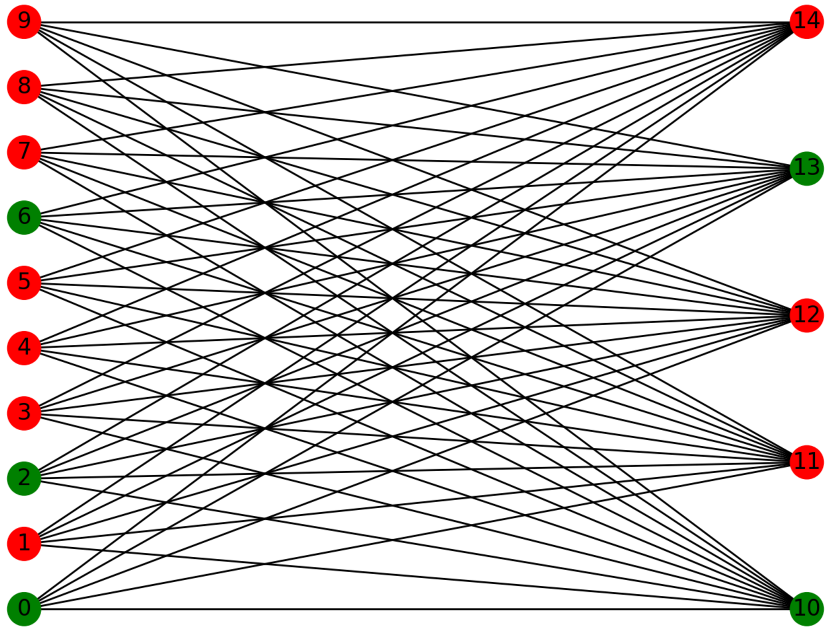

Figure 4.4: Realization of the voter model on the complete bipartite graph $K_{10,5}$. On the left is presented the set $L = \{0, \dots, 9\}$, while on the right $S = \{10, \dots, 14\}$. The circles represent the vertices of the graph and the black lines the edges, while the two opinions are represented with colors red and green inside the circles.

problem of finding the geometric rate of the time to absorption in the voter model, i.e. $\lambda_{\mathrm{V}}(K_{n,m}, \mu)$, to searching the time for which two distinct random walks meet for the first time on $K_{n,m}$.
To continue our analysis, we will exploit the symmetry among vertices within each group. Instead of following the opinion on each vertex, we will follow the number of "yes" opinions in each of the groups S and L. This leads to a Markov chain on the state space $\{0, \dots, n\} \times \{0, \dots, m\}$. Each state is an ordered pair (k, h), with k representing the number of "yes" in group L and h representing the number of "yes" in S. The only allowed transitions are the following

(i) $(k, h) \to (k + 1, h)$. This happens if a "no" vertex in L is sampled and adopts a "yes" from S. Recall that the evolution consists in choosing uniformly a vertex and, independently, in choosing a neighbor whose opinion is taken from the first. Thus, the above transition happens with probability $\frac{n-k}{n+m}\frac{h}{m}$.

(ii) $(k, h) \to (k - 1, h)$. This happens if a "yes" vertex in L is sampled and adopts a "no" from S. The probability is therefore $\frac{k}{n+m}\frac{m-h}{m}$.

(iii) $(k, h) \to (k, h + 1)$. This happens if a "no" vertex in S is sampled and adopts a "yes" from L. The probability is therefore $\frac{m-h}{n+m}\frac{k}{n}$.

(iv) $(k,h) \to (k,h-1)$. This happens if a "yes" vertex in S is sampled and adopts a "no" from L. The probability is therefore $\frac{h}{n+m}\frac{n-k}{n}$.

(v) $(k,h) \to (k,h)$. This happens with probability $\frac{k}{n+m}\frac{h}{m} + \frac{h}{n+m}\frac{k}{m} + \frac{n-k}{n+m}\frac{m-h}{m} + \frac{m-h}{n+m}\frac{n-k}{n} = \frac{kh+(n-k)(m-h)}{nm}$.

Of course, $(0,0)$,(n,m) are the unique absorbing states, and the set BP collapses into two states, $(0,m)$ and $(n,0)$, not accessible from any other state. Thus eliminating these four states, the new chain is irreducible. As a result, it possesses a unique QSD which we denote by $\mu_{n,m}$. Furthermore, the absorption time for the new chain from any initial state coincides with the time to absorption for the voter model starting from any state with matching numbers of opinions in both S and L. Thus, if we apply Proposition 4.1 to the new chain, we get that the eigenvalue corresponding to $\mu_{n,m}$ is equal to $\lambda_{n,m}$. Now fix any state (k,h) of our chain. Adapting all the possible transitions to the corresponding state, we obtain the following eigenvalue equation for $\mu_{n,m}$:

$$\begin{aligned}\lambda_{n,m}\,\mu_{n,m}(k,h) = \mu_{n,m}(k,h)&\frac{kh+(n-k)(m-h)}{nm} \\ + \mu_{n,m}(k-1,h)&\frac{(n-k+1)h}{(n+m)m} + \mu_{n,m}(k+1,h)\frac{(k+1)(m-h)}{(n+m)m} \\ + \mu_{n,m}(k,h-1)&\frac{(m-h+1)k}{(n+m)n} + \mu_{n,m}(k,h+1)\frac{(h+1)(n-k)}{(n+m)n}.\end{aligned} \tag{4.15}$$

At this point, we can finally exploit the duality of the voter model with the coalescing random walks. We are going to use it strongly in the next proof.

Proposition 4.6.

$$\begin{aligned}\lambda_{n,m} = \lambda_{CRW}(K_{n,m}) &= 1 - \frac{2}{n+m}\left(1 - \sqrt{1 - \frac{1}{2n} - \frac{1}{2m}}\right) \\ &= 1 - \frac{\gamma_{n,m}}{n+m},\end{aligned}$$

where

$$\gamma_{n,m} = 2\left(1 - \sqrt{1 - \frac{1}{2n} - \frac{1}{2m}}\right).$$

Proof. We assume first $m > 1$ and, by (4.14), $n \in \mathbb{N}$. From (4.11) it is enough to consider only two coalescing random walks on $K_{n,m}$. We have to look at the evolution of $\mathbf{Y}$ in Section 4.2. The two CRW paths can be in either one of the following states:

1. Both walks are in different vertices of L.

2. Both walks are in different vertices of S.

3. One walk is in S and the other one in L.

4. They are both at the same vertex.

Label these four states of the system as 1, 2, 3, 4, respectively. Of course, 4 is the absorbing state for the CRW, so we will omit it from our calculations. Let us consider the state 1. We know that at every time step only one of the two random walks can possibly move, therefore the probability that the system moves from the state 1 to the state 2 is zero. While the probability to move from 1 to 3 is equal to the sum (recall that the RWs moves independently) of the probability of the events {the first RW moves and goes in S} and {the second RW moves and goes in S}; both events happen with probability $\frac{1}{n+m}$, thus the probability to pass from the state 1 to the state 3 is $\frac{2}{n+m}$. The probability to stay put from each of the states 1,2,3 is equal to the probability that none of the two random walks moves, i.e. $1-\frac{2}{n+m}$. The same reasoning can be applied starting from state 2. We need to be more careful starting from state 3, since we do not want that the CRWs meet at the same vertex. Thus, from state 3 the system can transition to 1 or 2 with respective probabilities $\frac{1}{n+m}\frac{n-1}{n}$ and $\frac{1}{n+m}\frac{m-1}{m}$. As a result the substochastic transition function on states 1, 2, 3 is

$$\begin{pmatrix} \frac{n+m+2}{n+m} & 0 & \frac{2}{n+m} \\ 0 & \frac{n+m+2}{n+m} & \frac{2}{n+m} \\ \frac{1}{n+m}\frac{n-1}{n} & \frac{1}{n+m}\frac{m-1}{m} & \frac{n+m+2}{n+m} \end{pmatrix}. \tag{4.16}$$

Since from both states 1 and 2 the transitions are either to themselves, with the same probability, or to state 3, with the complementary probability, the lumping[4] conditions holds, so we can consolidate these states into one, leading to the matrix

$$\begin{pmatrix} \frac{n+m+2}{n+m} & \frac{2}{n+m} \\ \frac{1}{n+m}(\frac{n-1}{n}+\frac{m-1}{m}) & \frac{n+m+2}{n+m} \end{pmatrix}. \tag{4.17}$$

The characteristic equation is

$$\left(\lambda-\frac{n+m-2}{n+m}\right)^2-\frac{2}{(n+m)^2}\frac{2mn-m-n}{nm}=0.$$

[4]Given a stochastic matrix P, we say that P is a lumpable matrix on a partition T of the state space if and only if, for any subset t_i and t_j in the partition and for any states $n, n' \in t_i$, one has $\sum_{k\in t_j} p(n,k)=\sum_{k\in t_j} p(n',k)$.

Therefore the two eigenvalues, λ_+ and λ_-, are given by

$$\begin{aligned}\lambda_\pm &= 1 - \frac{2}{n+m} \pm \frac{1}{n+m}\sqrt{4 - \frac{2}{n} - \frac{2}{m}} \\ &= 1 - \frac{2}{n+m}\left(1 \mp \sqrt{1 - \frac{1}{2n} - \frac{1}{2m}}\right),\end{aligned}$$

and the largest eigenvalue is obtained by choosing λ_-(with the "-" sign), giving the expression in the statement.
It remains to consider $m = 1$ and $n \geq 3$. In this case, state 2 is not possible. We therefore eliminate the second row and column from (4.16), ending up with the matrix (4.17) and then continue as before. □

We now give a direct relation between $\lambda_{n,m}$ and $\mu_{n,m}$.

Proposition 4.7.

$$\lambda_{n,m} = 1 - \frac{2}{n+m}(\mu_{n,m}(1,0) + \mu_{n,m}(0,1)).$$

Proof. Let (K, H) be the random vector representing the number of "yes" in L and S, respectively, whose distribution is $\mu_{n,m}$. Recall that the set of absorbing states is composed by $(0,0)$ and (n,m), thus, since $\mu_{n,m}$ is not supported on them, we have $\mu_{n,m}(0,0) = \mu_{n,m}(n,m) = 0$. We can therefore sum on both sides of (4.15) over $-1 \leq k \leq n+1$ and $-1 \leq h \leq m+1$, while eliminating from the sum the pairs $(k,h) = (0,0)$ and $(k,h) = (n,m)$ to obtain

$$\begin{aligned}\lambda_{n,m} &= \frac{1}{nm}\big(\mathbb{E}[KH] + \mathbb{E}[(n-K)(m-H)]\big) \\ &+ \frac{1}{m(n+m)}\big(\mathbb{E}[(n-K)H] - \mu_{n,m}(n-1,m)m\big) \\ &+ \frac{1}{m(n+m)}\big(\mathbb{E}[(m-H)K] - \mu_{n,m}(1,0)m\big) \\ &+ \frac{1}{n(n+m)}\big(\mathbb{E}[(m-H)K] - \mu_{n,m}(n,m-1)n\big) \\ &+ \frac{1}{n(n+m)}\big(\mathbb{E}[(n-K)H] - \mu_{n,m}(0,1)n\big) \\ &= 1 - \frac{1}{n+m}(\mu_{n,m}(n-1,m) + \mu_{n,m}(1,0) + \mu_{n,m}(n,m-1) + \mu_{n,m}(0,1)) \\ &= 1 - \frac{2}{n+m}(\mu_{n,m}(1,0) + \mu_{n,m}(0,1)),\end{aligned}$$

where the last equality follows from invariance under relabeling of the two opinions. □

An interesting fact, as we will see in the next chapter, is that in the case of complete tripartite graph (and then in the k-partite) a connection between λ and μ as in (4.7) can still be proved, but a direct computation of λ as in (4.6) will not be possible.

Remark 11. As a secondary result, we now show how to compute $\mu_{n,m}(\jmath, 0)$ in the special case where $m = 1$. As the state (0,1) is in the boundary points BP and not in the support of $\mu_{n,1}$, its measure vanishes and using Proposition 4.6 along with Proposition 4.7 gives

$$\mu_{n,1}(1,0) = \frac{n+1}{2}(1-\lambda_{n,1}) = 1 - \sqrt{\frac{1}{2} - \frac{1}{2n}} = \frac{\gamma_{n,1}}{2}. \tag{4.18}$$

Writing out the eigenvalue equation (4.15) in the case $m = 1$ and $h = 0$ leave us with

$$\Big(\lambda_{n,1} - \frac{n-k}{n}\Big)\mu_{n,1}(k,0) = \mu_{n,1}(k+1,0)\frac{k+1}{n+1} + \mu_{n,1}(k,1)\frac{n-k}{n(n+1)}.$$

Due to invariance under relabelling of the two opinions, we can write $\mu_{n,1}(k,1) = \mu_{n,1}(n-k,0)$. In addition, from Proposition 4.6 we can rewrite $\lambda_{n,1}$ as

$$\lambda_{n,1} - \frac{n-k}{n} = \frac{k}{n} - \frac{\gamma_{n,1}}{n+1},$$

so that we can plug it into the eigenvalue equation above and get

$$\Big(\frac{k}{n} - \frac{\gamma_{n,1}}{n+1}\Big)\mu_{n,1}(k,0) = \mu_{n,1}(k+1,0)\frac{k+1}{n+1} + \mu_{n,1}(n-k,0)\frac{n-k}{n(n+1)}. \tag{4.19}$$

This nonlocal recurrence relation can be solved through iteration. Having calculated $\mu_{n,1}(1,0)$ in (4.18) and recalling that $\mu_{n,1}(n,0) = 0$, we can plug $k = n-1$ into (4.19) and obtain

$$\Big(\frac{n-1}{n} - \frac{\gamma_{n,1}}{n+1}\Big)\mu_{n,1}(n-1,0) = \mu_{n,1}(1,0)\frac{1}{n(n+1)},$$

or

$$\mu_{n,1}(n-1,0) = \frac{\gamma_{n,1}}{2(n^2 - n\gamma_{n,1} - 1)}.$$

We can repeat this procedure inductively. Having calculated $\mu_{n,1}(\jmath,0)$ and $\mu_{n,1}(n-\jmath,0)$ for $\jmath = 1,\dots,k < n$, we can use (4.19) to recover $\mu_{n,1}(k+1,0)$ and then $\mu_{n,1}(n-k-1,0)$.

In the next technical lemma we show the asymptotic behaviour of $\mu_{n,m}$ as $n \to \infty$, in the case where $h \neq 0$.

Lemma 4.8. *Suppose $k \geq 0$ and $h > 0$. Then*

$$\lim_{n\to\infty} \mu_{n,m}(k,h) = 0.$$

Proof. Assume both $k, h > 0$. From the eigenvalue equation for $\mu_{n,m}$ (4.15), we have

$$\Big(\lambda_{n,m} - \frac{kh + (n-k)(m-h)}{nm}\Big)\mu_{n,m}(k,h) = (1+o(1))\frac{h}{m}\mu_{n,m}(k-1,h) + O\Big(\frac{1}{n}\Big), \quad (4.20)$$

as $n \to \infty$. From Proposition 4.6, we know that $\lambda_{n,m} \to 1$ as $n \to \infty$. Moreover, the coefficient of $\mu_{n,m}(h,k)$ on the left-hand side of (4.20) tends to $\frac{h}{m} + o(1)$ as $n \to \infty$. Hence, since $h > 0$, we get

$$\mu_{n,m}(k,h) = (1+o(1))\mu_{n,m}(k-1,h) + O\Big(\frac{1}{n}\Big), \quad \text{as } n \to \infty. \quad (4.21)$$

Now suppose for sake of contradiction that $\limsup_{m\to\infty} \mu_{n,m}(k',h') = \epsilon > 0$ for some $k' \geq 0$ and $h' > 0$. Then (4.21) implies that $\lim_{j\to\infty} \mu_{n_j,m}(k'+1,h') = \epsilon$ along some subsequence $\{n_j\}_{j\geq 0}$. Similarly, (4.21) can be used again to show that $\lim_{j\to\infty} \mu_{n_j,m}(k'+2,h') = \epsilon$ along the same subsequence. Reasoning inductively both forwards and backwards in k, it follows that $\lim_{j\to\infty} \mu_{n_j,m}(k,h') = \epsilon$ for all $k \geq 0$ and $h' > 0$. In particular, for j large enough we have $\mu_{n_j,m}(k'+2,h') > \frac{\epsilon}{2}$ for $0 \leq k \leq \lceil \frac{2}{\epsilon} \rceil$. Hence $\sum_{0\leq k\leq n} \mu_{n_j,m}(k,h') \geq \sum_{0\leq k\leq \lceil \frac{2}{\epsilon}\rceil} \mu_{n_j,m}(k,h') > 1$, a contradiction. □

As a consequence, when we will look for the limit distribution of $\mu_{n,m}$, we will have to concern only to the values on the states $(k,0)$, $k \geq 0$.
Before going on we recall from [9] that the *Sibuya distribution* with parameter $\gamma \in (0,1]$ is a probability distribution on $\mathbb{Z}_+$ with probability mass function f_γ and probability generating function ϕ_γ given by

$$\begin{aligned} f_\gamma(k) &= \frac{\gamma}{k!}\prod_{j=1}^{k-1}(j-\gamma), \quad k \in \mathbb{Z}_+, \\ \phi_\gamma(z) &= 1-(1-z)^\gamma, \quad |z| < 1. \end{aligned} \quad (4.22)$$

When $\gamma \in (0,1)$, the Sibuya distribution is heavy tailed and, in this case, f_γ decays according to a power law with

$$f_\gamma(k) \sim \frac{1}{\pi}\sin(\gamma\pi)\Gamma(1-\gamma)\frac{1}{k^{\gamma+1}}, \quad \text{as } k \to \infty.$$

We denote this probability distribution by $\mathrm{Sib}(\gamma)$. See [2] for a more detailed list of references about it. Furthermore, define

$$\gamma_m := 2\left(1 - \sqrt{1-\frac{2}{m}}\right), \quad (4.23)$$

and note that $\gamma_{n,m} \to \gamma_m$ as $n \to \infty$, where $\gamma_{n,m}$ was defined in Proposition 4.6.
In the following proposition we calculate the pointwise limit of $\mu_{n,m}$ as the size of the large partition tends to infinity.

Proposition 4.9. *Let f_{γ_m} be as (4.22), with γ_m defined in (4.23). Then, for $k \geq 0$*

$$\mu_{\infty,m}(k,0) := \lim_{n\to\infty} \mu_{n,m}(k,0) = \frac{1}{2} f_{\gamma_m}(k).$$

Proof. For Lemma 4.8 we know that $\lim_{n\to\infty} \mu_{n,m}(0,1) = 0$. Therefore it follows from Proposition 4.6 and Proposition 4.7 that

$$\mu_{\infty,m}(1,0) = \frac{\gamma_m}{2}. \tag{4.24}$$

Returning to (4.15) with $k \geq 1$ and $h = 0$ (for $h > 0$ we know that $\mu_{\infty,m}(k,h) = 0$), we can write

$$\lambda_{n,m}\, \mu_{n,m}(k,0) = \mu_{n,m}(k,0)\frac{n-k}{n} + \mu_{n,m}(k+1,0)\frac{k+1}{n+m} + \mu_{n,m}(k,1)\frac{n-k}{n(n+m)}.$$

If we rewrite now $\lambda_{n,m}$ using Proposition 4.6, rearranging the latter equation we can write

$$\mu_{n,m}(k+1,0) = \frac{n+m}{k+1}\left(\mu_{n,m}(k,0)\left(\frac{k}{n} - \frac{\gamma_{n,m}}{n+m}\right) - \mu_{n,m}(k,1)\frac{n-k}{n(n+m)}\right). \tag{4.25}$$

Letting $n \to \infty$ in (4.25) and recalling that $\gamma_{n,m} \to \gamma_m$ and $\mu_{n,m}(k,1) \to 0$, we get

$$\mu_{\infty,m}(k+1,0) = \mu_{\infty,m}(k,0)\frac{k-\gamma_m}{k+1}.$$

We can now argue inductively starting from (4.24) to conclude that

$$\begin{aligned}\mu_{\infty,m}(k,0) &= \frac{\gamma_m}{2}\frac{1}{k!}\prod_{j=1}^{k-1}(j-\gamma_m) \\ &= \frac{1}{2} f_{\gamma_m}(k), \quad k \in \mathbb{Z}_+.\end{aligned}$$

□

Finally we can state the main result of this chapter. We recall once again that all the results are taken from the very recent paper of Ben-Ari, Panzo, Speegle and VandenBerg, in [2].

Theorem 4.10. *Let $C \sim$ Bern$(1/2)$ and $D \sim$ Sib(γ_m) be independent, where γ_m is defined in (4.23) and with Bern(p) we denote a Bernoulli random variable which takes values 0 and 1 with probability $1-p$ and p, respectively. Then the distribution of opinions under the QSD for the voter model on $K_{n,m}$ as $n \to \infty$ converges weakly to the following:*

(i) All vertices of S have opinion C.

(ii) All but D vertices in L have opinion C.

Proof. By the invariance under relabeling of the two opinions, we know that $\mu_{n,m}(k,h) = \mu_{n,m}(n-k, m-h)$. Hence Proposition 4.9 implies that

$$\lim_{n\to\infty} \mu_{n,m}(n-k,m) = \lim_{n\to\infty} \mu_{n,m}(k,0) = \frac{1}{2} f_{\gamma_m}(k), \quad k \in \mathbb{Z}_+. \tag{4.26}$$

Since we know that f_{γ_m} is a probability mass function, it follows from the Portmanteau theorem[5] that the QSDs for the voter model on $K_{n,m}$ converge weakly as $n \to \infty$ to the probability distribution where: S attains a consensus of all "0" or all "1" each with probability $1/2$, and, conditioned on the opinion of S, the number of vertices in L which are of a different opinion has probability mass function f_{γ_m}. □

Remark 12. The same result can be seen using a more direct approach. In fact, using the relabelling of the two opinions, from (4.26) it follows that dissent towards opinions 1 and 0 are analogous, that is

$$\mu_{\infty,m}\{\text{k disagreements towards opinion } 0\} = \frac{1}{2} f_{\gamma_m}(k), \quad \text{and}$$
$$\mu_{\infty,m}\{\text{k disagreements towards opinion } 1\} = \frac{1}{2} f_{\gamma_m}(k).$$

It follows that

$$\begin{aligned}\mu_{\infty,m}\{\exists \text{ disagreements towards opinion } 0\} &= \mu_{\infty,m}\{\exists \text{ disagreements towards opinion } 1\} \\ &= \sum_{k\in\mathbb{N}} \frac{1}{2} f_{\gamma_m}(k) = \frac{1}{2}.\end{aligned}$$

Therefore the system polarize itself to opinions 0 or 1 with equal probability. So, if we indicate with C the prevailing opinion of the system, we have that $C \sim \mathrm{Bern}(1/2)$. By Lemma 4.8, we know that the distribution of *all* the opinions in S, under the QSD as $n \to \infty$, is necessarily the dominant one, i.e. all the vertices in S has opinion C. Again from (4.26), it follows that

$$\mu_{\infty,m}\{\text{k disagreements in L} \,|\, C = \jmath\} = \frac{1}{2} f_{\gamma_m}(k), \quad \jmath \in \{0,1\},$$

that is

$$\mu_{\infty,m}\{\text{k disagreements in L}\,,\, C = \jmath\} = f_{\gamma_m}(k), \quad \jmath \in \{0,1\}.$$

[5]Adapted to our context, gives equivalents conditions for weak convergence of a sequence of probability distributions in a probability space. In particular, we can say that our sequence of probability measures $\{\mu_{n,m}\}_{n\in\mathbb{N}}$ converges weakly to a probability mass function f if $\lim_{t\to\infty} \mu_{n,m}(A) = f(A)$, for all continuity sets A of f, i.e. for all Borel sets A such that $f(\partial A) = 0$.

Thus, indicating with D the number of disagreements in L, we have that D has density mass function equal to $\mathrm{Sib}(\gamma_m)$.

An immediate consequence is the following

Corollary 4.11. *The distribution of the number of disagreements along edges in $K_{n,m}$ under the QSD tends to mD as $n \to \infty$, where $D \sim \mathrm{Sib}(\gamma_m)$.*

Chapter 5

Voter model on Complete k-partite Graphs

In this last chapter we will discuss two possible generalizations of the results given by Ben-Ari et al. in [2] that we treated in Chapter 4. In the first and main generalization we will consider the voter model evolving in a more general set of complete graphs, namely the k-partite ones, with $k \geq 2$. Our goal is to provide the limit behaviour of the QSD for the conditional process as the size of one of the k sets goes to infinity, while the other $k-1$ sizes remain fixed. This will be fully discussed in Section 5.1. Our conclusions shows a result that is completely similar to the one of the bipartite case. In fact, the distribution of opinions under the QSD for the voter model on the k-partite graph K converges to the following setting: all vertices of the fixed $k-1$ sets share the same opinion, which can be 0 or 1 according to a Bernoulli distribution of parameter 1/2, while the totality of dissenters is located in the set whose size explodes and is distributed according a Sibuya distribution of parameter $\gamma \in (0,1)$, where γ depends on the -fixed- sizes of the $k-1$ sets. Through Section 5.2 we will deal again with the case $k = 2$, i.e. going back to bipartite graphs, but this time we will suppose that both sizes of the sets diverges as the size of the first one goes to infinity. We will consider two main cases, depending on the fact that the two sizes grow with the same speed or one prevails over the other. We will show that in both situations the limit in distribution of the QSD does not converge to a probability measure, leading to the same result mentioned in Chapter 4 regarding the complete graphs K_n.

5.1 Complete k-partite graphs

Let $K_{m_0,m_1,\dots,m_{k-1}} = (V,E)$, $k \geq 2$, be a complete k-partite graph whose vertex set V is the disjoint union $\cup_{i=0}^{k-1} S_i$, where $|S_i| = m_i$ for all $i = 0,\dots,k-1$, and $m_i \leq m_0$ for all $i = 1,\dots,k-1$; while its edge set is the disjoint union $E = \cup_{i=0}^{k-1} E_i$, where $E_i = \{\{s_i, s_j\} : s_i \in S_i, s_j \in S_j, \forall j \neq i\}$, for all $i = 0,\dots,k-1$. Denote $K := K_{m_0,m_1,\dots,m_{k-1}}$. We will study QSDs for the voter model on K with two opinions, "0" and "1", also referred to as "no" and "yes", respectively. The evolution of the model is exactly the same as in the complete bipartite graph. Similarly, the set of the absorbing states Δ is

$$\Delta = \{\eta : \eta \equiv 1 \text{ or } \eta \equiv 0\}.$$

Instead of following the opinion on each vertex, we will follow the number of "yes" opinions in each of the groups S_i, $i = 0, \dots, k-1$.
This leads to a Markov chain on the state space $\Omega := \{0, \dots, m_0\} \times \{0, \dots, m_1\} \times \cdots \times \{0, \dots, m_{k-1}\}$. Each state is an ordered k-uple $(h_0, \dots, h_{k-1})$, with h_i representing the number of "yes" opinions in group S_i, for all $i = 0, \dots, k-1$. In view of the study of the asymptotic behaviour of the unique QSD for the process, we will consider the size of the first set S_0 to be dominant with respect to the others, and let $m_0 \to \infty$ while fixing all m_i, $i \neq 0$. In the following, we will write an arrow $(h_0, \dots, h_{k-1}) \to (h_0, \dots, h_i \pm 1, \dots, h_{k-1})$ meaning that in this changing of states only the i-th coordinate changes.
Let us introduce some other notations in other to simplify the expressions: we write $N := \sum_{i=0}^{k-1} m_i$ and $N_i := N - m_i$, for all $i = 0, \dots, k-1$. In this chain, the only allowed transitions are the following

1. $(h_0, \dots, h_{k-1}) \to (h_0, \dots, h_i+1, \dots, h_{k-1})$. This happens if a "no" vertex is sampled in S_i and a "yes" vertex is sampled on one of the others S_j, $j \neq i$. These events are independent because the uniform choice of the adjacent vertex is itself. The probability of such a transition is therefore
$$\frac{m_i - h_i}{N} \frac{\sum_{j \neq i} h_j}{N_i}, \quad \forall i \in \{0, \dots, k-1\}. \tag{5.1}$$

2. $(h_0, \dots, h_{k-1}) \to (h_0, \dots, h_i-1, \dots, h_{k-1})$. Similarly, this happens if a "yes" vertex is sampled in S_i and a "no" vertex is sampled on one of the others S_j, $j \neq i$. The probability of such a transition is therefore
$$\frac{h_i}{N}\left(1 - \frac{\sum_{j \neq i} h_j}{N_i}\right), \quad \forall i \in \{0, \dots, k-1\}. \tag{5.2}$$

3. $(h_0, \dots, h_{k-1}) \to (h_0, \dots, h_{k-1})$. This happens if a "yes" vertex is sampled in S_i and a "yes" vertex is sampled on one of the others S_j, $j \neq i$, for every $i \in \{0, \dots, k-1\}$, or if a "no" vertex is sampled in S_i and a "no" vertex is sampled on one of the others S_j, $j \neq i$, for every $i \in \{0, \dots, k-1\}$. The probability of such a transition is therefore
$$\frac{1}{N} \sum_{i=0}^{k-1} \left[\frac{h_i}{N_i} \sum_{j \neq i} h_j + \frac{(m_i - h_i)(N_i - \sum_{j \neq i} h_j)}{N_i}\right]. \tag{5.3}$$

Clearly, since there are only two opinions, the unique absorbing states are the all zeros k-uple $(0, \dots, 0)$ and $(m_0, \dots, m_{k-1})$. Let now fix any state $(h_0, \dots, h_{k-1})$ of our new chain. Moreover, let us denote with $\mu := \mu_{m_0, \dots, m_{k-1}}$ the unique QSD for the chain supported in $\Omega \setminus \Delta$, and with $\lambda := \lambda_{m_0, \dots, m_{k-1}}$ the corresponding eigenvalue. Under some irreducibility assumptions, as done in the bipartite case, the following eigenvalue equation for μ is fulfilled

$$\begin{aligned}\lambda\,\mu(h_0,\dots,h_{k-1}) = \mu(h_0,\dots,h_{k-1})\,\frac{1}{N}\sum_{i=0}^{k-1}\left[\frac{h_i}{N_i}\sum_{j\neq i}h_j + \frac{(m_i-h_i)(N_i-\sum_{j\neq i}h_j)}{N_i}\right]\\ +\sum_{i=0}^{k-1}\mu(h_0,\dots,h_i+1,\dots,h_{k-1})\frac{(h_i+1)}{N}\left(1-\frac{\sum_{j\neq i}h_j}{N_i}\right)\\ +\sum_{i=0}^{k-1}\mu(h_0,\dots,h_i-1,\dots,h_{k-1})\frac{(m_i-h_i+1)}{N}\frac{\sum_{j\neq i}h_j}{N_i}.\end{aligned} \tag{5.4}$$

Note that we adapted the transitions (5.1),(5.2) and (5.3) to the corresponding states of the eigenvalue equation.
From now on we will enumerate all our attempts of generalization following the steps of Chapter 4 and [2]. In particular, we will consider every intermediate result needed to prove Theorem 4.10 and we aim, whenever possible, to extend it to our context.

Generalization of Proposition 4.6

Just like in the bipartite graphs, it is enough to consider only two coalescing random walks on K. First of all we need to count in how many states the CRWs can be found, and, in order to write the transition matrix, we need to sort each of them. Let us start by the counting problem. It can be solved by enumeration:

1. There are k states corresponding to the events that both the coalescing random walks are in distinct vertices of the same set S_i, for all $i \in \{0,\dots,k-1\}$.

2. Since we want to let $m_0 \to \infty$, we start counting by S_0. There are surely $k-1$ states which corresponds to the event that one CRW is on S_0 and the other one in S_i, for every $i \in \{1,\dots,k-1\}$. Fix now a CRW on S_1. Since the state that one random walk is on S_0 and the other one on S_1 has been already counted before, there are $k-2$ states corresponding to have one CRW on S_1 and the other in S_i, $i \in \{2,\dots,k-1\}$. Thus, reasoning inductively and taking a general $j \in \{0,\dots,k-1\}$, we have $k-1-j$ states corresponding to have a CRW on S_j and the other one on S_i, for every $j < i \leq k-1$. It follows that for all these cases there are $\sum_{j=0}^{k-1} k-1-j = \frac{k^2-k}{2}$ states.

3. There is 1 state corresponding to have the two coalescing random walks on the same vertex in K.

Of course, similarly to the bipartite case, we will exclude the latter state because we are conditioning the system to never reach consensus states (the analogous for CRWs are clearly the coalescing states). We can therefore conclude that there are $\frac{k^2+k}{2}$ possible states for the two CRWs. It follows that the corresponding transition matrix will have a size of $\frac{k^2+k}{2} \times \frac{k^2+k}{2}$.
We now need to sort all these states. To do so, since the main set of our interest is S_0, we chose to define all the possible states as ordered couples (i, j), $i, j \in \{0, \ldots, k-1\}$, $i < j$, that represents the event in which one of the two CRWs is in S_i and the other one in S_j. The restriction $i < j$ refers to the fact that we consider the state (i, j) equivalent to the state (j, i), since we are not interested in which random walk is in S_i or S_j. We decided to sort the states in the following way: the first k are given by (i, i), $i = 0, \ldots, k-1$; then there are $k-1$ states given by $(0, j)$, $j \in \{1, \ldots, k-1\}$; then there are $k-2$ states given by $(1, j)$, $j \in \{2, \ldots, k-1\}$; proceeding this way we end up with the last state $(k-2, k-1)$. By way of example, all the possible (sorted) states with $k = 5$ are

$$(0,0),(1,1),(2,2),(3,3),(4,4),(0,1),(0,2),(0,3),(0,4),(1,2),(1,3),(1,4),(2,3),(2,4),(3,4).$$

We are left to find all the possible probabilities to move from a state to another, and to write the corresponding transition matrix P. Thanks to the way we ordered the states, it is useful to visualize P as a block matrix of the form

$$P = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right],$$

where A is a $k \times k$ matrix whose entries corresponds to the probabilities to pass from a state (i, i) to (j, j), $i, j \in \{0, \ldots, k-1\}$; B is a $k \times \frac{k^2-k}{2}$ matrix whose entries corresponds to the probabilities to pass from a state (i, i) to (j, l), $i, j, l \in \{0, \ldots, k-1\}$, $j < l$; C is a $\frac{k^2-k}{2} \times k$ matrix whose entries corresponds to the probabilities to pass from a state (j, l) to (i, i), $i, j, l \in \{0, \ldots, k-1\}$, $j < l$; and finally D is a $\frac{k^2-k}{2} \times \frac{k^2-k}{2}$ matrix whose entries corresponds to the probabilities to pass from a state (j, l) to (i, p), $i, j, l, p \in \{0, \ldots, k-1\}$, $j < l$, $i < p$. Recall that m_i is the size of the set S_i, $N = \sum_{i=0}^{k-1} m_i$, and that $N_i = N - m_i$ for all $i \in \{0, \ldots, k-1\}$. Thus, the passing probabilities are

- Let us start form block A. Since at most one of the two CRWs moves in a time step, the only nonzero probabilities are the ones for the system to stay put, that is

$$\mathbb{P}((i, i) \to (j, j)) = \delta_{i,j}\Big(1 - \frac{2}{N}\Big), \quad \forall i, j \in \{0, \ldots, k-1\}, \tag{5.5}$$

  where $\delta_{i,j} = 1$ if $i = j$ and 0 otherwise. It follows that A is a scalar matrix.

- Block B. Here we have the transitions $(i, i) \to (j, l)$. For the same reason as the block A case, the only nonzero probabilities are the ones in which either $i = j$ or

$i = l$. In this case, suppose $i = j$, one of the two random walks in S_i has to move to S_l, this happens if the vertex corresponding to the CRW in S_i is sampled and if the selected neighbour is a vertex in S_l. Thus we get

$$\mathbb{P}((i,i) \to (j,l)) = \delta_{i,j}\frac{2m_l}{NN_i} + \delta_{i,l}\frac{2m_j}{NN_i}, \quad \forall i,j,l \in \{0,\dots,k-1\}, j < l. \tag{5.6}$$

- In block C we are in the opposite situation compared to block B. Here we have the transitions $(j,l) \to (i,i)$. Even here the only nonzero probabilities are the ones in which either $i = j$ or $i = l$. The only precaution we have to take is that we want to rule out the case where the two CRWs meet at the same vertex (otherwise they would coalesce into a single random walk). For these reasons the corresponding probabilities are

$$\mathbb{P}((j,l) \to (i,i)) = \delta_{i,j}\frac{m_i - 1}{NN_l} + \delta_{i,l}\frac{m_i - 1}{NN_j}, \quad \forall i,j,l \in \{0,\dots,k-1\}, j < l. \tag{5.7}$$

- Finally, in block D we deal with the transitions $(j,l) \to (i,p)$. We have two cases: if $(j,l) = (i,p)$, then the probability of such transition equals to the probability for the system to stay put, that is

$$\mathbb{P}((j,l) \to (i,p)) = 1 - \frac{2}{N}, \tag{5.8}$$

for all $i,j,l,p \in \{0,\dots,k-1\}$, $j < l$, $i < p$ such that $(j,l) = (i,p)$. Consider now the case $(j,l) \neq (i,p)$. By the previous observations, the corresponding probabilities are

$$\mathbb{P}((j,l) \to (i,p)) = \delta_{j,i}\frac{m_p}{NN_l} + \delta_{j,p}\frac{m_i}{NN_l} + \delta_{l,i}\frac{m_p}{NN_j} + \delta_{l,p}\frac{m_i}{NN_j}, \tag{5.9}$$

for all $i,j,l,p \in \{0,\dots,k-1\}$, $j < l$, $i < p$ such that $(j,l) \neq (i,p)$.

Since we want to study λ, the spectral radius of P, as $m_0 \to \infty$, we can simplify the expressions above leaving only the probabilities that are $O\big(\frac{1}{m_0}\big)$ as $m_0 \to \infty$ and m_i fixed for every $i \neq 0$. Therefore the remaining nonzero probabilities are:

- For (5.5) in block A we are left with:

$$\mathbb{P}((i,i) \to (j,j)) = \delta_{i,j}\Big(1 - \frac{2}{N}\Big) \underset{m_0\to\infty}{\sim} \delta_{i,j}\Big(1 - \frac{2}{m_0}\Big), \tag{5.10}$$

for all $i,j \in \{0,\dots,k-1\}$.

- For (5.6) in block B we are left with:

- $(0,0) \to (0,j)$, $j \neq 0$,

$$\mathbb{P}((0,0) \to (0,j)) = \frac{2m_j}{NN_0} \underset{m_0\to\infty}{\sim} \frac{2m_j}{m_0 N_0}, \tag{5.11}$$

  for all $j \neq 0$.

- $(i,i) \to (0,j)$, $i,j \neq 0$,

$$\mathbb{P}((i,i) \to (0,j)) = \delta_{i,j} \frac{2m_0}{NN_j} \underset{m_0\to\infty}{\sim} \delta_{i,j} \frac{2}{m_0}, \tag{5.12}$$

  for all $i,j \neq 0$.

• For (5.7) in block C we are in the symmetric case considering block B, therefore we have:

- $(0,j) \to (0,0)$, $j \neq 0$,

$$\mathbb{P}((0,j) \to (0,0)) = \frac{m_0 - 1}{NN_j} \underset{m_0\to\infty}{\sim} \frac{1}{m_0}, \tag{5.13}$$

  for all $j \neq 0$.

- $(0,j) \to (i,i)$, $i,j \neq 0$,

$$\mathbb{P}((0,j) \to (i,i)) = \delta_{j,i} \frac{m_i - 1}{NN_0} \underset{m_0\to\infty}{\sim} \delta_{i,j} \frac{m_i - 1}{m_0 N_0}, \tag{5.14}$$

  for all $i,j \neq 0$.

• Finally, for (5.8) in block D we are in the same case of (5.10), and for (5.9) in the same block we get:

- $(0,i) \to (j,l)$, $i,j \neq 0$, $j < l$,

$$\mathbb{P}((0,i) \to (j,l)) = \delta_{i,j} \frac{m_l}{NN_0} + \delta_{i,l} \frac{m_j}{NN_0} \underset{m_0\to\infty}{\sim} \frac{1}{m_0 N_0} \big(\delta_{i,j}\, m_l + \delta_{i,l}\, m_j\big), \tag{5.15}$$

  for all $i,j \neq 0$, $j < l$.

- $(j,l) \to (0,i)$, $i,j \neq 0$, $j < l$,

$$\mathbb{P}((l,j) \to (0,i)) = \delta_{j,i} \frac{m_0}{NN_l} + \delta_{l,i} \frac{m_0}{NN_j} \underset{m_0\to\infty}{\sim} \frac{1}{m_0} \big(\delta_{j,i} + \delta_{l,i}\big), \tag{5.16}$$

  for all $i,j \neq 0$, $j < l$.

We were not able to explicitly find the spectral radius λ of P, but the relations from (5.10) to (5.16) will be useful to bound an important parameter in the next sections.

Generalization of Lemma 4.8

Assume that $h_i > 0$ for every $i \in \{0, \dots, k-1\}$. We want to study $\lim_{m_0 \to \infty} \mu(h_0, \dots, h_{k-1})$, in particular we claim that the limit goes to 0. Recall that $\mu = \mu_{m_0,\dots,m_{k-1}}$ significantly depends on the size of the sets S_i. Let us consider (5.4), we want to rewrite it only considering terms that are not $o(1)$ as $m_0 \to \infty$. If we look at (5.3), we note that it can be rewritten as

$$\sum_{i=0}^{k-1} \underbrace{\frac{1}{N}\frac{h_i}{N_i}}_{(\star)} \sum_{j \neq i} h_j + \sum_{i=0}^{k-1} \underbrace{\frac{(m_i - h_i)}{N}}_{(\star\star)} \Big(1 - \frac{\sum_{j \neq i} h_j}{N_i}\Big).$$

The terms in $(\star)$ are all $o(1)$ as $m_0 \to \infty$, since $N = \sum_{i=0}^{k-1} m_i$, as well as for the terms in $(\star\star)$ with index $i \neq 0$, while if $i = 0$ we are left with $\frac{m_0 - h_0}{N} \underset{m_0 \to \infty}{\longrightarrow} 1$. It follows that, as $m_0 \to \infty$, (5.3) equals to

$$1 - \frac{\sum_{j \neq 0} h_j}{N_0} + o(1).$$

Reasoning analogously, we get that all terms in (5.1) goes to 0 as $m_0 \to \infty$, while in (5.2) the only coefficient that survives is the one with $i = 0$. This leads to the following rewriting of (5.4)

$$\begin{aligned}\lambda\, \mu(h_0, \dots, h_{k-1}) &= \Big(1 - \frac{\sum_{j \neq 0} h_j}{N_0} + o(1)\Big)\, \mu(h_0, \dots, h_{k-1}) \\ &\quad + \frac{\sum_{j \neq 0} h_j}{N_0}\, \mu(h_0 - 1, h_1, \dots, h_{k-1}) + O\Big(\frac{1}{m_0}\Big), \qquad \text{as } m_0 \to \infty.\end{aligned}$$

Since, from the previous result[1], we know that $\lambda \to 1$ as $m_0 \to \infty$, we get

$$\mu(h_0, \dots, h_{k-1}) = (1 + o(1))\mu(h_0 - 1, h_1, \dots, h_{k-1}) + O\Big(\frac{1}{m_0}\Big), \tag{5.17}$$

which is the same result obtained in equation (4.21), in Chapter 4. Since the last part of the proof of Lemma 4.8 uses only (5.17), we can state that this result is also valid in the k-partite case. Therefore, taking $h_i \geq 0$ for all $i \in \{0, \dots, k-1\}$ and such that $\sum_{j \neq 0} h_j \neq 0$, i.e. such that at least one $h_j > 0$, $j \neq 0$, we have

$$\lim_{m_0 \to \infty} \mu(h_0, \dots, h_{k-1}) = 0. \tag{5.18}$$

[1] We did not fine explicitly the value of λ, but we can state that $\lambda \to 1$ as $m_0 \to \infty$ because the matrix P of the previous section tends to the identity matrix as $m_0 \to \infty$, therefore its spectral radius must tend to 1.

Generalization of Proposition 4.7

Let $(H_0, \ldots, H_{k-1})$ be a random vector representing the number of "yes" in S_i for all $i = 1, \ldots, k-1$, respectively, whose distribution is the unique QSD for the process $\mu = \mu_{m_0,\ldots,m_{k-1}}$. We recall that, since the all zeros k-uple $(0, \ldots, 0)$ and $(m_0, \ldots, m_{k-1})$ belong to Δ (the set of absorbing states), $\mu(0, \ldots, 0) = \mu(m_0, \ldots, m_{k-1}) = 0$ because the support of μ is $\Omega \backslash \Delta$. Therefore we can sum on both sides of (5.4) over $-1 \leq h_i \leq m_i+1$, $i = 0, \ldots, k-1$, while eliminating from the sum the pairs $(h_0, \ldots, h_{k-1}) = (0, \ldots, 0)$ and $(h_0, \ldots, h_{k-1}) = (m_0, \ldots, m_{k-1})$. To this aim, let us first consider separately the general terms of the sums over $i = 0, \ldots, k-1$ on the right-hand side of (5.4). After a re-indexing, setting again $N := \sum_{i=0}^{k-1} m_i$ and $N_i := N - m_i$, from the $h_i + 1$ part we obtain

$$\frac{1}{N}\left[\frac{1}{N_i}\mathbb{E}\Big[H_i(N_i - \sum_{j\neq i} H_j)\Big] - \mu(e_i)\right], \quad \forall i \in \{0, \ldots, k-1\},$$

where e_i represent the null k-uple with value 1 in the i-th position. Similarly, from the $h_i - 1$ part we have

$$\frac{1}{N}\left[\frac{1}{N_i}\mathbb{E}\Big[(m_i - H_i)\sum_{j\neq i} H_j\Big] - \mu(m_0, \ldots, m_i - 1, \ldots, m_{k-1})\right], \quad \forall i \in \{0, \ldots, k-1\};$$

while

$$\frac{1}{N}\frac{1}{N_i}\left[\mathbb{E}\Big[H_i \sum_{j\neq i} H_j\Big] + \mathbb{E}\Big[(m_i - H_i)(N_i - \sum_{j\neq i} H_j)\Big]\right], \quad \forall i \in \{0, \ldots, k-1\},$$

corresponds to the general term of the sum of the probability for the system to stay put. Thus, putting all together in (5.4), it follows that

$$\begin{aligned}
\lambda = \frac{1}{N}\sum_{i=0}^{k-1}\Bigg\{&\frac{1}{N_i}\mathbb{E}\Big[H_i \sum_{j\neq i} H_j\Big] + \frac{1}{N_i}\mathbb{E}\Big[(m_i - H_i)(N_i - \sum_{j\neq i} H_j)\Big] \\
&+ \left[\frac{1}{N_i}\mathbb{E}\Big[H_i(N_i - \sum_{j\neq i} H_j)\Big] - \mu(e_i)\right] \\
&+ \left[\frac{1}{N_i}\mathbb{E}\Big[(m_i - H_i)\sum_{j\neq i} H_j\Big] - \mu(m_0, \ldots, m_i - 1, \ldots, m_{k-1})\right]\Bigg\}.
\end{aligned}$$

After some computations, we notice that all the means cancel each other out, leading us to the following expression

$$\lambda = \frac{1}{N}\sum_{i=0}^{k-1}\Big[m_i - \mu(e_i) - \mu(m_0, \ldots, m_i - 1, \ldots, m_{k-1})\Big].$$

Finally, the invariance under relabeling of the two opinions allows us to write $\mu(m_0, \dots, m_i - 1, \dots, m_{k-1})$ as $\mu(e_i)$. Thus the final result is

$$\lambda = 1 - \frac{2}{N} \sum_{i=0}^{k-1} \mu(e_i). \tag{5.19}$$

Generalization of Proposition 4.9

From (5.19) it follows that $\lim_{m_o \to \infty} \lambda = 1$. Moreover, from Section 5.1, in the generalization of Proposition 4.6, we can say that there exists $\gamma := \gamma(m_1, \dots, m_{k-1})$ such that

$$\lambda = 1 - \frac{\gamma}{N} + o\Big(\frac{1}{m_0}\Big), \quad \text{as } m_0 \to \infty, \tag{5.20}$$

where we recall that $N = \sum_{i=0}^{k-1} m_i$. Set $\mu_\infty(\cdot) := \lim_{m_0 \to \infty} \mu_{m_0, \dots, m_{k-1}}(\cdot) = \lim_{m_0 \to \infty} \mu(\cdot)$, and consider the states $(h_0, \dots, h_{k-1})$ such that $h_0 \geq 1$, $h_1 = \cdots = h_{k-1} = 0$. By (5.18) the latter are the only states in which μ_∞ may not vanish. Therefore we can write the eigenvalue equation (5.4) with the only states $(h_0, 0, \dots, 0)$, $h_0 \geq 1$, obtaining

$$\begin{aligned}\lambda\, \mu(h_0, 0, \dots, 0) &= \mu(h_0, 0, \dots, 0)\, \frac{1}{N} \Big[m_0 - h_0 + \sum_{i \neq 0} \frac{m_i (N_i - h_0)}{N_i} \Big] \\ &\quad + \mu(h_0 + 1, 0, \dots, 0)\, \frac{h_0 + 1}{N} + \frac{1}{N} \Big(1 - \frac{h_0}{N}\Big) \sum_{i \neq 0} \mu(h_0, e_i),\end{aligned}$$

where e_i is the null $k-1$-uple with 1 in position i, and $N_i = N - m_i$. After rearranging the terms and substituting $\lambda = 1 - \frac{\gamma}{N}$ we get

$$\begin{aligned}\mu(h_0 + 1, 0, \dots, 0) &= \mu(h_0, 0, \dots, 0)\, \frac{N}{h_0 + 1} \Big[1 - \frac{\gamma}{N} - \frac{1}{N} \Big(m_0 - h_0 + N_0 - \sum_{i \neq 0} \frac{m_i h_0}{N_i} \Big) \Big] \\ &\quad - \frac{1}{h_0 + 1} \Big(1 - \frac{h_0}{N}\Big) \sum_{i \neq 0} \mu(h_0, e_i).\end{aligned}$$

Observe that we can neglect the last term on the right-hand side since, as $m_0 \to \infty$, we have that $\mu_\infty(h_0, e_i) = 0$. Similarly, $\sum_{i \neq 0} \frac{m_i h_0}{N_i} \to 0$ as $m_0 \to \infty$. Thus, taking the limit of $m_0 \to \infty$, we are left with

$$\mu_\infty(h_0 + 1, 0, \dots, 0) = \mu_\infty(h_0, 0, \dots, 0)\, \frac{h_0 - \gamma}{h_0 + 1}, \tag{5.21}$$

for every $h_0 \geq 1$. Note that if we write λ as $1 - \frac{\gamma}{N}$, and in (5.19) let $m_0 \to \infty$, we get

$$\mu_\infty(1, 0, \dots, 0) = \frac{\gamma}{2}. \tag{5.22}$$

It follows that we can solve the recursive equation (5.21) with the initial condition (5.22). The final result is

$$\mu_\infty(h_0, 0, \dots, 0) = \frac{\gamma}{2} \frac{1}{h_0!} \prod_{j=1}^{h_0-1} (j-\gamma), \quad \forall h_0 \geq 1. \tag{5.23}$$

It is immediate that the latter equivalence implies

$$\mu_\infty(h_0, 0, \dots, 0) = \frac{1}{2} f_\gamma(h_0), \quad \forall h_0 \geq 1, \tag{5.24}$$

where f_γ is the probability mass function of the Sibuya distribution of parameter $\gamma \geq 0$. The last claim that we need to prove is that $\gamma \leq 1$. Then we could conclude that, as $m_0 \to \infty$, the distribution of dissenters under the QSD μ tends to the Sibuya distribution of parameter $\gamma \in (0,1)$.
To this aim we recall that the first order limit behaviour of the transition matrix P, as $m_0 \to \infty$, is given by the relations (5.10)-(5.16). Since P is a non-negative matrix, it holds that its spectral radius λ is bounded as follows

$$\min_i \sum_j p_{i,j} \leq \lambda \leq \max_i \sum_j p_{i,j},$$

where $(p_{i,j})$, $i, j \in \{1, \dots, \frac{k^k+k}{2}\}$, are the elements of the matrix P. The upper bound is clearly 1, due to the fact that P is sub-stochastic, while we are interested in the lower bound. Using (5.10)-(5.16), we find that the sum of every row of P gives either 1 or $1 - \frac{1}{m_0 N_0}$. Therefore we have that

$$\lambda \geq 1 - \frac{1}{m_0 N_0}.$$

As stated at the beginning of this section, we can consider (5.20) as the approximation of λ as $m_0 \to \infty$. With such rewriting, the latter inequality becomes

$$\gamma < \frac{N}{m_0 N_0} \underset{m_0 \to \infty}{\sim} \frac{1}{N_0} < 1.$$

Since we are interested in the asymptotic behaviour of λ, we can conclude that under our hypothesis $\gamma \in (0,1)$.
Eventually, from (5.24), it follows a result similar to Theorem 4.10 for complete bipartite graphs.

Theorem 5.1. *Let $C \sim \mathrm{Bern}(1/2)$ and $D \sim \mathrm{Sib}(\gamma)$ be independent, where $\gamma \in (0,1)$ depends on $m_1, \dots, m_{k-1}$ and with $\mathrm{Bern}(p)$ we denote a Bernoulli random variable which takes values 0 and 1 with probability $1-p$ and p, respectively. Then the distribution of opinions under the QSD for the voter model on the complete k-partite graph K, with $k \geq 2$, as $m_0 \to \infty$ converges weakly to the following:*

(i) All vertices of every S_i, for all $i \neq 0$, have opinion C.

(ii) All but D vertices in S_0 have opinion C.

In particular, in the k-partite case there is a dominant opinion, which can be 0 or 1 chosen with Bernoulli distribution of parameter 1/2, that brings to consensus k-1 sets over k, and all the dissenters are concentrated in the biggest set (the one whose size goes to infinity) with density given by the Sibuya distribution of parameter $\gamma \in (0, 1)$. We can use the invariance under relabeling of the two opinions to rewrite this result as follows

$$\lim_{m_0 \to \infty} \mu_{m_0,\dots,m_{k-1}}(m_0 - h_0, m_1, \dots, m_{k-1}) = \lim_{m_0 \to \infty} \mu_{m_0,\dots,m_{k-1}}(h_0, 0, \dots, 0) = \frac{1}{2} f_\gamma(h_0),$$

for all $h_0 \geq 1$.

5.2 Bipartite case with both $n, m \to \infty$

In the following section we introduce our second attempt of generalization regarding the work of Ben-Ari et al. in [2]. This time, instead of taking k sets and studying the asymptotic behaviour of the QSD as the size of one set goes to infinity, we return to bipartite graphs $K_{n,m}$, and we consider the case in which both the sizes n and m tend to infinity. To this aim, let us rewrite $m = g(n)$, where $g : \mathbb{N} \to \mathbb{N}$ is a regular function such that

$$\lim_{n \to +\infty} g(n) = +\infty.$$

Clearly, all the results that do not involve the limit of $n \to \infty$ in Chapter 4 still hold in our settings. In particular, the eigenvalue equation for the unique QSD $\mu := \mu_{n,g(n)}$, with eigenvalue $\lambda := \lambda_{n,g(n)}$, for the process is still

$$\begin{aligned}\lambda\, \mu(k,h) = \mu(k,h) &\frac{kh + (n-k)(g(n)-h)}{ng(n)} \\ &+ \mu(k-1,h)\frac{(n-k+1)h}{(n+g(n))g(n)} + \mu(k+1,h)\frac{(k+1)(g(n)-h)}{(n+g(n))g(n)} \\ &+ \mu(k,h-1)\frac{(g(n)-h+1)k}{(n+g(n))n} + \mu(k,h+1)\frac{(h+1)(n-k)}{(n+g(n))n}.\end{aligned} \tag{5.25}$$

Moreover, if we consider the transition matrix of the (two) coalescing random walk processes used to derive λ, we can notice that each of its entries do not depend on the growth of n and $g(n)$ as $n \to \infty$. This implies that the corresponding spectral radius, i.e.

λ, will be the same whether $g(n)$ increases, for example, exponentially or logarithmically with n. In fact its value will be

$$\lambda = 1 - \frac{2}{n+g(n)}\left(1 - \sqrt{1 - \frac{1}{2n} - \frac{1}{2g(n)}}\right) = 1 - \frac{\gamma}{n+g(n)}, \tag{5.26}$$

where $\gamma := \gamma_{n,g(n)} = 2\left(1 - \sqrt{1 - \frac{1}{2n} - \frac{1}{2g(n)}}\right)$. It follows that $\gamma \to 0$ and $\lambda \to 1$ as $n \to \infty$, for any function $g(\cdot)$.
Also the result that links the eigenvalue λ to the values of QSD μ in $(0,1)$ and $(1,0)$ still holds

$$\lambda = 1 - \frac{2}{n+g(n)}(\mu(1,0) + \mu(0,1)).$$

Let us now define $\mu_\infty(k,h) := \lim_{n\to\infty} \mu(k,h) = \lim_{n\to\infty} \mu_{n,g(n)}(k,h)$ as the limit in distribution of the QSD, for any state $(k,h) \in \{0,\dots,n\} \times \{0,\dots,g(n)\}$. If we rewrite in the latter relation λ as in (5.26), we get

$$\mu(1,0) + \mu(0,1) = \frac{\gamma}{2} \xrightarrow[n\to\infty]{} 0,$$

that is

$$\mu_\infty(1,0) = \mu_\infty(0,1) = 0. \tag{5.27}$$

This equation shows a first relevant difference between the case in which only $n \to \infty$ and the one in which both $n, m \to \infty$. Another significant difference regards the limit distribution itself. In their article, Ben-Ari et al. showed that for any $k \geq 0$ and $h > 0$, $\lim_{n\to\infty} \mu_{n,m}(k,h) = 0$; we now see that in our case such statement can not hold in general. If we take the eigenvalue equation (5.25) and rewrite λ as in (5.26) we obtain

$$\begin{aligned}\left(1 - \frac{\gamma}{n+g(n)}\right)\mu(k,h) = \mu(k,h)&\frac{kh + (n-k)(g(n)-h)}{ng(n)} \\ &+ \mu(k-1,h)\frac{(n-k+1)h}{(n+g(n))g(n)} + \mu(k+1,h)\frac{(k+1)(g(n)-h)}{(n+g(n))g(n)} \\ &+ \mu(k,h-1)\frac{(g(n)-h+1)k}{(n+g(n))n} + \mu(k,h+1)\frac{(h+1)(n-k)}{(n+g(n))n},\end{aligned}$$

for all $k, h \geq 0$. We can get rid of every term that is $o(1)$ as $n \to \infty$. Thus, after a rearrangement we are left with

$$\begin{aligned}-\gamma\,\mu(k,h) = -\mu(k,h)\left[h\left(1 + \frac{n}{g(n)}\right) + k\left(1 + \frac{g(n)}{n}\right)\right] + \mu(k-1,h)\,h\frac{n}{g(n)} \\ + \mu(k+1,h)\,(k+1) + \mu(k,h-1)\,k\frac{g(n)}{n} + \mu(k,h+1)\,(h+1).\end{aligned} \tag{5.28}$$

Therefore there are two relevant cases to treat, depending on the fact that $\frac{n}{g(n)}$ tends to ∞ (equivalently 0) or a positive constant as $n \to \infty$. Let us consider first the latter case. Suppose that $\frac{n}{g(n)} \to c > 0$ as $n \to \infty$. Recalling that $\gamma \to 0$ as $n \to \infty$, it follows that if we let $n \to \infty$, (5.28) becomes

$$\begin{aligned} \mu_\infty(k,h)\left(hc + h + k + \frac{k}{c}\right) &= \mu_\infty(k-1,h)\,hc + \mu_\infty(k+1,h)\,(k+1) \\ &\quad + \mu_\infty(k,h-1)\,\frac{k}{c} + \mu_\infty(k,h+1)\,(h+1), \end{aligned} \tag{5.29}$$

for all $k, h \geq 0$. Let us take $k = 0$ in the previous equation, with $h \geq 0$. We get the following recursive formula in $h \geq 0$

$$\mu_\infty(0,h)\,\big(hc + h\big) = \mu_\infty(1,h) + \mu_\infty(0,h+1)\,(h+1).$$

Using the initial conditions (5.27) and the fact that μ_∞ is not supported on the states $(0,0)$ and $(0,-1)$, it follows that

$$\mu_\infty(1,h) = \mu_\infty(0,h) = 0 \quad \forall h \geq 0. \tag{5.30}$$

Reasoning similarly, if we consider (5.29) with $h = 0$ and $k \geq 0$, we obtain another recursive relation that can be solved still using (5.27) and the fact that μ_∞ is not supported on $(0,0)$ and $(-1,0)$. The result is

$$\mu_\infty(k,1) = \mu_\infty(k,0) = 0 \quad \forall k \geq 0. \tag{5.31}$$

At this point, let us go back to (5.29) and take $k = 1$, $h \geq 0$ (or $h = 1$ and $k \geq 0$). By (5.30) (respectively (5.31)), it follows that $\mu_\infty(2,h) = 0$ for all $h \geq 0$ ($\mu_\infty(k,2) = 0$ for all $k \geq 0$). Reasoning inductively we conclude that

$$\mu_\infty(k,h) = 0 \quad \forall k, h \geq 0. \tag{5.32}$$

Thus, we conclude that in the case in which $\frac{n}{g(n)} \to c > 0$ as $n \to \infty$, the limit behaviour of the QSD for the process does not converge to a probability distribution, in the same way as we have seen in the case of the complete graph K_n, in Example 1.
Suppose now that $\frac{n}{g(n)} \to \infty$ as $n \to \infty$, so that $\frac{g(n)}{n} \to 0$ as $n \to \infty$. We can rewrite (5.28) as follows

$$\begin{aligned} h\,\frac{n}{g(n)}\big[\mu(k-1,h) - \mu(k,h)\big] &= \mu(k,h)\left[k + h + k\,\frac{g(n)}{n} - \gamma\right] - \mu(k+1,h)\,(k+1) \\ &\quad - \mu(k,h-1)\,k\,\frac{g(n)}{n} - \mu(k,h+1)\,(h+1), \end{aligned} \tag{5.33}$$

for all $h, k \geq 0$. If we consider $h = 0$ the left-hand side vanishes, thus taking the limit $n \to \infty$ we obtain

$$\mu_\infty(k, 0)\, k = \mu_\infty(k+1, 0)\,(k+1) + \mu_\infty(k, 1), \quad \forall k \geq 0.$$

Similarly to the previous case, using the initial conditions (5.27), we have

$$\mu_\infty(k, 0) = \mu_\infty(k, 1) = 0, \quad \forall k \geq 0.$$

Let us now consider $h \neq 0$ and divide, from both sides, the equation (5.33) by $\frac{n}{g(n)}$. Leaving only the terms that are not $o(1)$ as $n \to \infty$, we get

$$h[\mu(k-1, h) - \mu(k, h)] = o(1), \quad \text{as } n \to \infty. \tag{5.34}$$

Thus, necessarily

$$\mu_\infty(k, h) = \mu_\infty(k-1, h), \quad \forall k, h \geq 0.$$

Using the fact that μ_∞ is not supported on the states $(-1, h)$, for every $h \geq 0$, it follows that

$$\mu_\infty(k, h) = 0, \tag{5.35}$$

for all $k, h \geq 0$. Leading to the same result of the case $\frac{n}{g(n)} \to c > 0$ as $n \to \infty$, in (5.32). We finally conclude that in both cases $\frac{n}{g(n)} \to c > 0$ and $\frac{n}{g(n)} \to \infty$, as $n \to \infty$, there exists no QSD for the asymptotic system, exactly as it happens in the complete graph K_n; the continuous-time analogue of such case was treated in [10].
A possible interpretation to the fact that the limit distribution vanishes on all states is that the mass of the system, through the limit, will be concentrated out of the support of μ_∞. This implies that there must be probability mass only on the states $(0, 0)$ and (∞, ∞), as $n \to \infty$.

Bibliography

[1] D. Aldous and J. A. Fill. *Reversible Markov Chains and Random Walks on Graphs.*, Unfinished monograph, 1994-2002.

[2] I. Ben-Ari, H. Panzo, P. Speegle and R. O. VandenBerg. *Quasi-Stationary Distributions for the Voter Model on Complete Bipartite Graphs*, Lat. Am. J. Probab. Math. Stat. 18, pages 421–437, 2021.

[3] P. Brémaud. *Markov chains: Gibbs fields, Monte Carlo simulation and queues*, Springer, 1999.

[4] P. Collet, S. Martínez and J. San Martín. *Quasi-Stationary Distributions. Markov Chains, Diffusions and Dynamical Systems*, Springer, 2013.

[5] J. T. Cox. *Coalescing random walks and voter model consensus times on the tourus* $\mathbb{Z}^d$. Syracuse University, The annals of probability, Vol. 17, No. 4, pages 1333-1366, 1989.

[6] J. N. Darroch and E. Seneta.*On quasi-stationary distributions in absorbing discrete-time finite Markov chains.* J. Appl. Probability, Volume 2, pages 88–100, 1965.

[7] J. N. Darroch and E. Seneta.*On quasi-stationary distributions in absorbing continuous-time finite Markov chains.* J. Appl. Probability, Volume 4, pages 192–196, 1967.

[8] D. A. Dawson. *Introductory Lectures on Stochastic Population Systems*, School of Mathematics and Statistics, Carleton University, Ottawa, Canada, 2017.

[9] L. Devroye. *A triptych of discrete distributions related to the stable law.* Statist. Probab. Lett., 18 (5), pages 349–351, 1993.

[10] R. Dickman, and R. Vidigal. *Quasi-stationary distributions for stochastic processes with an absorbing state.* J. Phys. A, 35 (5), pages 1147–1166, 2002.

[11] S. N. Evans. and D. Steinsaltz. *Markov mortality models: Implications of quasistationarity and varying initial conditions.* Theo. Pop. Bio., Volume 65, pages 319–337, 2004.

[12] J. Fernley and M. Ortgiese. *Voter models on subcritical inhomogeneous random graphs*, arXiv:1911.13187 [math.PR], 2020.

[13] P. A. Ferrari, H. Kesten, S. Martínez, and P. Picco. *Existence of quasi-stationary distributions. A renewal dynamical approach.* Ann. Probab., Volume 23, Issue 2, 1995.

[14] F. R. Gantmacher. *The theory of matrices.* Vols. 1, 2. Translated by K. A. Hirsch. Chelsea Publishing Co., New York, 1959.

[15] R. A. Holley and T. M. Liggett. *Ergodic theorems for weakly interacting infinite systems and the voter model*, The Annals of Probability, Vol. 3, No. 4, pages 643-663, 1975.

[16] T. M. Liggett. *Interacting Particle Systems*, Springer-Verlag, 1985.

[17] S. Méléard and D. Villemonais. *Quasi-stationary distributions and population processes*, arXiv:1112.4732 [math.PR], 2011.

[18] R. I. Oliveira. *On the coalescence time of reversible random walks.*, Transaction of the AMS, Volume 364, Number 4, pages 2109–2128, 2012.

[19] S. I. Resnick. *Adventures in Stochastic Processes*, Birkhauser, 1992.

[20] E. A. van Doorn. *Quasi-stationary distributions and convergence to quasi-stationarity of birth-death processes.* Adv. in Appl. Probab., Volume 23, Issue 4, pages 683–700, 1991.