Diffusion limit of the Poisson encounter-mating model

by

Murad Ramanovski

A Dissertation Submitted to the
Graduate School of Sciences and Engineering
in Partial Fulfillment of the Requirements for
the Degree of

Master of Science

in

Mathematics

KOÇ UNIVERSITY

March 2019

Diffusion limit of the Poisson encounter-mating model

Koç University

Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a master's thesis by

Murad Ramanovski

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Committee Members:

Prof. Mine Çağlar (Advisor)

Assoc. Prof. Atilla Yılmaz

Asst. Prof. Ümit Işlak

Date:

To my family and friends

ABSTRACT

Encounter-mating models are used to represent mating scenarios in finite zoological populations consisting of males and females, partitioned into different types. Using them, one is able to draw conclusions about the mating process, and ultimately, the resulting pattern of permanent male-female pairs formed between different types, called the mating pattern. Recently, Gün and Yılmaz (*Acta Applicandae Mathematicae* 148, 1 (2017), 71-102) developed the so-called stochastic encounter-mating (SEM) model to generalize and unify encounter-mating models previously presented in the literature. One particular scenario in the SEM is when the mating clocks of the individuals, within the population being modeled, are Poisson. In this scenario, the process that keeps record of the pair-type formation becomes a continuous-time Markov chain. Gün and Yılmaz show that, despite some inconsistencies (i.e. up to error terms), this process can be incorporated into a framework of density-dependent Markov processes, and derive two separate laws of large numbers, one for the mating process, and another for the mating pattern (*Advances in Applied Probability* 49, 4 (2017) 1201-1229). In this thesis, we discuss the central limit theorem (CLT) counterparts of these two results. We present two proofs of the CLT for the pair-type process, and provide evidence gathered from MATLAB simulations, to argue the existence of a CLT for the mating pattern.

ÖZETÇE

Karşılaşma-eşleşme modelleri, çok sayıda erkek ve dişi türü olan sonlu zoolojik popülasyonlardaki eşleşme senaryolarını temsil etmek için kullanılır. Bunlar kullanılarak, çiftleşme süreci ve en nihayetinde eşleşme düzeni diye adlandırılan farklı türler arasında kalıcı erkek-dişi çift oluşumunun son düzenine ilişkin sonuçlar çıkarılabilir. Yakın geçmişte, Gün ve Yılmaz (*Acta Applicandae Mathematicae* 148, 1 (2017), 71-102), literatürde önceden sunulan karşılaşma-eşleşme modellerini genelleyen ve birleştiren stokastik karşılaşma-eşleşme (SEM) modelini geliştirmişlerdir. Modellenmekte olan popülasyondaki bireylerin eşleşme saatlerinin Poisson dağılımlarına sahip olduğu durum SEM modelinin kapsadığı senaryolardan önemli bir tanesidir. Bu durum, çift-türü oluşumunu izleyen süreci sürekli-zamanlı Markov zinciri haline getirir. Daha sonra, Gün ve Yılmaz (*Advances in Applied Probability* 49, 4 (2017), 1201-1229), modelin bazı tutarsızlıklara rağmen, sürecin yoğunluğa bağlı Markov süreçleri teorisi kapsamında incelenebileceğini göstermişlerdir. Ayrıca, bahsedilen çift-tür süreci ve nihai eşleşme düzeni için iki ayrı büyük sayılar kanunu elde etmişlerdir. Bu tezde, bahsi geçen bulguların merkezi limit teorem karşılıklarını ele alacağız. Stokastik çift-türü sürecine dair merkezi limit teoremin iki farklı ispatını sunacağız. Eşleşme biçim düzenine ait olan merkez limit teoremin doğruluğu ise yaptığımız MATLAB simülasyonlarıyla desteklenecektir.

ACKNOWLEDGMENTS

I am extremely grateful to my advisor Prof. Mine Çağlar for her helpful advice, and support during my studies at Koç University. I would also like to express my deep and sincere gratitude to my former advisor, Assoc. Prof. Atilla Yılmaz, for his insight, patient approach, invaluable discussions, and most of all guidance, for it has influenced how I now view mathematics. Last, but not least, I want to thank Asst. Prof. Ümit Işlak for selflessly investing his time and knowledge in helping not only me, but a number of other students nurture their passion for probability. His enthusiasm at Boğaziçi University was truly an inspiration.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

1.1 Motivation, background and early works

In population genetics, when observing how traits (or more specifically their genotypes) are passed from generation to generation in species that reproduce sexually, it is important to determine how those traits are being distributed among mating pairs. One possible way that this can be done is by keeping record of (and analyzing) the so-called mating pattern (i.e. who has mated with whom) of the population. An alternative approach would be, instead of looking at the mating pattern, which is just the end-result of an ongoing mating process, to examine the mating process itself, and the underlying preferences of the mating individuals as being a part of a dynamical mechanism. It is quite the misfortune that, from a researchers' perspective, and mostly due to practical reasons, the first approach seems far more fetching, and therefore is prone to be mistakenly more preferred [5].

As mentioned in [7], one of the reasons for such an issue is that since the mating pattern can be observed in the natural habitat of the population in question, it is relatively less costly to try and infer the mating preferences of individuals by observing the pattern, as opposed to trying to measure the preferences directly (which could require time-consuming controlled experiments). There was, however, another and particularly important reason for such mistakes to be made. This was the distressing lack of a clear distinction between mating patterns and mating preferences in the literature until the 90's.

Due to this matter, terms such as "assortative mating", and "random mating" were used carelessly to describe either the pattern or the preferences, when clearly the former is a population property, whereas the latter is a property of the mating process. In other words, one may run into scenarios where although the individuals in the mating mechanism have assortative preferences (i.e. they prefer one type over another), the end result still gives rise to a mating pattern with zero correlations between different types, thus falsely implying random mating. This matter was discussed by Burley in [3] where she addressed the ambiguity of the term assortative mating, and in turn used other definitions such as "homogamy" and "heterogamy", which had been already adopted to refer to positive and negative correlations between similar phenotypes in a population pattern. More importantly, she demonstrated that several different mating preferences may result in the same pattern thus giving evidence that the two are not interchangeable.

1.2 Encounter-mating models

Additional works in the literature that made distinctions similar to Burley's include [14], [12], and [16], although in these, mating preferences are not defined precisely enough. This gives rise to difficulties whenever one attempts to utilize said works in mating models. This predicament is pointed out in [5], where Gimelfarb introduces two encounter-mating (EM) models, to investigate how different mating mechanisms influence the resulting mating pattern of a population. In both models, the population is comprised of n males and n females, each of which belongs to one of $k \geq 2$ different types, and the mating preferences are defined as probabilities that given a courtship between two types occurred, it will result in mating. The mating procedure can be visualized with a scheme of two urns (one for each sex), each containing n balls.

The first model, which is named individual EM, is constructed by picking one ball from each urn uniformly at random, and initiating an encounter, i.e. a temporary pair between the two. Then, depending on the types of the balls, this temporary bond becomes permanent with a certain probability (say p_{ij} if we have an i-type female and a j-type male, where i, j are some natural numbers between 1 and k), in which case the newly formed pair is stored in a third urn. If the bond fails to become permanent (i.e., the courtship does not result in mating), both the male and female ball are returned to their respective urns and are available to be picked again from the pool of singles. The second model, called mass EM, is only different from the first one in the encounter step, where instead of having one encounter at a time, we assign to each single male a single female picked uniformly at random. Again each of these temporary pairs that successfully become permanent are put into the third urn, and the rest are returned to their respective urns for another round of mass encounter. In both of the models, the encounter-mating steps are repeated until the pool of singles is depleted, and everyone has successfully mated.

In his paper, [5] Gimelfarb observed that for the same set of preferences (i.e. the probabilities p_{ij} that a temporary ij-bond will become permanent), the individual and the mass EM yield different mating patterns, thus concluding that the pattern is not exclusively dependent on the preferences. Moreover he addressed the converse problem, i.e. whether plugging different preferences into either model will make them yield similar mating patterns. This turns out to be an important issue and in Chapter 3 we review it in more detail.

Gimelfarb is not the first who worked on these kind of models (see [15], where Romney utilized the individual EM to model marriages in community, and [13], who analysed Romney's model and was able to derive equations for the expected mating pattern for small populations). However his work is of particular interest for this thesis as it motivates the stochastic encounter-mating (SEM) model, introduced in [7] as a generalization of both the individual and the mass EM. The SEM draws its strength from employing the concept of random firing times at which temporary pairs are formed just like in [5]. More precisely, to each male/female a random counting process denoting the individuals' firing times is attached. The distributions of these processes can be chosen appropriately to obtain a variety of different mating scenarios, two special cases of which are the individual (when the processes are Poisson) and mass (when the processes are Bernoulli) EM models that Gimelfarb introduces in [5].

1.3 Overview of the rest of the thesis

As is shown in Gün and Yılmaz's paper [7] and will be explained in this text, whenever the processes that give the firing times are Poisson, a particular pair-type process (which we will define later on) assigned to the SEM model becomes a continuous-time Markov chain. Moreover this process can be fit into a theoretical framework for density-dependent Markov processes developed by Kurtz (see [4, 10, Chapter 11], [9–11]). We present this theory in Chapter 2, by firstly defining what a family of density-dependent processes is, and providing a couple of examples that should endow the reader with some intuition. We then state a law of large numbers and a central limit theorem for these processes. Along the way we establish notation that will be used throughout the text.

In Chapter 3, we give details of the Poisson SEM model (we will write PEM shortly). In particular we explain how it is constructed, how the pair-type process assigned to it (approximately) works within the theory given in Chapter 2, and how therefore, it satisfies a law of large numbers and has a deterministic limit. This functional LLN for the mating process is then extended to a LLN for the mating pattern of the population. We mention how some of Gimelfarb's non-rigorous results were proven and generalized rigorously in [6], and how two of his conjectures regarding the relationship between preferences and patterns were proven and strengthened by Gün and Yılmaz.

Chapters 4 and 5 contain the main results of this thesis. We present a functional central limit theorem to complement the LLN for the pair-type process, and an additional CLT to complement the LLN for the mating pattern of the population. The former (i.e. the FCLT) is proved in Chapter 4, in two different ways. The first proof is shorter and more elegant, while the second one is inspired by [4] and [1, Chapter 4], and should be found enjoyable by any reader who studied through the text of Ethier and Kurtz, as it heavily relies on techniques presented there. Chapter 5 is dedicated to the CLT for the mating pattern. Although we do not present an analytic proof for it, we provide results of numerous MATLAB simulations to provide positive evidence.

Chapter 2

DENSITY-DEPENDENT PURE-JUMP MARKOV PROCESSES

In population dynamics, density dependence refers to when the rates of growth of some variable of interest in a given population process, are dependent on the density of that particular variable. Although the theory of density-dependent population processes was motivated within the field of population dynamics, the definition of a population can take a much broader spectrum of meanings. In its most general context, population means a set of similar particles. These can be molecules (if the stochastic process models a chemical reaction), or individuals in an epidemic, birth-death, or encounter-mating model.

The purpose of this chapter is to introduce the theory of density-dependent pure-jump Markov processes, which are used in models such as the ones we mentioned above. The main approximation and limiting results for these processes were obtained by Kurtz [9–11], and a significant exposition on these results is given in [4]. We therefore follow the latter text closely, however, with minor notation alterations to fit our purposes.

2.1 *Definition and examples*

Let $\mathcal{E} \subset \mathbb{R}^d$, and let $\mathbb{Z}^d$ be the d-dimensional integer lattice. We attach to each nonzero l in $\mathbb{Z}^d$, a nonnegative Borel-measurable function $F_l : \mathcal{E} \to \mathbb{R}$ such that $\sum_l F_l(y) < \infty$, $\forall y \in \mathcal{E}$. Having this setup, we are ready for the following definition.

Definition 2.1. We call $\{n^{-1}Q^{(n)} : n = 1, 2, \dots\}$ a *density-dependent family* of processes corresponding to the collection $\{F_l : l \in \mathbb{Z}^d, l \neq 0\}$ if they are defined on a common probabiity space $(\Omega, \mathcal{F}, \mathbb{P})$, and for each $n \in \mathbb{N}$, $\left(Q^{(n)}(t)\right)_{t\geq 0}$ is a continuous-time Markov chain with

- $Q^{(n)}(0) = q_n \in \mathbb{R}^d$ being non-random,
- state space $\mathcal{E}_n^* := \mathbb{Z}^d \cap \{ny : y \in \mathcal{E}\}$
- l-jump probabilities constructed via each F_l as follows:

$$\mathbb{P}_x(Q^{(n)}(h) = y) = \begin{cases} nF_l(n^{-1}x)h + o(h), \text{ if } y = x + l; \\ 1 - \sum_l nF_l(n^{-1}x)h + o(h), \text{ if } y = x; \end{cases}$$

where for any $\Gamma \subset \mathcal{E}_n^*$, $\mathbb{P}_x(Q^{(n)}(h) \in \Gamma) := \mathbb{P}(Q^{(n)}(t+h) \in \Gamma \mid Q^{(n)}(t) = x)$ denotes the usual conditional probability.

It is worth noting several details regarding the above definition.

- Since $Q^{(n)}(t)$ lies in the state space $\mathcal{E}_n^*$, the rescaled quantity $n^{-1}Q^{(n)}(t)$ takes values in $\mathcal{E}_n := \mathcal{E} \cap \{n^{-1}l : l \in \mathbb{Z}^d\}$.
- It was implicitly required that if $n^{-1}x \in \mathcal{E}_n$ and $F_l(n^{-1}x) > 0$, then we have $n^{-1}(x+l) \in \mathcal{E}_n$, too. For example, if $\mathcal{E}$ were to be bounded, we would require $\{l : \sup_y F_l(y) > 0\}$ to be finite.
- It is immediately implied by the definition that the probability of each succeeding jump is dependent on the current state x only via its ratio with the parameter n, thus giving the process its "density-dependent" nature.

Say $\mathcal{E}$ and $\{F_l : l \in \mathbb{Z}^d, l \neq 0\}$ are such that for each $n \in \mathbb{N}$, there exists a separate probability space on which a process $\left(\widehat{Q}^{(n)}(t)\right)_{t\geq 0}$ satisfies the distributional conditions of Definition 2.1. Then finding a common space $(\Omega, \mathcal{F}, \mathbb{P})$ on which we can define a family of processes $\{Q^{(n)}\}_{n\in\mathbb{N}}$ (where $Q^{(n)}$ and $\widehat{Q}^{(n)}$ have the same finite-dimensional distributions) is no obstacle at all. To see this, let $x \in \mathcal{E}_n^*$, $\Gamma \subset \mathcal{E}_n^*$, and $\Gamma_x := \{l \in \mathbb{Z}^d : x + l \in \Gamma\}$. Then define the transition function $\mu_n(x, \Gamma)$, and the waiting time parameter function $\lambda_n(x)$ to be

$$\mu_n(x, \Gamma) := \frac{\sum_{l\in\Gamma_x} F_l(n^{-1}x)}{\sum_{l\in\mathbb{Z}^d} F_l(n^{-1}x)}, \quad \text{and} \quad \lambda_n(x) := n \sum_{l\in\mathbb{Z}^d} F_l(n^{-1}x).$$

Next, let some discrete-time Markov chain $\{X_n(s) : s \in \mathbb{N} \cup \{0\}\}$ satisfy $X_n(0) = q_n$, and have the jump probabilities $\mathbb{P}(X_n(s+1) \in \Gamma \mid X_n(s) = x) = \mu_n(x, \Gamma)$. As shown in [4, Chapter 4.2],

$$\begin{aligned} A_n f(x) &:= \lambda_n(x) \int \big(f(y) - f(x)\big)\mu_n(x, dy) \\ &= \sum_l nF_l(n^{-1}x)\big(f(x+l) - f(x)\big) \end{aligned}$$

defines a bounded linear operator on the Banach space of Borel functions $f : \mathcal{E}_n^* \to \mathbb{R}$ such that $\sup_x |f(x)| < \infty$. Moreover, any continuous Markov chain that satisfies the distributional conditions in Definition 2.1 has A_n as its infinitesimal generator, and X_n as its embedded discrete-time chain.

Now, let $\{J_l(t) : l \in \mathbb{Z}^d, l \neq 0\}$ be a collection of independent standard Poisson processes defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and $\mathbb{Z}^d \cup \{\infty\}$ be the one-point compactification of the integer lattice. Set $Q^{(n)}(t)$ to be given by the equation

$$Q^{(n)}(t) = \begin{cases} q_n + \sum_l l J_l\Big(n \int_0^t F_l(n^{-1}Q^{(n)}(s))ds\Big) \text{ for } \ t < \tau_\infty, \\ \infty \quad \text{for } \ t \geq \tau_\infty, \end{cases} \tag{2.1}$$

where $\tau_\infty := \inf\{t : Q^{(n)}(t-) = \infty\}$ is the first time the process makes infinitely many jumps. Then by [4, Theorem 6.4.1], $Q^{(n)}$ has infinitesimal generator A_n, and satisfies Definition 2.1 before the first infinity of jumps. Furthermore, since for all n, $Q^{(n)}$ is constructed via the same Poisson processes, the entire sequence $\{n^{-1}Q^{(n)}\}_{n\in\mathbb{N}}$ resides on $(\Omega, \mathcal{F}, \mathbb{P})$, which makes it a density-dependent family.

This Poisson construction of $\{Q^{(n)}\}_{n\in\mathbb{N}}$ proves to be of much importance later on. Intuitively it makes sense, since for each l, conditioned on $Q^{(n)}(t) = x$, we expect the probability that $J_l\Big(n \int_0^\cdot F_l(n^{-1}Q^{(n)}(s))ds\Big)$ has a jump within $[t, t + \Delta t)$ to be equal to $nF_l(n^{-1}x)\Delta t + o(\Delta t)$. Conveniently enough, this is exactly what we need as it corresponds to the third condition in the definition of $Q^{(n)}$.

We end this section by introducing two standard examples of density-dependent processes, which will hopefully provide the reader with some intuition.

2.1.1 Birth-and-death models

Let n be the area of a city, in which at time t, a population of $k^{(n)}(t)$ people resides. Here, one matter of interest is the logistic growth behavior of the population in the long term. In real-world scenarios, we generally expect both the birth rate $p^{(n)}_{k,k+1}(t)$ and death rate $p^{(n)}_{k,k-1}(t)$ in the city to depend on the population density $k^{(n)}(t)/n$ via some functions λ and μ, respectively. Another reasonable assumption is that the rates approximately increase (or decrease) proportionally to $k^{(n)}(t)$.

As an example, one could suppose that given some $b > 0$, there are initially $k(0) > b$ individuals and that to each (current and newborn) one an independent Poisson process of rate a is attached to denote their times of giving birth. Similarly, let each individual die with rate b, and emigrate with rate proportional to the population density, say $cn^{-1}k^{(n)}(t)$ for some $c > 0$, i.e. the more crowded the city is, the more individuals are prone to leave to another city. Then, given the two assumptions we have made above, the birth and death/emigration rates are governed by $\lambda(x) = a$ and $\mu(x) = b + cx$ via the following two equations:

$$p^{(n)}_{k,k+1}(t) = k^{(n)}(t)\lambda\left(\frac{k^{(n)}(t)}{n}\right) = k^{(n)}(t)a = na\frac{k^{(n)}(t)}{n} =: nF^{(n)}_1(n^{-1}k^{(n)}(t)),$$
$$p^{(n)}_{k,k-1}(t) = k^{(n)}(t)\mu\left(\frac{k^{(n)}(t)}{n}\right) = n\left(b\frac{k^{(n)}(t)}{n} + c\frac{k^{(n)}(t)^2}{n^2}\right) =: nF^{(n)}_{-1}(n^{-1}k^{(n)}(t)).$$

They give the corresponding density-dependent Markov process $K^{(n)}(t) := n^{-1}k^{(n)}(t)$ with a state space $\mathcal{E}_n := n^{-1}\mathbb{N} \cup \{0\}$ and the following jump probabilities:

$$\begin{aligned}\mathbb{P}(K^{(n)}(t+\Delta t) = i + n^{-1} \mid K^{(n)}(t) = i) &= nF_1(i)\Delta t + o(\Delta t),\\ \mathbb{P}(K^{(n)}(t+\Delta t) = i - n^{-1} \mid K^{(n)}(t) = i) &= nF_{-1}(i)\Delta t + o(\Delta t).\end{aligned} \tag{2.2}$$

Loosely speaking, it is reasonable to expect that after some time, the density process $K^{(n)}(t)$ will reach (and stay in) a neighborhood of the equilibrium point $\hat{e} := c^{-1}(a-b)$ where $p^{(n)}_{i,i+1}(t)/p^{(n)}_{i,i-1}(t) \approx 1$. To see this, notice that for $k^{(n)}(t)/n > \hat{e}$, the rates yield $p^{(n)}_{i,i-1}(t) > p^{(n)}_{i,i+1}(t)$. This means that on average more individuals will leave/emigrate than be born (with the opposite being true for $k^{(n)}(t)/n < \hat{e}$) i.e. the density will maintain a tendency to move towards the equilibrium.

2.1.2 Epidemic models

Let there be a given population with a fixed size n, in which we want to observe the spread of an epidemic. At time $t = 0$ when we begin our observation, we count a number of i_0 individuals susceptible to the disease, and $n - i_0 = j_0$ infected individuals. We further assume that each susceptible individual makes contact at a rate λ with another random individual, until it becomes infected itself (if the one contacted carried the disease). To do this, to each healthy individual we attach one of a collection of independent identically distributed (i.i.d.) Poisson processes with rate λ to denote the contact times. Moreover, once a person is infected, let the waiting time until they recover and become immune be exponentially distributed with rate μ. Then the two dimensional process $Q^{(n)}(t) := (X_1^{(n)}(t), X_2^{(n)}(t))$, where the components denote the number of susceptible and infected individuals respectively, is Markov with $Q^{(n)}(0) = (i_0, j_0)$, and transition probabilities

$$\begin{aligned} \mathbb{P}(Q^{(n)}(t+\Delta t) =& (i-1, j+1) \mid Q^{(n)}(t) = (i,j)) = \lambda i \frac{j}{n} \Delta t + o(\Delta t), \\ \mathbb{P}(Q^{(n)}(t+\Delta t) =& (i, j-1) \mid Q^{(n)}(t) = (i,j)) = \mu j \Delta t + o(\Delta t) \end{aligned} \tag{2.3}$$

which makes sense because of two reasons. First, we would expect the rate $p^{(n)}_{(-1,+1)}$ with which someone healthy would encounter someone infected should be directly proportional to both the number of healthy individuals and the ratio of infected ones. Second, since the infectious period of any individual carrying the disease is exponentially distributed, the rate $p^{(n)}_{(0,-1)}$ with which someone infected recovers is directly proportional to the number of infected individuals.

Let the real-valued functions $F_{(-1,+1)}(x,y) := xy\lambda$, and $F_{(0,-1)}(x,y) := y\mu$ be defined on $\mathbb{R}^2$. Then, for any time t, we can rewrite the transition rates as

$$\begin{aligned} p^{(n)}_{(-1,+1)}(t) &= X_1^{(n)}(t) \frac{X_2^{(n)}(t)}{n} \lambda = n \frac{X_1^{(n)}(t)}{n} \frac{X_2^{(n)}(t)}{n} \lambda =: n F^{(n)}_{(-1,+1)}(n^{-1} Q^{(n)}(t)) \\ p^{(n)}_{(0,-1)}(t) &= \mu X_2^{(n)}(t) = n\mu \frac{X_2^{(n)}(t)}{n} =: n F^{(n)}_{(0,-1)}(n^{-1} Q^{(n)}(t)) \end{aligned}$$

which, when plugged into (2.3) shows that $\{n^{-1} Q^{(n)}\}_{n \in \mathbb{N}}$ is a density-dependent family.

2.2 Fluid limits and the law of large lumbers

An important property of the examples above (and density-dependent processes in general) is that under mild conditions, they tend to “lose” their randomness and approach a specific deterministic limit for large n. One can, for instance, take the birth-and-death model $K^{(n)}(t) = n^{-1}k^{(n)}(t)$ from the previous section and use the jump probabilities in (2.2) to get

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t}\Big(\mathbb{E}[K^{(n)}(t+\Delta t) \mid K^{(n)}(t)] - K^{(n)}(t)\Big) = F_1(K^{(n)}(t)) - F_{-1}(K^{(n)}(t)),$$

the deterministic analogue of which is the differential equation

$$\frac{d}{dt}K(t) = F_1(K(t)) - F_{-1}(K(t)). \tag{2.4}$$

One may ask if $K^{(n)}$ converges to the deterministic solution K of the differential equation above. This is indeed the case, because provided that $\lim_{n\to\infty} K^{(n)}(0) = K(0)$, almost all paths of $K^{(n)}$ approach K uniformly on bounded time intervals.

The general case for convergence of density-dependent processes to solutions of ordinary differential equations was initially given by Kurtz [9, 11], firstly proving convergence in probability, then almost sure convergence. This result is referred to as the law of large numbers for density-dependent processes and we present it by closely following Ethier and Kurtz's exposition in [4, Chapter 11]. Before we do that, we state the integral version of the well-known Gronwall's inequality, which is used in the proof of Theorem 2.3 (and later in proving our main results in Chapter 4).

Lemma 2.2. (Gronwall's inequality) [4, page 498] Let μ be a Borel measure on $[0,\infty)$, let $\varepsilon \geq 0$, and let f be a nonnegative Borel measurable function that is bounded on bounded intervals and satisfies

$$f(t) \leq \varepsilon + \int_{[0,t)} f(s)\mu(ds), \qquad t \geq 0.$$

Then the following holds:

$$f(t) \leq \varepsilon e^{\mu[0,t)}, \qquad t \geq 0.$$

Theorem 2.3. [4, Theorem 11.2.1] Let $\{n^{-1}Q^{(n)}\}_{n\in\mathbb{N}}$ be the density-dependent family in Definition 2.1, and assume $Q^{(n)}$ satisfies (2.1). Define the function $F(x) := \sum_l lF_l(x)$, and suppose that for each compact $K \subset \mathcal{E}$, we have

$$\sum_l \mid l \mid \sup_{x\in K} F_l(x) < \infty \tag{2.5}$$

and that there exists a constant M_K such that

$$|F(x) - F(y)| \leq M_K|x-y|, \qquad x, y \in K. \tag{2.6}$$

If $n^{-1}q_n \to q$, then for any finite time $T > 0$, we have

$$\mathbb{P}\Big(\omega : \lim_{n\to\infty} \sup_{t\leq T} |n^{-1}Q^{(n)}(t) - Q(t)| = 0\Big) = 1,$$

where $Q(t)$ is the unique solution (provided that it exists) of the integral equation

$$Q(t) = q + \int_0^t F(Q(s))ds. \tag{2.7}$$

Proof. We observe that the validity of the theorem depends on the values of $F_l(x)$ in some neighborhood of $\{Q(t) : t \leq T\}$. Then, let us without loss of generality, assume that conditions (2.5) and (2.6) hold for $\mathcal{E}$, instead of some compact neighborhood K of $\{Q(t) : t \leq T\}$. Defining $\widehat{F}_l := \sup_{x\in\mathcal{E}} F_l(x)$, and $\varepsilon_n(t)$ to be

$$\varepsilon_n(t) := \sup_{u\leq t} \Big| n^{-1}Q^{(n)}(u) - n^{-1}q_n - \int_0^u F(n^{-1}Q^{(n)}(s))ds \Big|,$$

we have by writing $\widehat{J}_l(t) := J_l(t) - t$,

$$\begin{aligned}\lim_{n\to\infty} \varepsilon_n(t) = & \lim_{n\to\infty} \sup_{u\leq t} \Big| \sum_l ln^{-1} \widehat{J}_l\Big(n \int_0^u F_l^{(n)}\big(n^{-1}Q^{(n)}(s)\big)ds\Big) \Big| \\ & \leq \lim_{n\to\infty} \sum_l \mid l \mid n^{-1} \sup_{u\leq t} \big| \widehat{J}_l(n\widehat{F}_l u) \big| \leq \lim_{n\to\infty} \sum_l \mid l \mid n^{-1} \big(J_l(n\widehat{F}_l t) + n\widehat{F}_l t \big)\end{aligned}$$

where the last inequality is term-wise, i.e. for every natural n, and $l \in \mathbb{Z}^d$, we have

$$\sup_{u\leq t} \big| \widehat{J}_l(n\widehat{F}_l u) \big| \leq J_l(n\widehat{F}_l t) + n\widehat{F}_l t.$$

Then by the LLN, for sufficiently large n, the rightmost expression in the brackets is bounded from above by $3\widehat{F}_l t$, so by (2.5) and the dominated convergence theorem (DCT), we can interchange the limit and the sum to obtain

$$\lim_{n\to\infty}\sum_l \mid l \mid n^{-1}\big(J_l(n\widehat{F}_l t) + n\widehat{F}_l t\big) = \sum_l \lim_{n\to\infty} \mid l \mid n^{-1}\big(J_l(n\widehat{F}_l t) + n\widehat{F}_l t\big).$$

Then by another application of a generalized DCT [4, page 492], we get

$$\lim_{n\to\infty}\sum_l \mid l \mid n^{-1}\sup_{u\le t}\big|\widehat{J}_l(n\widehat{F}_l u)\big| = \sum_l \lim_{n\to\infty} \mid l \mid n^{-1}\sup_{u\le t}\big|\widehat{J}_l(n\widehat{F}_l u)\big| = 0, \text{ a.s.,}$$

where the convergence to 0 is obtained by a standard use of Doob's martingale inequality, followed by Borel-Cantelli's lemma. Finally, since F is M-Lipschitz continuous, we have

$$\big|n^{-1}Q^{(n)}(t) - Q(t)\big| \le \big|n^{-1}q_n - q\big| + \varepsilon_n(t) + M\int_0^t \big|n^{-1}Q^{(n)}(s) - Q(s)\big| ds,$$

and due to Gronwall's inequality,

$$\sup_{t\le T}\big|n^{-1}Q^{(n)}(t) - Q(t)\big| \le \big(\big|n^{-1}q_n - q\big| + \varepsilon_n(T)\big)e^{Mt} \to 0, \quad \text{a.s.}$$

as $n \to \infty$ because $\varepsilon_n(T) \to 0$ a.s. due to our discussion above, and $\big|n^{-1}q_n - q\big| \to 0$ because we posed it as a condition. This finishes the proof. $\square$

In other words, given that the drift function F is Lipschitz continuous and bounded on compact subsets of $\mathcal{E}$, almost all paths of a density-dependent process $n^{-1}Q^{(n)}$ do indeed converge (uniformly on compact sets) to the solution of (2.7). This deterministic function Q bears a special name, *"the fluid limit"* of the sequence $\{n^{-1}Q^{(n)}\}_{n\in\mathbb{N}}$.

Since we have shown that the logistic growth model in Section 1 is density-dependent, we can use Theorem 2.3 to see if the process $K^{(n)}$ actually converges to the solution of (2.4). Inspecting how the drift function was defined for $K^{(n)}(t)$ we see that $F_1(x) = ax$ and $F_{-1}(x) = bx + cx^2$, both of which are Lipschitz continuous on compact sets in $\mathcal{E} = [0,\infty)$. Assuming that $K^{(n)}(0) \to k$, where $K^{(n)}(0)$ and $k > 0$ are non-random, we get that a.s. $\sup_{t\le T}|K^{(n)}(t) - K(t)| \to 0$, as $n \to \infty$, where $K(t)$ is the solution to the equation

$$\frac{d}{dt}K(t) = (a-b)K(t) - cK(t)^2, \quad K(0) = k,$$

and is written explicitly as

$$K(t) := \frac{k\hat{e}}{k + (\hat{e} - k)\exp(-(a-b)t)},$$

where $\hat{e} := c^{-1}(a-b)$. A noteworthy observation is that as $t \to \infty$, $K(t) \to \hat{e}$, similarly to how the stochastic process $K^{(n)}$ behaved in the logistic growth model.

2.3 Diffusion limits and the central limit theorem

Having the law of large numbers provides information about how $Q^{(n)}$ acts when rescaled by n^{-1}, as then the process is very well approximated by the fluid limit Q. A second matter of interest would be whether there is a useful approximation result for the distribution of the process. For this purpose one is inclined to investigate the fluctuations of $n^{-1}Q^{(n)}(t)$ around $Q(t)$ by observing the distribution of the process

$$V^{(n)}(t) := \sqrt{n}(n^{-1}Q^{(n)}(t) - Q(t)), \text{ as } n \to \infty.$$

Already, this hints at a CLT for density-dependent Markov processes. Indeed, for $Q^{(n)}(t)$ satisfying (2.1), we can rewrite $V^{(n)}(t)$ to be equal to

$$V^{(n)}(0) + \sqrt{n}\Big(\sum_l \frac{l}{n} J_l\Big(n\int_0^t F_l\big(n^{-1}Q^{(n)}(s)\big)ds\Big) - \int_0^t F(Q(s))ds\Big)$$

$$= v_n + \sum_l lW_l^{(n)}\Big(\int_0^t F_l\big(n^{-1}Q^{(n)}(s)\big)ds\Big) + \sqrt{n}\int_0^t \big(F(n^{-1}Q^{(n)}(s)) - F\big(Q(s)\big)\big)ds$$

where $W_l^{(n)}(t) := n^{-1/2}(J_l(nt) - nt)$. A couple of comments and predictions are in place regarding the last expression.

- Because J_l is a standard Poisson process, it is known that by the functional CLT, $W_l^{(n)}(\cdot)$ converges to a d-dimensional standard Brownian motion $W_l(\cdot)$. This suggests that $\sum_l lW_l^{(n)}\Big(\int_0^\cdot F_l\big(n^{-1}Q^{(n)}(s)\big)ds\Big)$ may converge in distribution to the limit candidate $\sum_l lW_l\Big(\int_0^\cdot F_l\big(Q(s)\big)ds\Big)$.

- Provided that the drift function F and its derivative $\partial F := (\partial_j F_i)$ are continuous in a neighborhood of $Q(t)$, one is able to use a first-order Taylor approximation to write

$$\begin{aligned}&\sqrt{n}\int_0^t \big(F(n^{-1}Q^{(n)}(s)) - F\big(Q(s)\big)\big)ds\\ &= \int_0^t \partial F(Q(s))V^{(n)}(s)ds + \int_0^t O\big(|n^{-1}Q^{(n)}(s) - Q(s)|\big)V^{(n)}(s)ds,\end{aligned}$$

  where Theorem 2.1 suggests that the far-right term may vanish.

These two observations suggest the possibility that the process $V^{(n)}(t)$ converges in distribution to the diffusion $V(t)$, defined as the solution of the equation

$$V(t) = v + \sum_l lW_l\Big(\int_0^t F_l\big(Q(s)\big)ds\Big) + \int_0^t \partial F\big(Q(s)\big)V(s)ds. \tag{2.8}$$

This is indeed the case, under some continuity conditions, as our next theorem states. So that no confusions are encountered, the notation $V^{(n)} \Rightarrow V$ will mean that the stochastic processes are considered as random variables taking values in the space $D([0,\infty),\mathbb{R}^d)$ (i.e. the space of all càdlàg functions from $[0,\infty)$ to $\mathbb{R}^d$) equipped with the Skorohod $\mathcal{J}_1$ topology, and in this sense $V^{(n)}$ converges to V in distribution.

Theorem 2.4. Let $Q^{(n)}$ satisfy (2.1) and Q be the solution of (2.7). Moreover, suppose that F and ∂F are continuous on $\mathcal{E}$, and that for each compact set $K \in \mathcal{E}$, we have

$$\sum_l |\, l\,|^2 \sup_{x\in K} F_l(x) < \infty.$$

If $V^{(n)}$ satisfies $\lim_{n\to\infty} V^{(n)}(0) \to V(0) = v$, then we have $V^{(n)} \Rightarrow V$, where $\big(V(t)\big)_{t\geq 0}$ is the solution to equation (2.8).

Proof. The reader is directed to [4, Theorem 11.2.3]. A similar proof, (an adaptation of which is used later in our own work in Chapter 4.3) is given in [1, Theorem 5.4]. □

The main implication of Theorem 2.4 is that we are able to use the approximation $n^{-1}Q^{(n)} \approx Q + n^{-1/2}V$ for large n, which is useful because V is a d-dimensional Gaussian process whose finite-dimensional distributions are completely determined by F. More specifically, define $G : \mathcal{E} \to \mathbb{R}^{d\times d}$ as $G(x) := \sum_l ll^T F_l(x)$, and let the function $\Phi(t,s)$ be the solution of the forward Kolmogorov equation

$$\frac{\partial}{\partial t}\Phi(t,s) = \partial F(Q(t))\Phi(t,s), \quad \Phi(s,s) = I. \tag{2.9}$$

Then, $V(t)$ has expectation and covariance functions respectively given by

$$\begin{aligned}\mathbb{E}[V(t)] =& \Phi(t,0)v, \qquad \text{and}\\ \mathbf{Cov}[V(t),V(r)] =& \int_0^{r\wedge t} \Phi(t,s)G(Q(s))[\Phi(r,s)]^T ds.\end{aligned}$$

Remark 2.5. The covariance and expectation function can be defined similarly by using $\Phi^*(t,s)$, the solution of the backward Kolmogorov equation

$$\frac{\partial}{\partial s}\Phi^*(t,s) = -\Phi^*(t,s)\partial F(Q(s)), \quad \Phi^*(t,t) = I.$$

This approach is used in [1] and will be used later in this text (see Theorem 4.15).

Chapter 3

THE POISSON ENCOUNTER-MATING MODEL

Recently, motivated by Gimelfarb's work, Gün and Yılmaz [7] introduced the so-called stochastic encounter-mating (SEM) model, which combines both the individual EM and mass EM models that Gimelfarb proposed. Indeed, by attaching to each individual in the population an independent mating clock (in the form of a point process), Gün and Yılmaz generalize the former models to encompass a plethora of different scenarios, two special cases of which are the individual and mass EM.

Specifically, when the point processes are Poisson with rates dependent on the sex and type of each individual, the discrete Markov chain embedded in the pair-type process $(Q^{(n)}(t))_{t\geq 0}$ that emerges from the SEM (we define $Q^{(n)}$ properly later in this chapter), is precisely the individual EM. On the other hand, if the point processes were Bernoulli (also, with sex and type dependent rates) the pair-type process acts the same way as in the mass EM. In this chapter we only give details about the Poisson case of the SEM (we write PEM from now on), as it is the one related to our results in Chapter 4, and refer the reader to [7] for the general exposition. However, we do list a number of results achieved for both the Poisson and Bernoulli scenarios. These results rigorously prove some of Gimelfarb's arguments, and settle (as well as strengthen) two of his conjectures regarding scenarios that produce random-like mating patterns, due to the lack of correlation between different types.

3.1 Construction and properties

3.1.1 Population setup

Let the sets Z_f and Z_m each have n elements, which we will conveniently refer to as females and males, respectively. We categorize the individuals into different types and to do this we choose a fixed natural number $k \geq 2$ to represent the number of types in question. Furthermore we require each individual $\zeta \in Z := Z_f \cup Z_m$ to belong to exactly one type; this gives rise to a population of $2n$ individuals (called an n-population) in which we have

$$\sum_{i=1}^{k} x_i^{(n)} = \sum_{j=1}^{k} y_j^{(n)} = n, \tag{3.1}$$

where $x_i^{(n)}$ is the number of i-type females and $y_j^{(n)}$ the number of j-type males.

3.1.2 Encounter stage

Suppose that to each $\zeta \in Z$ a mating clock that fires at random times is attached, and a counting process $N_\zeta(t)$ keeps track of the number of times the clock fired up to time $t \geq 0$. Let the counting processes $\{N_\zeta(t) : \zeta \in Z\}$ satisfy the following conditions:

- $\{N_\zeta(t) : \zeta \in Z\}$ are independent Poisson processes on $[0, \infty)$;
- each Poisson process $N_\zeta(t)$ has rate α_i (resp. β_j) if ζ is a i-female (resp. j-male);
- for each $i, j \in [k] := \{1, 2, \ldots, k\}$ we have $\alpha_i + \beta_j > 0$.

The role of the Poisson clocks is that they denote the encounter times: whenever a mating clock fires, the corresponding single i-female (resp. j-male) picks a candidate single j-male (resp. i-female) uniformly at random and initiates an encounter. This encounter represents a temporary ij-bond. The bond is then assigned an independent Bernoulli random variable, with probability of success dependent only on the two types involved in the bond, namely $p_{ij} \in (0, 1]$.

3.1.3 Mating stage

Given that the temporary bond succeeds, it becomes permanent and the individuals who formed the ij-pair become unavailable for further mating and exit the singles pool. If the encounter fails however, both of them go back to the singles pool and continue with the encounter-mating process until they form a pair. This continues until the pool of singles is depleted and a total of n pairs have formed. Due to the conditions imposed on $N_\zeta(t)$, almost surely no two different encounters occur at the same time, so the process only assumes jumps of size 1. Also, since for each ij-pair both p_{ij} and $\alpha_i + \beta_j$ are positive, the time T_n when the process terminates is almost surely finite.

3.2 *The pair-type process $Q^{(n)}$ and its Markov property*

To count the formation of pairs between different types, we define the nonnegative integer valued $k \times k$ matrix $Q^{(n)}(t) := (Q^{(n)}_{ij}(t))_{i,j\in[k]}$ such that $Q^{(n)}_{ij}(t)$ is the number of permanent bonds between i-females and j-males formed up to time $t \geq 0$. Let us define, for a given matrix $M \in \mathbb{R}^{k\times k}$, the following notation for the i-th row (respectively j-th column and total) sum as

$$M_{i,\cdot} := \sum_{j'=1}^{k} M_{ij'}, \quad M_{\cdot,j} := \sum_{i'=1}^{k} M_{i'j}, \text{ and } M_{tot} := \sum_{i',j'\in[k]} M_{i'j'}.$$

Then, using this notation the expressions that give the remaining number of single i-females, j-males, and unformed pairs at any time t can be written as

$$x_i^{(n)} - Q^{(n)}_{i,\cdot}(t), \;\; y_j^{(n)} - Q^{(n)}_{\cdot,j}(t), \text{ and } \;n - Q^{(n)}_{tot}(t), \tag{3.2}$$

respectively. The process $Q^{(n)}$ naturally takes values in the space $\mathcal{E}^*_n$ defined as

$$\mathcal{E}^*_n := \left\{ M \in \mathcal{M}^{k\times k}(\mathbb{N}\cup\{0\}) : M_{i,\cdot} \leq x_i^{(n)}, M_{\cdot,j} \leq y_j^{(n)}, \forall i,j \in [k] \right\} \tag{3.3}$$

where $\mathcal{M}^{k\times k}(\mathcal{A}) := \mathcal{A}^{k\times k}$, for some $\mathcal{A} \subseteq \mathbb{R}$.

Due to the Poisson nature of $\{N_\zeta(t)\}_{\zeta\in Z}$, waiting times between clock firings are exponential, and therefore have the memoryless property. Combined with the fact that only one clock fires at a time (and therefore only one permanent bond can be formed at a time), this implies that $Q^{(n)}$ is a continuous-time Markov chain, with jumps being the canonical base vectors of the vector space $\mathcal{M}^{k\times k}(\mathbb{R}) = \mathbb{R}^{k\times k}$. We will henceforth denote them as I^{ij}.

Next, as none of the Poisson clocks have fired and therefore no bonds have yet formed by time $t = 0$, we obviously have $Q^{(n)}(0) = M_0$, where M_0 is the zero $k \times k$ matrix. Lastly, to obtain the transition rates for $Q^{(n)}$ we just have to make an elementary observation. Namely, if at a certain time, $t > 0$, the process $Q^{(n)}(t)$ is at the state $M \in \mathcal{E}_n^*$, then exactly $x_i^{(n)} - M_{i,\cdot}$ single i-females fire, each at a rate α_i, and form a temporary bond with a j-male (out of the remaining $n - M_{tot}$ males) with probability $\frac{y_j^{(n)} - M_{\cdot,j}}{n - M_{tot}}$. However such a ij-bond becomes permanent with probability p_{ij}. By combining this with the second way to form a ij-pair, which is by having single j-males fire at the single i-females at rate β_j, we see that the transition rates of $Q^{(n)}$ are given by the function

$$\rho^{(n)}(M, M') = \sum_{i,j \in [k]} \pi_{ij} \frac{(x_i^{(n)} - M_{i,\cdot})(y_j^{(n)} - M_{\cdot,j})}{n - M_{tot}} \mathbb{1}_{\{M' = M + I^{ij}, M_{tot} < n\}} \tag{3.4}$$

where $\pi_{ij} := p_{ij}(\alpha_i + \beta_j)$, and the indicator function $\mathbb{1}_{\{M'=M+I^{ij},M_{tot}<n\}}$ ensures that we have only unit jumps, until a total of n-pairs are formed, at which point the process stops.

Although we have specified all necessary characteristics of the pair-type process $Q^{(n)}$, one may notice that we have refrained from placing it on an appropriate probability space, and therefore from using any probability notation. This is done deliberately, and although defining the space induced by $Q^{(n)}$ is achievable (and done in rigorous detail in [7]), this is not the space we shall be working with. Instead, we will use the Poisson characterization (see equation (2.1)) given in Chapter 2, to define a common space on which $Q^{(n)}$ is measurable for all n.

3.3 The Poisson fine balance condition and panmixia

An important observation involving the expression at (3.4) is that as long as π_{ij} remains unchanged, how we adjust the parameters p_{ij}, α_i, and β_j is of no importance to the jumping rates (and therefore the distribution) of the pair-type process $Q^{(n)}$. For example, we may tune the values of α_i and β_j so that only females and no males fire, and still not change the finite-dimensional distributions of $Q^{(n)}$. This motivates the following definition.

Definition 3.1. Given an n-population Z with types satisfying (3.1), let $Q^{(n)}$ be its corresponding pair-type process. Then the $k \times k$ matrix Π defined as

$$\Pi := (\pi_{ij})_{i,j\in[k]},$$

where $\pi_{ij} := p_{ij}(\alpha_i + \beta_j)$ is called the *encounter-mating* (or shortly EM) law of $Q^{(n)}$.

Remark 3.2. In fact, this is not the way Gün and Yılmaz define the EM law, in [7]. The correct, general approach is to include both the matrix Π *and* the joint distributions of the point processes $\{N_\zeta(t)\}_{\zeta\in Z}$. Since in this thesis we only cover the case where $\{N_\zeta(t)\}_{\zeta\in Z}$ are Poisson, we took the liberty of narrowing this definition down to just the matrix Π.

In order to fully appreciate the EM law, we need to discuss some of Gimelfarbs results in more detail. When inspecting mating in a population consisting of different types of males and females, one is inclined to learn whether the resulting mating patterns exhibit correlations between different types. As one would expect, no correlations between, say, type-i females and type-j males, would correspond to the frequency of type ij-pairs being just the product of each of the constituting types' frequencies. The term used in the literature to refer to this is "panmixia" (or "random mating"), and represents a concept of great significance in population genetics.

Among others, it was Gimelfarb [5] who pointed out that it is a misfortune that the terms panmixia and random mating are used interchangeably, when in fact the latter implies the former but not vice versa. It is for this reason, he says, that the term panmixia should be used only when discussing mating patterns with no correlations between certain types, whereas the term random mating should describe mating processes with no type-dependent mating preferences whatsoever. He supports his claim by performing simulations with different mating preferences, some of which result in panmictic mating patterns. Due to similarities shared by the preferences which did give rise to panmixia, Gimelfarb conjectures two things:

i) if there exist positive reals $\bar{\alpha}_1, \bar{\alpha}_2, \ldots, \bar{\alpha}_k$ and $\bar{\beta}_1, \bar{\beta}_2, \ldots, \bar{\beta}_k$ such that $\forall i, j \in [k]$, $p_{ij} = \bar{\alpha}_i + \bar{\beta}_j$, then as $n \to \infty$ the individual EM results in panmixia;

ii) if there exist positive reals $\bar{\gamma}_1, \bar{\gamma}_2, \ldots, \bar{\gamma}_k$ and $\bar{\delta}_1, \bar{\delta}_2, \ldots, \bar{\delta}_k$ such that $\forall i, j \in [k]$, $p_{ij} = 1 - \bar{\gamma}_i \bar{\delta}_j$, then as $n \to \infty$ the mass EM results in panmixia.

However, in most of his analysis, Gimelfarb simplifies his calculations by replacing what would normally be random quantities (for example ratios of pair-types present in the population) by their expected values. He adopts these approximations by saying that they are reliable for large populations, which is why his conjectures are stated in asymptotic context. Moreover, he does not provide a proof for his first conjecture (instead he supports his claim with simulations) whereas he does present an argument as a proof for the second conjecture. However this argument lacks rigor due to his simplifying approximations.

To approach this issue, Gün and Yılmaz define two so-called *fine balance conditions*, one each for the Poisson and Bernoulli cases. The Poisson fine balance condition is a statement about the EM law Π of the PEM, requiring it to satisfy

$$\pi_{ij} + \pi_{i'j'} = \pi_{ij'} + \pi_{i'j} \quad \text{for any} \quad i, j, i', j' \in [k], \tag{3.5}$$

and it is proved in the same paper that this condition indeed implies panmixia, via the following theorem.

Theorem 3.3. [7, Theorem 3.6] In a population satisfying (3.1), let the EM law of $Q^{(n)}$ have the Poisson fine balance property (3.5). Then, given that $(\Omega, \mathcal{F}, \mathbb{P})$ is the space induced by $Q^{(n)}$, for every $t \geq 0$ and $M \in \mathcal{E}_n^*$ we have

$$\mathbb{P}(Q^{(n)}(t) = M) = \frac{(\prod_i x_i^{(n)}!)(\prod_j y_j^{(n)}!)}{n!(\prod_{i,j} m_{ij}!)} \sum_{\substack{M' \in \widehat{\mathcal{E}_n^*} \\ M' \geq M}} \prod_{i,j} \frac{(1 - e^{-\pi_{ij} t})^{m_{ij}} (e^{-\pi_{ij} t})^{m'_{ij} - m_{ij}}}{(m'_{ij} - m_{ij})!}$$

where the matrix relation $M' \geq M$ is defined term-wise, and

$$\widehat{\mathcal{E}_n^*} := \left\{ M \in \mathcal{M}^{k \times k}(\mathbb{N} \cup \{0\}) : M_{i,\cdot} = x_i^{(n)}, M_{\cdot,j} = y_j^{(n)}, \forall i, j \in [k] \right\}.$$

Similarly,

$$\mathbb{P}(Q^{(n)}(t) = M) = \frac{(\prod_i x_i^{(n)}!)(\prod_j y_j^{(n)}!)}{n!(\prod_{i,j} m_{ij}!)} \quad \text{for } M \in \widehat{\mathcal{E}_n^*},$$
$$\mathbb{E}[Q_{ij}^{(n)}(t)] = \frac{x_i^{(n)} y_j^{(n)}}{n}(1 - e^{-\pi_{ij}t}), \quad \text{and} \quad \mathbb{E}[Q_{ij}^{(n)}(T_n)] = \frac{x_i^{(n)} y_j^{(n)}}{n}.$$

However, Gün and Yılmaz also show the converse [7, Proposition 3.4], i.e. that $\mathbb{E}[Q_{ij}^{(n)}(T_n)] = n^{-1}x_i^{(n)}y_j^{(n)}$ implies that Π satisfies equation (3.5). Finally, by proving that this property is equivalent to Gimelfarb's condition in his first hypothesis (i), they not only settle Gimelfarb's conjecture but also strengthen it by showing that panmictic mating patterns in the individual EM model happen *if and only if* condition (i) holds.

3.4 *Fluid limit and LLN for the mating pattern*

If we were to assume Poisson fine balance, then by Theorem 3.3, for any fixed $t > 0$, the expected values of $n^{-1}Q_{ij}^{(n)}(t)$ and $n^{-1}Q_{ij}^{(n)}(T_n)$ are $(n^{-1}x_i^{(n)})(n^{-1}y_j^{(n)})(1 - e^{-\pi_{ij}t})$ and $(n^{-1}x_i^{(n)})(n^{-1}y_j^{(n)})$ respectively. In other words, the expected ratio of ij-pairs depends on the product of corresponding ratios of i-females and j-males. Now, suppose that the zoological n-population Z we consider is only a part of the entire species and that as $n \to \infty$, the ratios of the k different types converge, i.e.

$$n^{-1}x_i^{(n)} \to x_i, \quad \text{and} \quad n^{-1}y_j^{(n)} \to y_j \quad \forall i, j \in [k]. \tag{3.6}$$

It is evident that the expectations of $n^{-1}Q^{(n)}(t)$ and $n^{-1}Q^{(n)}(T_n)$ converge to the values $x_i y_j(1 - e^{-\pi_{ij}t})$ and $x_i y_j$, respectively. However, a much stronger and more general result about the rescaled pair-type process $n^{-1}Q^{(n)}$ can be proved. In fact, even without the fine balance condition, the sequence $\{n^{-1}Q^{(n)}\}_{n\in\mathbb{N}}$ converges to a fluid limit (recall Theorem 2.3, and the discussion in the succeeding paragraph) although it only partially fits into the framework of density-dependent processes we have covered in Chapter 2. Gün and Yılmaz deal with these discrepancies in [6] and prove two LLN-type results, one for the pair-type process (i.e. the fluid limit), and another for the resulting mating pattern. This section closely follows the work in that paper.

3.4.1 Approximate density-dependence and inconsistencies

Let the vector space $\mathcal{M}^{k\times k}(\mathbb{R})$ be equipped with the norm $|\cdot|$, defined as

$$|M| := \max_{i,j\in[k]} |M_{ij}|, \quad \text{for } M \in \mathcal{M}^{k\times k}(\mathbb{R}), \tag{3.7}$$

and have its canonical base $\{I^{ij} : i,j \in [k]\}$. Clearly, $\mathcal{E}_n := \{n^{-1}M : M \in \mathcal{E}_n^*\}$ is the state space of the rescaled process $n^{-1}Q^{(n)}$, and is explicitly defined as

$$\mathcal{E}_n := \big\{M \in \mathcal{M}^{k\times k}(n^{-1}\mathbb{N}\cup\{0\}) : M_{i,\cdot} \le n^{-1}x_i^{(n)}, M_{\cdot,j} \le n^{-1}y_j^{(n)}, \forall i,j \in [k]\big\}. \tag{3.8}$$

Since as $n \to \infty$ we assumed (3.6), we also introduce a limiting space for the infinite population case, which is

$$\mathcal{E} := \big\{M \in \mathcal{M}^{k\times k}([0,\infty)) : M_{i,\cdot} \le x_i, M_{\cdot,j} \le y_j, \forall i,j \in [k]\big\}. \tag{3.9}$$

Note that for every n, $\mathcal{E}_n \setminus \mathcal{E} \neq \emptyset$. Next, we define $F^{(n)} : \mathcal{E}_n \to \mathcal{M}^{k\times k}([0,\infty))$ and $F : \mathcal{M}^{k\times k}([0,\infty)) \to \mathcal{M}^{k\times k}([0,\infty))$ as

$$\begin{aligned} F_{ij}^{(n)}(M) &:= \pi_{ij}\frac{(n^{-1}x_i^{(n)} - M_{i,\cdot})(n^{-1}y_j^{(n)} - M_{\cdot,j})}{1 - M_{tot}}\mathbb{1}_{\{M_{tot}<1\}} \\ F_{ij}(M) &:= \pi_{ij}\frac{(x_i - M_{i,\cdot})(y_j - M_{\cdot,j})}{1 - M_{tot}}\mathbb{1}_{\{M_{tot}<1\}}. \end{aligned} \tag{3.10}$$

Taking into perspective the transition rates of $Q^{(n)}$ in (3.4), we observe that they are related to $F^{(n)}$ in the following way:

$$\rho^{(n)}(M, M + I^{ij}) = nF_{ij}^{(n)}(n^{-1}M).$$

Although the rates here clearly depend on the density of the process at any given state, the family $\{n^{-1}Q^{(n)}\}_{n\in\mathbb{N}}$ fails to meet the requirements given in Definition 2.1. Indeed, the matrix-valued drift function $F^{(n)}$ is n-dependent, and thus Theorem 2.3 cannot tell us whether $n^{-1}Q^{(n)}$ converges to a fluid limit as $n \to \infty$. If it did converge, however, this limit would be the function $Q(t) : [0,\infty) \to \mathcal{E}$, given as the unique solution of the integral equation

$$Q(t) = \int_0^t F\big(Q(s)\big)ds. \tag{3.11}$$

3.4.2 Presentation and statement of the theorems

A result due to Kurtz [11, Theorem 2.2] covers a wider class of processes than Theorem 2.3, but even for this theorem to guarantee a fluid limit two conditions are needed: firstly, for all n, F has to be Lipschitz continuous on $\mathcal{E} \cup \mathcal{E}_n$ and secondly, the convergence $|F^{(n)}(M) - F(M)| = O(n^{-1})$ needs to hold uniformly on $\mathcal{E}_n$. In the particular case of the PEM model, however, it can be proved that a weaker convergence rate of $o(1)$ holds instead. This makes the rescaled pair-type process family a density-dependent one, but only approximately, i.e. as $n \to \infty$. Moreover, the fact that $\mathcal{E}_n \subseteq \mathcal{E}$ does not hold poses difficulties if one would normally want to use the Lipschitz continuity of F around a neighborhood of the function $Q(t)$.

By working through these discrepancies, it can be proved that even though the sequence $\{n^{-1}Q^{(n)}\}_{n\in\mathbb{N}}$ used for the PEM model is not density-dependent, it does admit a fluid limit as the population size converges to infinity. We note that since the pair-type process can make at most n jumps, the Poisson characterization from Chapter 2 can be used to write

$$Q^{(n)}(t) = \sum_{i,j\in[k]} I^{ij} J_{ij}\Big(n \int_0^t F_{ij}^{(n)}\big(n^{-1}Q^{(n)}(s)\big)ds\Big), \quad t \geq 0, \tag{3.12}$$

for any n, on some sample space $(\Omega, \mathcal{F}, \mathbb{P})$ induced by the independent standard Poisson processes $\{J_{ij}(t)\}_{i,j\in[k]}$. Using this representation (as in Definition 2.1), Gün and Yılmaz proved the analogue of Theorem 2.3, which establishes the convergence to a deterministic function on bounded time intervals.

Theorem 3.4. In an n-population Z, let Π be the EM law of the PEM model , $F^{(n)}$ and F be given as in (3.10), and the pair-type process be given by (3.12). Then the unique solution $Q(t)$ of equation (3.11) exists, and under condition (3.6), for any $T \in [0, \infty)$,

$$\mathbb{P}\Big(\omega : \lim_{n\to\infty} \sup_{t\leq T} |n^{-1}Q^{(n)}(t) - Q(t)| \to 0\Big) = 1.$$

Proof. The reader is referred to [6, Theorem 2.1], as the proof relies on several auxiliary lemmas that deal with the discrepancies we have mentioned above. $\square$

Another question is whether any information about the rescaled mating pattern $n^{-1}Q^{(n)}(T_n)$ (recall that $T_n = \inf\{t \geq 0 : Q^{(n)}_{tot}(t) = n\}$), can be extracted from the deterministic function $Q(t)$. The answer is yes, and is given by the following LLN which can be considered as an extension to the fluid limit theorem above.

Theorem 3.5. [6, Theorem 2.2] Let $n^{-1}Q^{(n)}(T_n)$ denote the rescaled mating pattern, taking values in

$$\mathcal{E}_n' := \{M \in \mathcal{M}^{k\times k}([0,\infty)) : M_{i,\cdot} = n^{-1}x_i^{(n)}, M_{\cdot,j} = n^{-1}y_j^{(n)}; \forall i,j \in [k]\}.$$

Then as $n \to \infty$, $\mathbb{P}$-a.s.

$$n^{-1}Q^{(n)}(T_n) \to Q(\infty),$$

where the matrix $Q(\infty) := \lim_{t\to\infty} Q(t)$ exists, and takes values in

$$\mathcal{E}' := \{M \in \mathcal{M}^{k\times k}([0,\infty)) : M_{i,\cdot} = x_i, M_{\cdot,j} = y_j; \forall i,j \in [k]\}.$$

Informally, one can think of $Q(t)$ as how the pair-type process would act for infinite populations. To see how the last two theorems extend and strengthen Theorem 3.3, one needs to know how $Q(t)$ behaves when the PEM law Π satisfies condition (3.5). Conveniently enough, even though the function does not, in general, have a closed form, in such a case it can be written explicitly (see [6, Section 3.2] for details) as

$$Q_{ij}(t) = x_i y_j(1 - e^{-\pi_{ij}t}) \quad \text{and} \quad Q_{ij}(\infty) = xy,$$

which corresponds to what the expected values converge to in our discussion at the beginning of this section. This motivates the following definition from [6], where Gün and Yılmaz precisely state, and solidify the notion of panmictic species, adding mathematical rigor to Gimelfarb's concept.

Definition 3.6. An infinite population is said to be panmictic if, given the ratios of each type $x_1, \dots, x_k, y_1, \dots, y_k$, the following limit holds:

$$\lim_{n\to\infty} n^{-1}Q^{(n)}_{ij}(T_n) = x_i y_j.$$

If every infinite population of a species is panmictic, then the species itself is panmictic.

Chapter 4

DIFFUSION LIMIT

4.1 Discussion, statement and preliminary results

As presented in the previous chapter, and proven in [6], the fluid limit theorem and the LLN for the mating pattern show that under assumption (3.6), both the continuous-time matrix-valued Markov chain $(n^{-1}Q^{(n)}(t))_{t\geq 0}$ and the random matrix $n^{-1}Q^{(n)}(T_n)$ converge $\mathbb{P}$-a.s. to their corresponding deterministic limits, as $n \to \infty$. This gives motivation for one to go on further to prove the CLT counterparts of the above-mentioned theorems. Specifically, given the sequence $\{n^{-1}Q^{(n)}\}_{n\in\mathbb{N}}$, as a complement to the fluid limit Q, one can try to prove a distributional convergence to (and existence thereof) a diffusion limit V for the sequence $\{V^{(n)}\}_{n\in\mathbb{N}}$, where

$$V^{(n)}(t) := \sqrt{n}\big(n^{-1}Q^{(n)}(t) - Q(t)\big), \quad t \geq 0. \tag{4.1}$$

Furthermore, by an analogous reasoning, one could furthermore attempt to prove a CLT for the mating pattern by analyzing the limiting behavior of $\{V^{(n)}(T_n)\}_{n\in\mathbb{N}}$, where

$$V^{(n)}(T_n) := \sqrt{n}\big(n^{-1}Q^{(n)}(T_n) - Q(\infty)\big).$$

This chapter is intended to provide a rigorous verification to the former of the two claims, under a stronger version of assumption (3.6) which requires

$$\frac{x_i^{(n)} - nx_i}{\sqrt{n}} \to 0, \qquad \text{and} \qquad \frac{y_j^{(n)} - ny_j}{\sqrt{n}} \to 0 \tag{4.2}$$

for all $i, j \in [k]$. We present two different proofs for the result, the first one via a coupling approach, and the second one via a modification of the proof of Theorem 2.4 provided in [4, Chapter 11.2]. We write $V^{(n)}(t) \Rightarrow V(t)$ to denote distributional convergence of the $\mathcal{M}^{k\times k}(\mathbb{R})$-valued random variable $V^{(n)}(t)$ for some fixed t. However, when we write $V^{(n)} \Rightarrow V$ or $V^{(n)}(\cdot) \Rightarrow V(\cdot)$, we denote distributional convergence as well, but this time it is the processes $V^{(n)}$ and V that are considered as random variables taking values in the space $D([0,T),E)$ equipped with the Skorohod $\mathcal{J}_1$ topology.

Throughout this chapter, $T \in [0,\infty)$, and the state space E is either $\mathcal{M}^{k\times k}(\mathbb{R})$ (used in Section 4.2), or $\mathbb{R}^{k^2}$ (used in Section 4.3). Whenever the need to make the formal transition from a matrix vector space to a Euclidean one arises, the canonical base $\{I^{ij}\}_{i,j\in[k]}$ used for $\mathcal{M}^{k\times k}(\mathbb{R})$, will be substituted with $\{\ell_{ij}\}_{i,j\in[k]}$, where $\ell_{ij} \in \mathbb{R}^{k^2}$ is the canonical $\big((j-1)k+i\big)$-th unit vector. To help visualize this transition, one can take a matrix $M \in \mathcal{M}^{k\times k}(\mathbb{R})$ and stack the columns on top of one another (until a single column vector has formed), so that columns further on the left would be above the ones on the right.

The reader should be aware of a slight abuse of notation, prominent in Section 4.3, where we work in $\mathbb{R}^{k^2}$. Due to the fact that there exists an isometry between $\mathcal{M}^{k\times k}(\mathbb{R})$ endowed with the $|\cdot|$ norm defined in (3.7), and $\mathbb{R}^{k^2}$ endowed with the well-known $|\cdot|_\infty$ norm (explicitly defined in Section 4.3), we refrain ourselves from writing $|\cdot|_\infty$, and use $|\cdot|$ instead.

Since it will be extensively used throughout the chapter, we recall that for $F^{(n)}$ given in (3.10), and a collection of independent Poisson processes $\{J_{ij}(t)\}_{i,j\in[k]}$ defined on a common probability space $(\Omega,\mathcal{F},\mathbb{P})$, the equation

$$Q^{(n)}(t) = \sum_{i,j\in[k]} I^{ij} J_{ij}\Big(n\int_0^t F_{ij}^{(n)}\big(n^{-1}Q^{(n)}(s)\big)ds\Big) \tag{4.3}$$

uniquely defines the pair-type process for any n on $(\Omega,\mathcal{F},\mathbb{P})$. This characterization, as we will see, is crucial for our analysis, as it will be used in both of the proofs to our main result, which we state next.

Theorem 4.1. Let, for an n-population Z, Π be the EM law for the PEM model, and $F^{(n)}$ and F be given as in (3.10). Moreover, let $Q^{(n)}$ be the associated pair-type process, and $V^{(n)}$ be defined as in (4.1), where Q is the fluid limit. Then, given the condition in (4.2), for any $T \in [0,\infty)$, $V^{(n)} \Rightarrow V$ in $D([0,T),\mathcal{M}^{k\times k}(\mathbb{R}))$ as $n \to \infty$, where the limiting process $V(t) = \sum_{i,j\in[k]} I^{ij} V_{ij}(t)$ taking values in $\mathcal{M}^{k\times k}(\mathbb{R})$, is given by the set of equations

$$V_{ij}(t) = W_{ij}(Q_{ij}(t)) + \sum_{i',j'\in[k]} \int_0^t \frac{\partial F_{ij}(Q(s))}{\partial M_{i'j'}} V_{i'j'}(s)ds, \qquad i,j \in [k], \tag{4.4}$$

where $\{W_{ij}\}_{i,j\in[k]}$ are independent standard Brownian motions.

Due to the different nature of the two proofs each will be presented within a separate subsection, and both can be read independently of each other. We also choose not to list all of the preliminary results and definitions here at the beginning, but instead, mention some of them whenever the need arises. The first two are technical lemmas regarding how the EM law of the PEM model (see Definition 3.1), relates to the nature of the functions $F^{(n)}$ and F given in (3.10). They will be stated without proofs, although the curious reader is referred to [6, Section 2.1] for more details.

Lemma 4.2. [6, Lemma 2.1] Let $\mathcal{E}_n$, $\mathcal{E}$, $F^{(n)}$, and F be given as in displays (3.8) to (3.10), and Π be the EM law. Moreover, let $n \in \mathbb{N}$ and $i, j, i', j' \in [k]$. Then the following hold.

a) For every $M \in \mathcal{E}_n$ and $M' \in \mathcal{E}$

$$0 \le F_{ij}^{(n)}(M) \le \pi_{ij}, \quad \text{and} \quad 0 \le F_{ij}(M') \le \pi_{ij}.$$

b) For $M \in \mathcal{M}^{k\times k}([0,\infty))$ with $M_{tot} < 1$,

$$\frac{\partial F_{ij}(M)}{\partial M_{i'j'}} = \pi_{ij}\left[\left(\frac{x_i - M_{i,\cdot}}{1-M_{tot}}\right)\left(\frac{y_j - M_{\cdot,j}}{1-M_{tot}}\right) - \left(\frac{x_i - M_{i,\cdot}}{1-M_{tot}}\right)\delta_{jj'} - \left(\frac{y_j - M_{\cdot,j}}{1-M_{tot}}\right)\delta_{ii'}\right],$$

where δ_{ij} denotes the Kronecker delta function. In particular,

$$\left|\frac{\partial F_{ij}(M)}{\partial M_{i'j'}}\right| \le \begin{cases} \pi_{ij} & \text{if } M_{tot} < 1 \text{ and } M \in \mathcal{E}, \\ \frac{3\pi_{ij}}{(1-M_{tot})^2} & \text{if } M_{tot} < 1 \text{ and } M \notin \mathcal{E}. \end{cases}$$

c) For every $M \in \mathcal{E}_n$ with $M_{tot} < 1$,

$$\left|F_{ij}^{(n)}(M) - F_{ij}(M)\right| \le \begin{cases} \pi_{ij}\left(\left|n^{-1}x_i^{(n)} - x_i\right| + \left|n^{-1}y_j^{(n)} - y_j\right|\right) & \text{if } M \in \mathcal{E}_n \cap \mathcal{E}, \\ \pi_{ij}\left(\left|n^{-1}x_i^{(n)} - x_i\right| + \frac{1}{1-M_{tot}}\left|n^{-1}y_j^{(n)} - y_j\right|\right) & \text{if } M \in \mathcal{E}_n \setminus \mathcal{E}. \end{cases}$$

Lemma 4.3. [6, Lemma 2.2] For any $T \in [0, \infty)$ and $c > \bar{\pi} := |\Pi|$,

$$\mathbb{P}\Big(1 - n^{-1} Q^{(n)}_{tot}(T) \geq e^{-cT} \text{ for sufficiently large } n\Big) = 1.$$

Remark 4.4. Under assumption (3.6), part (c) in Lemma 4.2 originally states an error bound of $o(1)$. The stronger assumption (4.2) however, yields an error bound of $o(n^{-1/2})$. Lemma 4.3 on the other hand, when combined with Lemma 4.2(b) ensures that

$$\left| \frac{\partial F_{ij}(n^{-1} Q^{(n)}(t))}{\partial M_{i'j'}} \right| \leq \frac{3\pi_{ij}}{(1 - n^{-1} Q^{(n)}_{tot}(t))^2} \leq 3\bar{\pi} e^{2cT}$$

for sufficiently large n on a subset Ω_T with $\mathbb{P}(\Omega_T) = 1$. Thus almost surely, for sufficiently large n, $\{n^{-1} Q^{(n)}(t) : t \leq T\}$ stays in a neighborhood of $\{Q(t) : t \leq T\}$ where F can be taken as Lipschitz continuous.

4.2 Proof by a coupling method

By using the drift function $F^{(n)}$ and the collection $\{J_{ij}(t)\}_{i,j \in [k]}$, we have successfully defined $Q^{(n)}(t)$ in (4.3) for any $n \in \mathbb{N}$. Using the function F (given in (3.10)) in a completely analogous way, we now introduce the family of $\mathcal{E}_n$-valued processes $\{\widehat{Q}^{(n)}(t)\}_{n \in \mathbb{N}}$ where, for each n, $\widehat{Q}^{(n)}(t)$ is the solution of the equation

$$\widehat{Q}^{(n)}(t) = \sum_{i,j \in [k]} I^{ij} J_{ij}\Big(n \int_0^t F_{ij}\big(n^{-1} \widehat{Q}^{(n)}(s)\big) ds\Big). \tag{4.5}$$

Clearly, the newly defined family acts far more conveniently with respect to the density-dependent framework presented in Chapter 2. Indeed, $\widehat{Q}^{(n)}$ satisfies the criteria in Definition 2.1, and we have the following result.

Theorem 4.5. Define the process $\widehat{V}^{(n)}(t)$ as

$$\widehat{V}^{(n)}(t) := \sqrt{n}(n^{-1} \widehat{Q}^{(n)}(t) - Q(t)),$$

where $Q(t)$ is the solution of (3.11). Then $\widehat{V}^{(n)} \Rightarrow V$ in $D([0,T), \mathcal{M}^{k \times k}(\mathbb{R}))$, where the limiting process $V(t) = \sum_{i,j \in [k]} I^{ij} V_{ij}(t)$ taking values in $\mathcal{M}^{k \times k}(\mathbb{R})$, is the solution of the set of equations in (4.4).

Proof. By Lemma 4.2(a),(b) and Lemma 4.3, F is bounded, and for a sufficiently large n, Lipschitz continuous in a neighborhood of $Q(t)$ (see Remark 4.4). We then have, by Theorem 2.3,

$$\lim_{n\to\infty} \sup_{t\le T} |\widehat{Q}^{(n)}(t) - Q(t)| = 0, \qquad \mathbb{P}\text{-a.s.}$$

for any $T \in [0,\infty)$. Moreover, for sufficiently large n, ∂F is also continuous by the same lemmas, and since the equations in (4.4) are the same as the one in (2.8), but in matrix form (see the discussion two paragraphs above Theorem 4.1) with $v_0 = 0$, the conclusion follows from Theorem 2.4. $\square$

Since we have coupled $\widehat{V}^{(n)}$ and $V^{(n)}$ on the same sample space, if we prove that for an arbitrary fixed $\varphi > 0$, the set on which they differ more than φ gets arbitrarily small in probability as n gets larger, then by Slutsky's theorem [4, Corollary 3.3.3] both processes have the same distributional limit. This turns out to be case if we look at finite intervals.

Let then, for an arbitrary $T \in [0,\infty)$, $V^{(n)}(t)$ be defined only for $t \in [0,T]$. Observe that we can rewrite $V^{(n)} = A^{(n)} + \widehat{V}^{(n)}$, where for any $i,j \in [k]$,

$$\begin{aligned} &A_{ij}^{(n)}(t) := \\ &:= n^{-1/2}\Big(J_{ij}\Big(n\int_0^t F_{ij}^{(n)}\big(n^{-1}Q^{(n)}(s)\big)ds\Big) - J_{ij}\Big(n\int_0^t F_{ij}\big(n^{-1}\widehat{Q}^{(n)}(s)\big)ds\Big)\Big). \end{aligned} \tag{4.6}$$

The following lemma plays a key role in our ultimate goal of showing that for any $\varphi > 0$, we have $\mathbb{P}(\sup_{t\le T}|A^{(n)}(t)| > \varphi) \to 0$.

Lemma 4.6. Let ε and T^* be given positive reals. Then $\exists\delta > 0$ dependent on ε, such that for every $i,j \in [k]$,

$$\mathbb{P}\Big(\sup_{0\le t\le nT^*} n^{-1/2}\Big|J_{ij}\big(t+\delta\sqrt{n}\big) - J_{ij}\big(t\big)\Big| > \varepsilon\Big) \to 0, \qquad \text{as } n\to\infty.$$

Proof. Let $k_n := \delta^{-1}T^*\sqrt{n}$ and let $\lceil k_n\rceil := \inf\{k \in \mathbb{N} : k \ge k_n\}$. Observe that $[0,nT^*]$ can be covered with the elements of

$$\mathbb{K}_n := \big\{[\delta\sqrt{n}m, \delta\sqrt{n}(m+1)] : m = 0,1,...,\lceil k_n\rceil\big\}.$$

Next, notice that for an arbitrary $t_0 \in [0, nT^*]$, the subinterval $[t_0, t_0+\delta\sqrt{n}]$ is covered by at most two neighboring intervals in $\mathbb{K}_n$. Therefore at least a half of the difference $J_{ij}\big(t_0 + \delta\sqrt{n}\big) - J_{ij}\big(t_0\big)$ belongs to one of these intervals. This implies the following:

$$\sup_{0\le t\le nT^*} \frac{1}{\sqrt{n}}\Big|J_{ij}\big(t+\delta\sqrt{n}\big) - J_{ij}\big(t\big)\Big| \le \sup_{0\le m\le\lceil k_n\rceil} \frac{2}{\sqrt{n}}\Big|J_{ij}\big(\delta\sqrt{n}(m+1)\big) - J_{ij}\big(\delta\sqrt{n}m\big)\Big|.$$

On the RHS, the $\lceil k_n\rceil + 1$ differences in the absolute brackets are in fact i.i.d. Poisson random variables, each with a parameter $\lambda_n = \delta\sqrt{n}$. Let us denote them by $X^{(m)}_{n,\delta}$. Then we have

$$\mathbb{P}\Big(\sup_{0\le m\le\lceil k_n\rceil} \frac{2}{\sqrt{n}} X^{(m)}_{n,\delta} > \varepsilon\Big) = 1 - \bigg(1 - \mathbb{P}\Big(X^{(1)}_{n,\delta} > \frac{\varepsilon\sqrt{n}}{2}\Big)\bigg)^{\lceil k_n\rceil+1}. \tag{4.7}$$

To bound the term on the right from above we use the following fact: given a Poisson random variable X with parameter λ_n, and reals $a, s > 0$, we know from the Markov inequality that

$$\mathbb{P}\big(X > a\big) = \mathbb{P}\big(e^{sX} > e^{sa}\big) \le \frac{\mathbb{E}[e^{sX}]}{e^{sa}} = \exp(-sa)\exp(\lambda_n(e^s - 1)) \le \exp(s(2\lambda_n - a))$$

for s picked small enough. Setting $X = X^{(1)}_{n,\delta}$, $a = \frac{\varepsilon\sqrt{n}}{2}$, $\lambda_n = \delta\sqrt{n}$, and picking $\delta < \frac{\varepsilon}{8}$, we see that

$$\mathbb{P}\bigg(X^{(1)}_{n,\delta} > \frac{\varepsilon\sqrt{n}}{2}\bigg) \le \exp\bigg(s\Big(2\delta\sqrt{n} - \frac{\varepsilon\sqrt{n}}{2}\Big)\bigg) = \exp\Big(\frac{s\sqrt{n}}{2}(4\delta - \varepsilon)\Big) \le \exp\Big(-\frac{s\varepsilon\sqrt{n}}{4}\Big)$$

which, when plugged into (4.7), yields

$$\begin{aligned}\mathbb{P}\Big(\sup_{0\le m\le\lceil k_n\rceil} \frac{2}{\sqrt{n}} X^{(m)}_{n,\delta} > \varepsilon\Big) &\le 1 - \Big(1 - \exp\Big(-\frac{1}{4}s\varepsilon\sqrt{n}\Big)\Big)^{\lceil\delta^{-1}T^*\sqrt{n}\rceil+1} \\ &\le 1 - \Big(1 - \exp\Big(-\frac{1}{4}s\varepsilon\sqrt{n}\Big)\Big)^{\lceil\varepsilon^{-1}8T^*\sqrt{n}\rceil+1} \to 0, \qquad \text{as } n\to\infty,\end{aligned}$$

due to the fact that $\lceil k_n\rceil + 1$ does not outgrow the exponential expression in the brackets. The result then follows. □

Back to investigating (4.6), notice that it can, in turn, be further decomposed into $A^{(n)}_{ij}(t) = B^{(n)}_{ij}(t) + C^{(n)}_{ij}(t)$, where

$$B^{(n)}_{ij}(t) := n^{-1/2}\Big(J_{ij}\Big(n\int_0^t F^{(n)}_{ij}\big(n^{-1}Q^{(n)}(s)\big)ds\Big) - J_{ij}\Big(n\int_0^t F_{ij}\big(n^{-1}Q^{(n)}(s)\big)ds\Big)\Big),$$

and

$$C_{ij}^{(n)}(t) := n^{-1/2}\Big(J_{ij}\Big(n\int_0^t F_{ij}\big(n^{-1}Q^{(n)}(s)\big)ds\Big) - J_{ij}\Big(n\int_0^t F_{ij}\big(n^{-1}\widehat{Q}^{(n)}(s)\big)ds\Big)\Big).$$

Our initial use of Lemma 4.6 is to control the term $B_{ij}^{(n)}(t)$, by proving its uniform convergence to 0 in probability. This is done in the following lemma. The second one, however, is in proving Lemma 4.8, where we control the term $A_{ij}^{(n)}(t)$ in the same way.

Lemma 4.7. Let $T \in [0,\infty)$ and $\varepsilon > 0$ be arbitrarily small. Then, given assumption (4.2), $\mathbb{P}\big(\sup_{t\leq T}|B^{(n)}(t)| > \varepsilon\big) \to 0$, as $n \to \infty$.

Proof. To prove the statement for $B^{(n)}(t)$, it suffices to show that it holds for any $B_{ij}^{(n)}(t)$. Under assumption (4.2), we know from Lemmas 4.2(c) and 4.3 that given $T > 0$ and $\delta > 0$, on a subset Ω_T with probability 1,

$$\big|F^{(n)}\big(n^{-1}Q^{(n)}(t)\big) - F(n^{-1}Q^{(n)}(t))\big| < n^{-1/2}\delta,$$

for sufficiently large n (see Remark 4.4). Next, we have $\int_0^t F_{ij}^{(n)}\big(n^{-1}Q^{(n)}(s)\big)ds \leq \bar{\pi}t$ due to Lemma 4.2(a), and by combining this with the previous statement we get that for large n,

$$\begin{aligned}\sup_{t\leq T}|B_{ij}^{(n)}(t)| &\leq \sup_{0\leq t\leq T} n^{-1/2}\Big|J_{ij}\Big(n\big(\bar{\pi}t + \frac{\delta T}{\sqrt{n}}\big)\Big) - J_{ij}\Big(n\bar{\pi}t\Big)\Big| \\ &= \sup_{0\leq t\leq n\bar{\pi}T} n^{-1/2}\Big|J_{ij}\big(t + \delta\sqrt{n}T\big) - J_{ij}\big(t\big)\Big|,\end{aligned}$$

by performing a change of variables. Then, by just plugging $\delta \leftarrow \delta T$ and $T^* \leftarrow \bar{\pi}T$ in Lemma 4.6 we obtain the wanted result. □

We proceed, as mentioned previously, to Lemma 4.8, whose proof contains most of the technical work needed for our main result, Theorem 4.1. Notice that we still have not proved any convergence result regarding $\sup_{t\leq T}|C^{(n)}(t)|$.

Lemma 4.8. Let $T \in [0,\infty)$. Then, under the assumption in (4.2), for any $\varphi > 0$, $\mathbb{P}\big(\sup_{t\leq T}|A^{(n)}(t)| > \varphi\big) \to 0$, as $n \to \infty$.

Proof. We know from Lemma 4.3 that given $T \in [0, \infty)$, there exist a positive integer valued N_T, with $\mathbb{P}(N_T < \infty) = 1$, such that for $n \geq N_T$, we can consider F to be Lipschitz continuous (see Remark 4.4) in a neighborhood of $\{Q(t) : t \leq T\}$. We use K to denote its Lipschitz constant. Consequently, on $\Omega_T := \{N_T < \infty\}$ we have the following upper bound for sufficiently large n and $t \leq T$:

$$n \int_0^t F_{ij}(n^{-1}Q^{(n)}(s))ds \leq n \int_0^t F_{ij}(n^{-1}\widehat{Q}^{(n)}(s))ds + \int_0^t K|Q^{(n)}(s) - \widehat{Q}^{(n)}(s)|ds. \quad (4.8)$$

Next, since the term $K \int_0^t |Q^{(n)}(t) - \widehat{Q}^{(n)}(t)|ds = K\sqrt{n} \int_0^t |A^{(n)}(s)|ds$ is of order $o(n)$ and $F_{ij} \leq \bar{\pi}$, given any small $\delta > 0$, for $0 \leq t < T$ the RHS in the above inequality can be bounded from above by

$$n \int_0^t F_{ij}\big(n^{-1}\widehat{Q}^{(n)}(s)\big)ds + K\sqrt{n} \int_0^t \big(\delta + |A^{(n)}(s)|\big)ds \leq n\bar{\pi}T + \delta\sqrt{n}KT \qquad (4.9)$$

for $n \geq L_T$, where $\mathbb{P}(N_T \leq L_T < \infty) = 1$. Moreover, there exists a random positive integer M_n such that

$$(M_n - 1)\delta T \leq \int_0^t |A^{(n)}(s)|ds \leq M_n \delta T. \qquad (4.10)$$

Let $\varepsilon > 0$ and $\eta > 0$ be arbitrarily small. By plugging $T^* \leftarrow \bar{\pi}T$ in Lemma 4.6 and then picking $\delta < \frac{\varepsilon}{8KT}$, we have

$$\mathbb{P}\Big(\sup_{0 \leq t \leq n\bar{\pi}T} n^{-1/2} \Big| J_{ij}\big(t + \delta\sqrt{n}KT\big) - J_{ij}\big(t\big) \Big| \leq \varepsilon \Big) > 1 - \eta$$

for $n \geq N_{\varepsilon,\eta}$, where $N_{\varepsilon,\eta} \in \mathbb{N}$. Let us denote the subset on which this condition is satisfied as $\Omega_{\varepsilon,\eta}$.

We switch temporarily to the less cumbersome notation $\beta_{ij}(t) := nF_{ij}\big(n^{-1}\widehat{Q}^{(n)}(t)\big)$. Then finally, due to the paths of J_{ij} being non-decreasing, for $n \geq \max(N_{\varepsilon,\eta}, L_T)$, we have on the set $\Omega_{\varepsilon,\eta} \cap \Omega_T$ that the term $|C_{ij}^{(n)}(t)|$ is bounded from above by

$$\begin{aligned}
& n^{-1/2} \Big| J_{ij}\Big(\int_0^t \beta_{ij}(s)ds + K \int_0^t |Q^{(n)}(s) - \widehat{Q}^{(n)}(s)|ds \Big) - J_{ij}\Big(\int_0^t \beta_{ij}(s)ds \Big) \Big| \\
& \leq n^{-1/2} \Big| J_{ij}\Big(\int_0^t \beta_{ij}(s)ds + K\sqrt{n} \int_0^t \big(\delta + |A^{(n)}(s)|\big)ds \Big) - J_{ij}\Big(\int_0^t \beta_{ij}(s)ds \Big) \Big| \\
& \leq n^{-1/2} \Big| J_{ij}\Big(\int_0^t \beta_{ij}(s)ds + K\sqrt{n}(M_n + 1)\delta T \Big) - J_{ij}\Big(\int_0^t \beta_{ij}(s)ds \Big) \Big| \\
& \leq n^{-1/2} \sum_{m=0}^{M_n} \Big| J_{ij}\Big(\int_0^t \beta_{ij}(s)ds + K\sqrt{n}(m+1)\delta T \Big) - J_{ij}\Big(\int_0^t \beta_{ij}(s)ds + K\sqrt{n}m\delta T \Big) \Big|
\end{aligned}$$

$$\leq (M_n + 1)\varepsilon \leq 2M_n\varepsilon \leq 2\varepsilon + 2\varepsilon(\delta T)^{-1} \int_0^t |A^{(n)}(s)| ds$$

where the first three upper bounds come from (4.8) - (4.10), the fourth is just the triangle inequality done $m+1$ times, after which we use (4.9) and finally (4.10). Then on the set $\Omega_{\varepsilon,\eta} \cap \Omega_T$ we obtain (by generalizing over $i, j \in [k]$),

$$|A^{(n)}(t)| \leq |B^{(n)}(t)| + 2\varepsilon + 2\varepsilon(\delta T)^{-1} \int_0^t |A^{(n)}(s)| ds$$

which, by Gronwall's inequality (Lemma 2.2) and Lemma 4.7 yields

$$\sup_{0 \leq t \leq T} |A^{(n)}(t)| \leq \sup_{0 \leq t \leq T} \left(|B^{(n)}(t)| + 2\varepsilon\right) \exp(2\varepsilon\delta^{-1}) \xrightarrow{p} 2\varepsilon \exp(2\varepsilon\delta^{-1})$$

as $n \to \infty$. Since we have uniform convergence to an abritrarily small value (dependent on ε), on a set with arbitrarily large probability (dependent on η), the lemma follows. □

Proof of Theorem 4.1. Having proved that $\sup_{0 \leq t \leq T} |A^{(n)}(t)| \xrightarrow{p} 0$ in Lemma 4.8, the rest is just a straightforward application of Slutsky's theorem [4, Corollary 3.3.3] and Theorem 4.5, from which the distributional convergence follows.

4.3 An adaptation of the standard proof

This approach utilizes notation and techniques that work best in a Euclidean space setup, as previously presented in [4, Theorem 11.2.3] and [1, Theorem 5.4]. For this purpose, we consider the random variables $V^{(n)}(t)$ and $V(t)$ to be taking values in $\mathbb{R}^{k^2}$, instead of $\mathcal{M}^{k \times k}(\mathbb{R})$. The norm used will be the standard maximum norm $|\cdot|_\infty$, defined as $|M|_\infty := \max_{1 \leq i \leq k^2} |m_i|$ for $M \in \mathbb{R}^{k^2}$. However, as we have mentioned before (see the discussion before Theorem 4.1), we will avoid writing $|M|_\infty$, and write $|M|$ instead.

Keeping the above discussion in perspective, and recalling that $\{\ell_{ij}\}_{i,j \in [k]}$ is the $\mathbb{R}^{k^2}$ equivalent to the canonical base $\{I^{ij}\}_{i,j \in [k]}$ for the vector space $\mathcal{M}^{k \times k}(\mathbb{R})$, we let $Q^{(n)}(t) := \sum_{i,j \in [k]} \ell_{ij} Q_{ij}^{(n)}(t)$ be the $\mathbb{R}^{k^2}$-valued analogue of the pair-type process. Similarly, let $V^{(n)}(t) := \sum_{i,j \in [k]} \ell_{ij} V_{ij}^{(n)}(t)$ be defined for $t \in [0, T]$, where $T < \infty$.

Within this new setup, the functions $F^{(n)}$, and F, given in (3.10) will still behave in the same way, except for the obvious relabeling of their corresponding domains and ranges, as subsets of $\mathbb{R}^{k^2}$. Then, observe that for each $i, j \in [k]$, we can rewrite $V_{ij}^{(n)}(t)$ as

$$n^{-1/2}\Big(J_{ij}\Big(n\int_0^t F_{ij}^{(n)}\big(n^{-1}Q^{(n)}(s)\big)ds\Big) - n\int_0^t F_{ij}^{(n)}\big(n^{-1}Q^{(n)}(s)\big)ds\Big) \tag{4.11}$$

$$+ n^{1/2}\Big(\int_0^t F_{ij}^{(n)}\big(n^{-1}Q^{(n)}(s)\big)ds - \int_0^t F_{ij}\big(n^{-1}Q^{(n)}(s)\big)ds\Big) \tag{4.12}$$

$$+ n^{1/2}\Big(\int_0^t F_{ij}\big(n^{-1}Q^{(n)}(s)\big)ds - \int_0^t F_{ij}(Q(s))ds\Big). \tag{4.13}$$

We shall start with the term in (4.12) which we will denote as $\alpha_{ij}^{(n)}(t)$. Since $t \in [0, T]$, due to Lemmas 4.2(c) and 4.3 we see that $\alpha_{ij}^{(n)}(t)$ is almost surely bounded from above by

$$\widehat{A}_{ij}^{(n)}(t) := \pi_{ij} t\left[\left|\frac{x_i^{(n)} - nx_i}{\sqrt{n}}\right| + e^{cT}\left|\frac{y_j^{(n)} - ny_j}{\sqrt{n}}\right|\right] \le \widehat{A}_{ij}^{(n)}(T)$$

for sufficiently large n, which given assumption (4.2), yields $\sup_{0\le t\le T} |\alpha_{ij}^{(n)}(t)| \xrightarrow{a.s.} 0$.

Next, we turn to the expression in (4.11) which we will denote as

$$U_{ij}^{(n)}(t) := W_{ij}^{(n)}\Big(\int_0^t F_{ij}^{(n)}\big(n^{-1}Q^{(n)}(s)\big)ds\Big),$$

where $W_{ij}^{(n)}(t) := n^{-1/2}(J_{ij}(nt) - nt)$. Since by the functional CLT, $W_{ij}^{(n)} \Rightarrow W_{ij}$ in $D([0, T), \mathbb{R})$ where $W_{ij}(t)$ is a standard Brownian motion, we expect to obtain an analogous convergence result for $U_{ij}^{(n)}$, too. Naturally, our limit candidate is

$$U_{ij}(t) := W_{ij}\Big(\int_0^t F_{ij}(Q(s))ds\Big) = W_{ij}\big(Q_{ij}(t)\big),$$

as $F^{(n)}$, $n^{-1}Q^{(n)}$ and $W_{ij}^{(n)}$ converge (each in its appropriate sense) to F, Q, and W_{ij}, respectively. Since $\mathbb{R}^{k^2}$ is separable, to prove this it suffices (by [4, Theorem 3.7.8]) to show the following two things:

- the sequence of processes $\{U_{ij}^{(n)}\}_{n\in\mathbb{N}}$ is relatively compact in $D([0, T), \mathbb{R})$;
- for any $0 < t_1, t_2, ..., t_d < T$, $(U_{ij}^{(n)}(t_1), ..., U_{ij}^{(n)}(t_d)) \Rightarrow (U_{ij}(t_1), ..., U_{ij}(t_d))$, i.e. we have a convergence in finite-dimensional distributions.

The next lemma proves the former of the two, i.e. it consists of showing two properties $\{U_{ij}^{(n)}\}_{n\in\mathbb{N}}$ has, that are equivalent to relative compactness (see [4, Corollary 3.7.4]). We hereby set the necessary terminology and definitions we will be using in the proof of the lemma, or more specifically, while showing the latter of the two properties..

Definition 4.9. A set of points $0 = t_0 < t_1 < .. < t_d = T$ is called δ-sparse if, for $\delta > 0$, we have $\min_{1\le i\le d} |t_i - t_{i-1}| > \delta$ (see [2, page 122]).

Definition 4.10. For positive δ, T, the modulus of continuity w' is defined as

$$w'(x,\delta,T) := \inf_{\{t_i\}} \max_i \sup_{s,t\in[t_{i-1},t_i)} |x(t) - x(s)|,$$

where $x(\cdot) \in D([0,T),\mathbb{R})$, and the infimum is taken over all possible δ-sparse partitions of $[0,T]$ (see [4, page 122] or [2, page 122]).

Lemma 4.11. For any $i, j \in [k]$, $\{U_{ij}^{(n)}\}_{n\in\mathbb{N}}$ satisfies the following two conditions:

a) for any $\eta > 0$, and nonnegative $t \in \mathbb{Q}$, there exists a compact set $\Gamma_{\eta,t} \subset \mathbb{R}$, such that

$$\liminf_{n\to\infty} \mathbb{P}\Big(U_{ij}^{(n)}(t) \in \Gamma_{\eta,t}^{\eta}\Big) \ge 1 - \eta \tag{4.14}$$

where for a given set $\Gamma \subset \mathbb{R}$, $\Gamma^{\eta} := \{x \in \mathbb{R} : \inf_{y\in\Gamma} |x - y| < \eta\}$;

b) for every $\eta > 0$ and $T > 0$, there exists $\delta > 0$ such that

$$\limsup_{n\to\infty} \mathbb{P}\Big(w'(U_{ij}^{(n)},\delta,T) \ge \eta\Big) \le \eta. \tag{4.15}$$

In other words, the sequence $\{U_{ij}^{(n)}\}_{n\in\mathbb{N}}$ is relatively compact by [4, Corollary 3.7.4].

Proof. a) Since by Lemma 4.2(a), for all n and $i, j \in [k]$, $F_{ij}^{(n)} \le \pi_{ij} \le \bar{\pi}$, where $\bar{\pi} := \max_{i,j\in[k]} |\pi_{ij}|$, by utilizing Doob's Martingale inequality we have the following:

$$\mathbb{P}\Big(\sup_{t\le T}\Big|W_{ij}^{(n)}\Big(\int_0^t F_{ij}^{(n)}(n^{-1}Q^{(n)}(s))ds\Big)\Big| > a\Big) \le \mathbb{P}\Big(\sup_{t\le T}\big|W_{ij}^{(n)}(\bar{\pi}t)\big| > a\Big)$$

$$\le \frac{\mathbb{E}[\big|W_{ij}^{(n)}(\bar{\pi}T)\big|^2]}{a^2} = \frac{\bar{\pi}T}{a^2} \to 0 \text{ as } a \to \infty, \text{ and } n \to \infty.$$

Therefore, for each $\eta > 0$ we can pick a sufficiently large a, so that we have

$$\mathbb{P}\Big(\sup_{t\leq T}\big|U_{ij}^{(n)}(t)\big| \leq a\Big) \geq 1-\eta$$

for all n. In other words $\{U_{ij}^{(n)}\}_{n\in\mathbb{N}}$ satisfies condition (4.14) if we pick $\Gamma_{\eta,t} = [-a, a]$.

b) By the functional CLT, $W_{ij}^{(n)} \Rightarrow W_{ij}$ in $D([0,T),\mathbb{R})$, therefore the sequence $\{W_{ij}^{(n)}\}_{n\in\mathbb{N}}$ is relatively compact. To see this note that since $\mathbb{R}$ is Polish (which means complete and separable), by [4, Theorem, 3.5.6] and [4, Theorem 3.1.7], so are the space $D([0,T),\mathbb{R})$ under the Skohorod $\mathcal{J}_1$ topology, and the space of probability measures on $D([0,T),\mathbb{R})$ equipped with the Prohorov metric ρ^* (see [4, page 96] for details). Then, $W_{ij}^{(n)} \Rightarrow W_{ij}$ implies $\rho^*(W_{ij}^{(n)}, W_{ij}) \to 0$ by [4, Theorem, 3.3.1]. Finally, this means that $\{W_{ij}^{(n)}\}_{n\in\mathbb{N}}$ is relatively compact by [4, Lemma 3.2.1] followed by [4, Theorem 3.2.2(b)(c)].

The relative compactness of $\{W_{ij}^{(n)}\}_{n\in\mathbb{N}}$ simplifies the proof, since it shortens otherwise lengthy calculations. Firstly note that for any $\omega \in \Omega$ we have

$$\begin{aligned}
&w'(U_{ij}^{(n)}(\omega,\cdot),\delta,T)\\
&= w'\Big(W_{ij}^{(n)}\Big(\int_0^{\cdot} F_{ij}^{(n)}(n^{-1}Q^{(n)}(\omega,s))ds\Big),\delta,T\Big) \leq w'(W_{ij}^{(n)}(\omega,\cdot),\bar{\pi}\delta,\bar{\pi}T)
\end{aligned}$$

because every $\bar{\pi}\delta$-sparse partition on the RHS is also included in the δ-sparse partitions in the LHS. But since $W_{ij}^{(n)}$ is relatively compact, we can for any $\eta > 0$, $T > 0$, find a $\delta > 0$ such that $\limsup_{n\to\infty}\mathbb{P}(w'(W_{ij}^{(n)},\delta,T) \geq \eta) \leq \eta$ by condition (7.14) in [4, Corollary 3.7.4]. This fact, combined with the inequality above finishes our proof of part b) and therefore, concludes the lemma. □

We end the second step towards our main goal by proving that $U^{(n)} \Rightarrow U$ in $D([0,T),\mathbb{R}^{k^2})$ in the following lemma.

Lemma 4.12. Let $U_{ij}^{(n)}(t)$ and $U_{ij}(t)$, be defined for $0 \leq t \leq T < \infty$. Then

a) for fixed t, $U_{ij}^{(n)}(t) \Rightarrow U_{ij}(t)$

b) for any $0 < t_1, t_2, ..., t_d < T$, $(U_{ij}^{(n)}(t_1), ..., U_{ij}^{(n)}(t_d)) \Rightarrow (U_{ij}(t_1), ..., U_{ij}(t_d))$

c) $U_{ij}^{(n)} \Rightarrow U_{ij}$, and ultimately $U^{(n)} \Rightarrow U$.

Proof. We begin by stating the fact that since, for any T, $W_{ij}^{(n)}(\cdot) \Rightarrow W_{ij}(\cdot)$ in $D([0,T),\mathbb{R})$ by the functional CLT, then $W_{ij}^{(n)}(Q_{ij}(\cdot)) \Rightarrow W_{ij}(Q_{ij}(\cdot))$ is also true, (as each $Q_{ij}(t) : [0,T] \to [0, Q_{ij}(T)]$, is continuously differentiable with positive derivative, therefore a bijection.

a) Consider, for a fixed t, the random variable

$$\Lambda_{ij}^{(n)}(t) := \int_0^t F_{ij}^{(n)}(n^{-1}Q^{(n)}(s))ds - \int_0^t F_{ij}(Q(s))ds.$$

By using details from the proof in [6, Theorem 2.1], we are able to write that $\mathbb{P}$-a.s.,

$$\begin{aligned}\sup_{t\le T} |\Lambda_{ij}^{(n)}(t)| \le \pi_{ij}T\Big(\big|n^{-1}x_i^{(n)} - x_i\big| + e^{cT}\big|n^{-1}y_j^{(n)} - y_j\big|\Big)\\ + 3k^2\pi_{ij}e^{2cT}\int_0^T \Big|n^{-1}Q^{(n)}(s) - Q(s)\Big|ds \to 0.\end{aligned}$$

Then, for a fixed $t \in [0,T]$, given any $\delta > 0, \eta > 0$, we can find a $N_{\delta,\eta}$, such that for $n \ge N_{\delta,\eta}$, we have $\mathbb{P}(\Omega_{\delta,\eta}) < \eta$, where $\Omega_{\delta,\eta} := \{\omega : |\Lambda_{ij}^{(n)}(t)| > \delta\}$. Then on this set $\Omega_{\delta,\eta}$, we have

$$\begin{aligned}U_{ij}^{(n)}(t) &= W_{ij}^{(n)}\Big(\int_0^t F_{ij}^{(n)}\big(n^{-1}Q^{(n)}(s)\big)ds\Big) = W_{ij}^{(n)}\Big(\int_0^t F_{ij}(Q(s))ds + \Lambda_{ij}^{(n)}(t)\Big)\\ &= \begin{cases} W_{ij}^{(n)}\big(Q_{ij}(t)\big) + \widehat{W}_{ij}^{(n)}\big(\Lambda_{ij}^{(n)}(t)\big) & \text{if } \Lambda_{ij}^{(n)}(t) > 0,\\ W_{ij}^{(n)}\big(Q_{ij}(t) - \delta\big) + \widetilde{W}_{ij}^{(n)}\big(\delta + \Lambda_{ij}^{(n)}(t)\big) & \text{if } \Lambda_{ij}^{(n)}(t) < 0,\end{cases}\end{aligned}$$

where $\widehat{W}_{ij}^{(n)}$ and $\widetilde{W}_{ij}^{(n)}$ are processes i.i.d. to $W_{ij}^{(n)}$. We were able to do this since $W_{ij}^{(n)}$ is a process with stationary and independent increments. In the case where $\Lambda_{ij}^{(n)}(t)$ is positive, it gets arbitrarily small with arbitrarily large probability (as we increase n), therefore we expect $\widehat{W}_{ij}^{(n)}\big(\Lambda_{ij}^{(n)}(t)\big)$ to have a negligible contribution. We make this argument rigorous now.

Let $\varepsilon > 0$ be arbitrarily small. Then, by our arguments above, for $n \ge N_{\varepsilon^3,\varepsilon}$,

$\mathbb{P}(|\Lambda_{ij}^{(n)}(t)| > \varepsilon^3) < \varepsilon$ holds, and we have

$$\mathbb{P}\Big(\Big|\widehat{W}_{ij}^{(n)}\big(\Lambda_{ij}^{(n)}(t)\big)\Big| > \varepsilon\Big) \leq \mathbb{P}\Big(\mathbb{1}_{\{|\Lambda_{ij}^{(n)}(t)|<\varepsilon^3\}}\big|\widehat{W}_{ij}^{(n)}\big(\Lambda_{ij}^{(n)}(t)\big)\big| > \varepsilon\Big) + \varepsilon$$
$$\leq \mathbb{P}\Big(\sup_{t\leq\varepsilon^3}\big|\widehat{W}_{ij}^{(n)}(t)\big| > \varepsilon\Big) + \varepsilon \leq \frac{\mathbb{E}\big[\big|\widehat{W}_{ij}^{(n)}(\varepsilon^3)\big|^2\big]}{\varepsilon^2} + \varepsilon = \frac{\mathbb{E}\Big[|J_{ij}\big(n\varepsilon^3\big) - n\varepsilon^3|^2\Big]}{n\varepsilon^2} + \varepsilon$$
$$= \frac{Var(J_{ij}(n\varepsilon^3))}{n\varepsilon^2} = \frac{n\varepsilon^3}{n\varepsilon^2} + \varepsilon = 2\varepsilon$$

or, in other words, $\widehat{W}_{ij}^{(n)}\big(\Lambda_{ij}^{(n)}(t)\big)$ converges to 0 in probability as $n \to \infty$. Analogous arguments are used for the case where $\Lambda_{ij}^{(n)}(t) < 0$. Indeed, we can rewrite

$$U_{ij}^{(n)}(t) = W_{ij}^{(n)}\big(Q_{ij}(t)\big) + \big(W_{ij}^{(n)}\big(Q_{ij}(t) - \delta\big) - W_{ij}^{(n)}\big(Q_{ij}(t)\big)\big) + \widetilde{W}_{ij}^{(n)}\big(\delta + \Lambda_{ij}^{(n)}(t)\big).$$

Here, the fact that δ can be picked arbitrarily small, tells us via Slutsky's theorem [4, Corollary 3.3.3] that $U_{ij}^{(n)}(t) \Rightarrow W_{ij}\big(Q_{ij}(t)\big)$ for a *fixed* $t \in [0, T]$, which concludes part a) of the proof.

b) Let $0 < t_1 < ... < t_d < T$. Notice that the d-dimensional, real valued vector $(U_{ij}^{(n)}(t_1), ..., U_{ij}^{(n)}(t_d))^T$ can be written in the form $\overrightarrow{U}_{ij}^{(n)} = \overrightarrow{W}_{ij}^{(n)} + \overrightarrow{\Psi}_{ij}^{(n)}$, or more specifically,

$$\begin{pmatrix} U_{ij}^{(n)}(t_1) \\ U_{ij}^{(n)}(t_2) \\ \vdots \\ U_{ij}^{(n)}(t_d) \end{pmatrix} = \begin{pmatrix} W_{ij}^{(n)}\big(Q(t_1)\big) \\ W_{ij}^{(n)}\big(Q(t_2)\big) \\ \vdots \\ W_{ij}^{(n)}\big(Q(t_d)\big) \end{pmatrix} + \begin{pmatrix} \Psi_{ij}^{(n)}(t_1) \\ \Psi_{ij}^{(n)}(t_2) \\ \vdots \\ \Psi_{ij}^{(n)}(t_d) \end{pmatrix}, \tag{4.16}$$

where $\overrightarrow{W}_{ij}^{(n)}$ can be written as

$$\begin{pmatrix} W_{ij}^{(n)}\big(Q(t_1)\big) \\ W_{ij}^{(n)}\big(Q(t_2)\big) \\ \vdots \\ W_{ij}^{(n)}\big(Q(t_d)\big) \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ \vdots & \ddots & 1 & 1 \\ 0 & 1 & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} W_{ij}^{(n)}\big(Q(t_d)\big) - W_{ij}^{(n)}\big(Q(t_{d-1})\big) \\ \vdots \\ W_{ij}^{(n)}\big(Q(t_2)\big) - W_{ij}^{(n)}\big(Q(t_1)\big) \\ W_{ij}^{(n)}\big(Q(t_1)\big) \end{pmatrix}, \tag{4.17}$$

and $\Psi_{ij}^{(n)}(t) :=$

$$:= \begin{cases} \widehat{W}_{ij}^{(n)}\big(\Lambda_{ij}^{(n)}(t)\big) & \text{if } \Lambda_{ij}^{(n)}(t) > 0, \\ \big(W_{ij}^{(n)}\big(Q_{ij}(t)-\delta\big) - W_{ij}^{(n)}\big(Q_{ij}(t)\big)\big) + \widetilde{W}_{ij}^{(n)}\big(\delta + \Lambda_{ij}^{(n)}(t)\big) & \text{if } \Lambda_{ij}^{(n)}(t) < 0. \end{cases} \tag{4.18}$$

In (4.17), the components of the vector on the right are independent random variables (due to $W_{ij}^{(n)}$ having stationary and independent increments). Consequently, this vector converges in distribution to $\big(W_{ij}\big(Q(t_d)\big) - W_{ij}\big(Q(t_{d-1})\big), ..., W_{ij}\big(Q(t_1)\big)\big)^T$, yielding $\overrightarrow{W}_{ij}^{(n)} \Rightarrow \overrightarrow{W}_{ij}$. Moreover in (4.18), even though we have not explicitly stated this through notation, the processes $\widetilde{W}_{ij}^{(n)}$ and $\widehat{W}_{ij}^{(n)}$ are distinct (and mutually independent) for every $t = t_1, ..., t_d$.

Finally, similarly to what we did in part a), we can pick an arbitrarily small $\delta > 0$ and an arbitrarily large n, to make $\overrightarrow{\Psi}_{ij}^{(n)}$ arbitrarily close to 0 in probability. Then the result follows by an application of Slutsky's theorem.

c) Since we have just obtained convergence in finite-dimensional distributions in part b) and relative compactness of $\{U_{ij}^{(n)}\}_{n\in\mathbb{N}}$ in Lemma 4.11, by [4, Theorem 3.7.8] we finally have $U_{ij}^{(n)} \Rightarrow U_{ij}$. Moreover, because by the multi-dimensional functional CLT [4, Theorem 5.1.2] we have $W^{(n)} := \sum_{ij} \ell_{ij} W_{ij}^{(n)} \Rightarrow \sum_{ij} \ell_{ij} W_{ij} = W$ in $D([0,T), \mathbb{R}^{k^2})$, $\{W^{(n)}\}_{n\in\mathbb{N}}$ is relatively compact (by the same arguments we used in the proof of Lemma 4.11), and by repeating our steps in parts a) and b) (i.e. by approximating, taking limits and then using independence of the components) we get $U^{(n)} \Rightarrow U$ in $D([0,T), \mathbb{R}^{k^2})$. □

By obtaining two convergence results for (4.11) and (4.12) we completed two major steps towards our final goal, i.e. proving Theorem 4.1. For the third one, we turn our attention to expression (4.13). We define $\beta^{(n)}(t)$ as

$$\beta^{(n)}(t) := \sqrt{n} \int_0^t \Big(F\big(n^{-1}Q^{(n)}(s)\big) - F\big(Q(s)\big) - n^{-1/2} \partial F\big(Q(s)\big) V^{(n)}(s) \Big) ds. \tag{4.19}$$

Next, ∂F (explicitly given in Lemma 4.2(b)) is continuous in a neighborhood around $\{Q(t) : t \leq T\}$, so on a set Ω_T with probability 1 and for a sufficiently large n, we can perform a first order Taylor approximation around $Q(t)$ to obtain

$$\begin{aligned}&\sqrt{n}\Big(F\big(n^{-1}Q^{(n)}(t)\big) - F\big(Q(t)\big)\Big)\\&= \sqrt{n}\partial F\big(Q(t)\big)\big(n^{-1}Q^{(n)}(t) - Q(t)\big) + \sqrt{n}O\big(\big|n^{-1}Q^{(n)}(t) - Q(t)\big|^2\big)\end{aligned}$$

which means that, for some finite $C > 0$, and n large enough, we have

$$\sqrt{n}\Big(F\big(n^{-1}Q^{(n)}(t)\big) - F\big(Q(t)\big)\Big) \leq \partial F\big(Q(t)\big)V^{(n)}(t) + C\big|V^{(n)}(t)\big|\big|n^{-1}Q^{(n)}(t) - Q(t)\big|$$

However, this yields an upper bound for $\beta^{(n)}(t)$. Namely, for $n \geq N$, where N is some random natural number, we now have

$$|\beta^{(n)}(t)| \leq C\int_0^t \big|V^{(n)}(s)\big|\big|n^{-1}Q^{(n)}(s) - Q(s)\big|ds. \tag{4.20}$$

It turns out that $\mathbb{P}$-a.s., $\sup_{t\leq T}|\beta^{(n)}(t)| \to 0$, as $n \to \infty$. We prove this in the following lemma.

Lemma 4.13. For $T \in [0,\infty)$ and $a \in \mathbb{R}$,

a) $V^{(n)}(t)$ is bounded in probability on $[0,T]$, i.e.

$$\lim_{n\to\infty}\lim_{a\to\infty}\mathbb{P}\Big(\sup_{t\leq T}|V^{(n)}(t)| > a\Big) = \lim_{a\to\infty}\lim_{n\to\infty}\mathbb{P}\Big(\sup_{t\leq T}|V^{(n)}(t)| > a\Big) = 0.$$

b) $\sup_{t\leq T}|\beta^{(n)}(t)| \to 0$, a.s. as $n \to \infty$.

Proof. a) We observe how $V_{ij}^{(n)}(t)$ was split into expressions (4.11), (4.12), and (4.13), and recall that by Lemmas 4.2(b) and 4.3, on a set Ω_T with probability 1, for sufficiently large n, F is Lipschitz continuous with some constant K. Due to this, we now have

$$\begin{aligned}\big|V_{ij}^{(n)}(t)\big| &\leq \Big|U_{ij}^{(n)}(t)\Big| + \big|\alpha_{ij}^{(n)}(t)\big| + n^{1/2}\Big|\int_0^t F_{ij}(n^{-1}Q^{(n)}(s))ds - \int_0^t F_{ij}(Q(s))ds\Big|\\&\leq \Big|U_{ij}^{(n)}(t)\Big| + \big|\alpha_{ij}^{(n)}(t)\big| + n^{1/2}K\int_0^t \Big|n^{-1}Q^{(n)}(s) - Q(s)\Big|ds\\&= \Big|U_{ij}^{(n)}(t)\Big| + \big|\alpha_{ij}^{(n)}(t)\big| + K\int_0^t \Big|V^{(n)}(s)\Big|ds.\end{aligned}$$

Then by taking maximum over $i, j \in [k]$, and Gronwall's inequality (Lemma 2.2) we have that on a set Ω_T with probability 1, for large n,

$$\sup_{t\leq T} |V^{(n)}(t)| \leq \sup_{t\leq T} \max_{i,j\in[k]} \Big(\Big|U_{ij}^{(n)}(t)\Big| + \Big|\alpha_{ij}^{(n)}(t)\Big| \Big) \exp(KT).$$

Since we have already proved that both $U^{(n)}(t)$ and $\alpha^{(n)}(t)$ are bounded in probability (see proof of Lemma 4.11(a) for $U^{(n)}(t)$ and the beginning of this section for $\alpha^{(n)}(t)$), the proof is done.

b) We know from part a) that almost surely, for sufficiently large n, the path $V_{ij}^{(n)}(t)$ is bounded from above by some finite number dependent only on ω. Then since $\lim_{n\to\infty} \sup_{t\leq T} \Big|n^{-1}Q^{(n)}(t) - Q(t)\Big| \to 0$ almost surely, we have by the bound in (4.20)

$$\mathbb{P}\Big(\omega : \sup_{t\leq T} |\beta^{(n)}(t)| \to 0, \text{ as } n \to \infty\Big) = 1,$$

and the claim is proved. □

The last two lemmas constitute a significant portion of the technical work used in our second proof of Theorem 4.1. However, because we are now working in a Euclidean framework, i.e. $Q^{(n)}, V^{(n)}$, and V are now vector-valued instead of matrix-valued, we need to also redefine our candidate limit process. V was initially defined to be the solution of the system of equations given in (4.4). When translated into a vector notation, which is more compact, these equations assume the following form:

$$V(t) = \sum_{i,j\in[k]} \ell_{ij} W_{ij}\big(Q_{ij}(t)\big) + \int_0^t \partial F\big(Q(s)\big)V(s)ds. \tag{4.21}$$

Here, we recall that ℓ_{ij} is the $(j-1)k+i$th canonical base unit vector of $\mathbb{R}^{k^2}$, and the $k^2 \times k^2$ matrix ∂F is as given in Lemma 4.2(b). Next, let $\Phi^*(t, s)$ be defined as the solution to the matrix equation

$$\frac{\partial}{\partial s}\Phi^*(t, s) = -\Phi^*(t, s)\partial F(Q(s)), \qquad \Phi^*(t, t) = I \tag{4.22}$$

(see [1, page 44]). What plays an important part in our second proof of Theorem 4.1, is that $V(t)$ can also be represented as

$$V(t) = U(t) + \int_0^t \Phi^*(t, s)\partial F(Q(s))U(s)ds. \tag{4.23}$$

To see this note that since $U(t) = V(t) - \int_0^t \partial F(Q(s))V(s)ds$, we have

$$\int_0^t \Phi^*(t,s)dU(s) = \int_0^t \Phi^*(t,s)dV(s) - \int_0^t \Phi^*(t,s)\partial F(Q(s))V(s)ds.$$

We see from the definition of $\Phi^*(t,s)$, that it is differentiable in s, and therefore locally of bounded variation. Consequently, we can use integration by parts on the RHS to get

$$\begin{aligned}
&= \Phi^*(t,t)V(t) - \Phi^*(t,0)V(0) - \int_0^t \frac{\partial}{\partial s}\Phi^*(t,s)V(s)ds - \int_0^t \Phi^*(t,s)\partial F(Q(s))V(s)ds \\
&= V(t) - 0 - \int_0^t \Big(\frac{\partial}{\partial s}\Phi^*(t,s) + \Phi^*(t,s)\partial F(Q(s))\Big)V(s)ds = V(t),
\end{aligned}$$

while the LHS equals

$$\begin{aligned}
&= \Phi^*(t,t)U(t) - \Phi^*(t,0)U(0) - \int_0^t \frac{\partial}{\partial s}\Phi^*(t,s)U(s)ds \\
&= U(t) + \int_0^t \Phi^*(t,s)\partial F(Q(s))U(s)ds.
\end{aligned}$$

Remark 4.14. It is our preference to use the backward Kolmogorov equation in (4.22), as Britton has done in [1, Chapter 5] to define $V(t)$. However we may have very well used the forward Kolmogorov equation (2.9), used by Kurtz in [4, page 458] to reach the same result as the two diffusion processes are identical.

Having completed most of the work needed, we are now ready to conclude this section (and chapter) by proving our main result, i.e. Theorem 4.1, for the second time, this time by using techniques from [4] and [1]. For the sake of consistency, we also restate it, but using vector notation.

Theorem 4.15. (restatement of Theorem 4.1) Let, for an n-population Z, Π be the EM law for the PEM model, and let $F^{(n)}$, F, and Q be the $\mathbb{R}^{k^2}$-valued equivalents of the functions given in (3.10), and (3.11), respectively. Moreover, let the continuous-time Markov chain $Q^{(n)} = \sum_{i,j\in[k]} \ell_{ij} Q_{ij}^{(n)}$ be the $\mathbb{R}^{k^2}$-valued analogue to the pair-type process associated to the PEM and $V^{(n)}$ be defined as in (4.1). Then, under condition (4.2), for an arbitrary $T \in [0,\infty)$, $V^{(n)} \Rightarrow V$ in $D([0,T),\mathbb{R}^{k^2})$ as $n \to \infty$, where $V(t)$ is the solution of (4.21).

Proof. Let us define

$$\widehat{U}^{(n)}(t) := V^{(n)}(t) - \int_0^t \partial F(Q(s))V^{(n)}(s)ds.$$

Looking at expressions (4.11) to (4.13), and how $\beta^{(n)}(t)$ was defined in (4.19) we see that

$$\widehat{U}^{(n)}(t) = U^{(n)}(t) + \alpha^{(n)}(t) + \beta^{(n)}(t) \Rightarrow U(t) \quad \text{in } D([0,T), \mathbb{R}^{k^2}),$$

as a consequence of Lemma 4.12(c) (where the result $U^{(n)} \Rightarrow U$ was obtained), Lemma 4.13(b) (where the fact that $\sup_{t \leq T} |\beta^{(n)}(t)| \xrightarrow{a.s} 0$ was proved), the similar fact that $\sup_{t \leq T} |\alpha^{(n)}(t)| \xrightarrow{a.s.} 0$ (which was shown in the very beginning of this section), and an application of Slutsky's theorem.

By performing the same integration-by-parts routine we did to show (4.23), but this time for $V^{(n)}(t)$ and $\widehat{U}^{(n)}(t)$, we get

$$V^{(n)}(t) = \widehat{U}^{(n)}(t) + \int_0^t \Phi^*(t,s)\partial F(Q(s))\widehat{U}^{(n)}(s)ds.$$

Finally, since the mapping $\Delta : D([0,T), \mathbb{R}^{k^2}) \to D([0,T), \mathbb{R}^{k^2})$, defined by

$$\Delta(u)(t) = u(t) + \int_0^t \Phi^*(t,s)\partial F(Q(s))u(s)ds$$

is continuous (see [4, page 459]), the continuous mapping theorem tells us that

$$V^{(n)} = \Delta(\widehat{U}^{(n)}) \Rightarrow \Delta(U) = V \quad \text{in } D([0,T), \mathbb{R}^{k^2}),$$

which concludes the proof. □

Chapter 5

CENTRAL LIMIT THEOREM FOR THE MATING PATTERN

5.1 Discussion and motivation

In the previous chapter by stating and proving Theorem 4.1, we showed that given any fixed $T < \infty$, we have $V^{(n)} \Rightarrow V$ in $D([0,T), \mathcal{M}^{k\times k}(\mathbb{R}))$, as $n \to \infty$. This is a CLT for the pair-type process, and stands as a complement to Theorem 3.4. We are still however, left with the matter of a CLT for the mating pattern unresolved. To approach this issue correctly, we take a closer look at the final section of Chapter 3, or more specifically, how Theorem 3.5 (the LLN for the mating pattern) extends Theorem 3.4 (the LLN for the pair-type process).

Recall that the deterministic function Q describes how the rescaled pair-type process $n^{-1}Q^{(n)}$ behaves for infinite populations (i.e. if we take $n \to \infty$). Then, the value $Q(\infty) = \lim_{t\to\infty} Q(t)$, becomes in essence, the mating pattern in this infinite population case. What Theorem 3.5 tells us is that we can flip the limiting procedure around. Indeed, by taking the time limit first, and the population size limit afterwards, we get by said theorem

$$\lim_{n\to\infty} \lim_{t\to\infty} n^{-1}Q^{(n)}(t) = \lim_{n\to\infty} n^{-1}Q^{(n)}(T_n) = Q(\infty), \tag{5.1}$$

since T_n is a.s. finite, thus showing that the order of limits does not matter. This observation will prove to be of use towards the end of this chapter, as we analyze the evidence obtained from our MATLAB simulations data to argue the existence of a CLT for the mating pattern. Before that however, we use this data to validate our theoretical results in Chapter 4. We will mainly focus on comparing the finite-dimensional distributions (f.d.d.) of V and $V^{(n)}$ for large n.

5.2 Distributional properties of V and $V^{(n)}$ and methods used

Obtaining the f.d.d. of $V^{(n)}$ requires a simulation procedure, which although straightforward, becomes relatively time-consuming for larger populations. The main idea is to initially construct the embedded discrete-time Markov chain of the pair-type process. In our case this would be the $\mathbb{R}^{k^2}$ equivalent of the chain $X^{(n)}(s)$, $s \in \mathbb{N} \cup \{0\}$ with state space $\mathcal{E}_n^*$ given as in (3.3) and jump probabilities

$$\mathbb{P}\Big(X^{(n)}(s+1) = M + I^{ij} \mid X^{(n)}(s) = M\Big) = \frac{F_{ij}^{(n)}(n^{-1}M)}{\sum_{i',j' \in [k]} F_{i'j'}^{(n)}(n^{-1}M)}. \tag{5.2}$$

As we have mentioned in the beginning of Chapter 4, to switch from the matrix to the vector version of $X^{(n)}$, as a helpful visualization, one can imagine stacking the matrix columns of top of one another, with columns further on the left being above the ones on the right. To make the transition from the discrete $X^{(n)}$ to the continuous $Q^{(n)}$, we use the waiting time parameter function $\lambda_n(M) = n \sum_{i'j' \in [k]} F_{i'j'}^{(n)}(n^{-1}M)$ which gives as outputs the rates of the exponential waiting times between successive jumps. Once $Q^{(n)}$ has been finally constructed on some arbitrary interval $[0, T]$, we proceed to construct $V^{(n)}$ on the same interval via $V^{(n)}(t) = n^{-1/2}Q^{(n)}(t) - \sqrt{n}Q(t)$. As a solution to the integral equation (3.11), Q can be retrieved numerically via different discretization routines (we give more details in the following section).

To check whether $V^{(n)}$'s f.d.d. actually converge to a desired limit, we need to first know how the process V is distributed. One obvious approach is to pick a large T, then simulate a discrete time step approximation of V over the interval $[0, T]$ a large number of times. Afterwards one would look at how the values of $V(t)$ at different time points $0 < t_1 < t_2 < ... < t_d < T$ are distributed. This can be done via several methods found in [8], the simplest of which is the Euler-Maruyama method (essentially a stochastic version of the classical Euler method).

In this thesis however, we use a different and much faster approach, that relies heavily on the nature of the process V, and does not require any simulations. Recall from the previous chapter that we can look at V as the solution of equation (4.21), taking values in $\mathbb{R}^{k^2}$, instead of $\mathcal{M}^{k \times k}(\mathbb{R})$. Re-adopting the notation from Section 4.3, we define the $k^2 \times k^2$ matrix-valued function

$$G(x) := \sum_{i,j\in[k]} \ell_{ij}\ell_{ij}^T F_{ij}(x) \tag{5.3}$$

and observe that $G(x)$ is a diagonal matrix with its diagonal entries given by the vector $\big(F_{11}(x), F_{21}(x), ..., F_{k-1\ k}(x), F_{kk}(x)\big)$. Furthermore, let $\Phi^*(t,s)$ be the solution of equation (4.22). In view of our discussion at the end of Chapter 2, it is then clear that V is a Gaussian process with $\mathbb{E}[V(t)] = 0$, and its f.d.d. uniquely determined by its covariance function

$$\mathbf{Cov}[V(t), V(r)] = \int_0^{t\wedge r} \Phi^*(t,s)G(Q(s))[\Phi^*(r,s)]^T ds. \tag{5.4}$$

Consequently, at each $t \in [0,T]$, $V(t)$ is a normally distributed vector with mean 0, and covariance matrix

$$\mathcal{C}(t) = \sum_{i,j\in[k^2]} I^{ij}\mathcal{C}_{ij}(t) := \mathbf{Cov}[V(t), V(t)]. \tag{5.5}$$

It is not clear, even in a Poisson fine balance case, whether an analytic solution for the covariance function in (5.4) exists. However, solving it numerically does not pose a challenge once one has retrieved Q and $\frac{d}{dt}Q(t)$. Indeed, the diagonal entries of $G(Q(t))$ are nothing but the corresponding entries of $\frac{d}{dt}Q(t)$. Moreover, after one observes that the matrix function $\partial F(Q(t))$ can be assembled similarly on $[0,T]$, $\Phi^*(t,s)$ is obtained by numerically solving equation (4.22). This approach, finalized with a numerical integration procedure, successfully yields the matrix $\mathbf{Cov}[V(t), V(r)]$, which can be later used to compare the f.d.d. of $V(t)$ with the empirical distributions obtained from the simulation data. For simplicity purposes however, in the following section we will fix $t = r$ and use $\mathcal{C}(t)$ to compare the distributions in the one-dimensional case (i.e. for singular time-point cases).

5.3 Results and future work

5.3.1 Simulation parameters and details

Implementing in MATLAB the methods just explained in the previous section, we obtained Figures 5.1 to 5.8 found at the end of this chapter. Specifically, we chose an n-population Z (recall that this means Z is comprised of n males and n females) where $n = 10^5$. The number of types was picked to be $k = 3$, and a pseudo-random number generator was used to fix the values

$$\begin{aligned}(x_1, x_2, x_3) &= (0.1708369602, 0.6700803438, 0.1590826959)\\ (y_1, y_2, y_3) &= (0.4550358617, 0.2419355183, 0.3030286200),\end{aligned} \tag{5.6}$$

for the limiting type-ratios in the PEM model, whereas the vectors $(x_1^{(n)}, \ldots, x_k^{(n)})$ and $(y_1^{(n)}, \ldots, y_k^{(n)})$, giving the type-ratios in Z were tuned so as to satisfy (4.2) and (3.1). Similarly,

$$\begin{aligned}(\alpha_1, \alpha_2, \alpha_3) &= (1.5481729240, 1.8284700382, 1.5921598655)\\ (\beta_1, \beta_2, \beta_3) &= (2.2575758178, 2.1596756204, 2.1998583932),\end{aligned} \tag{5.7}$$

were set to be the respective Poisson firing rates for the females and males. These combined with the randomly picked (also using the same pseudo-random number generator) pair-type preference matrix

$$P := (p_{ij})_{i,j\in[3]} = \begin{pmatrix} 0.7659078565 & 0.1877212287 & 0.4413092229\\ 0.5184179879 & 0.0807412688 & 0.1583098677\\ 0.2968005016 & 0.7384402962 & 0.8799370312 \end{pmatrix}, \tag{5.8}$$

fix the EM law for the PEM model (see Definition 3.1), in the form of the matrix

$$\Pi := (\pi_{ij})_{i,j\in[3]} = \begin{pmatrix} 2.9148528611 & 0.6960418844 & 1.6540407880\\ 2.1182796710 & 0.3220079405 & 0.6377241411\\ 1.1426034817 & 2.7705065075 & 3.3367372888 \end{pmatrix}. \tag{5.9}$$

Next, picking $T = 20$, and a discretization time step $h = 10^{-4}$, we used the Runge-Kutta method of order 4, to retrieve Q on the interval $[0, T]$. In fact, $Q(t)$, $\frac{d}{dt}Q(t) = F(Q(t))$, and $\partial F(Q(t))$ were all retrieved in the same MATLAB function, as each of them can be constructed from the rest (this is obvious from the way they are defined). Having retrieved $\partial F(Q(t))$, we used the classical Euler method to solve the backward Kolmogorov equation (4.22), and obtain Φ^*. This was finalized with applying the trapezoidal integration method on the equation in (5.4) to obtain the covariance function.

Using the parameters given in (5.6)-(5.8), and the methods explained in the previous section, numerical values for the matrix-valued function $\mathcal{C}(t)$ were obtained on a discrete partition of $[0, T]$ with sparsity 0.1. Moreover, using these same parameters, the vector $V^{(n)}$ was repeatedly simulated to create a random sample of size $N = 10^5$. A sample covariance matrix for $V^{(n)}(t)$ was calculated from the empirical data for each of a total of 24 time points, and the results plotted against the function $\mathcal{C}(t)$.

5.3.2 Discussion of results and emerging conjectures

Several things can be noted regarding the resulting plots given in Figures 5.1 to 5.8. First, it is strongly suggested that the Gaussian vector $V(t)$ assumes a distributional limit as $t \to \infty$. It is highly unlikely that this observation is due to chance. To check, we repeatedly calculated $\mathcal{C}(t)$, each time with a different seed for the pseudo-random number generator, to obtain parameter values different from the ones given in (5.6)-(5.8). Each time $\mathcal{C}(t)$ was plotted, it showed similar limiting behavior, and because of this, we henceforth assume the existence of this limit and write $V(\infty) := \lim_{t\to\infty} V(t)$, where the convergence is distributional. A rigorous proof of this existence however, remains an open problem and is left as a future project.

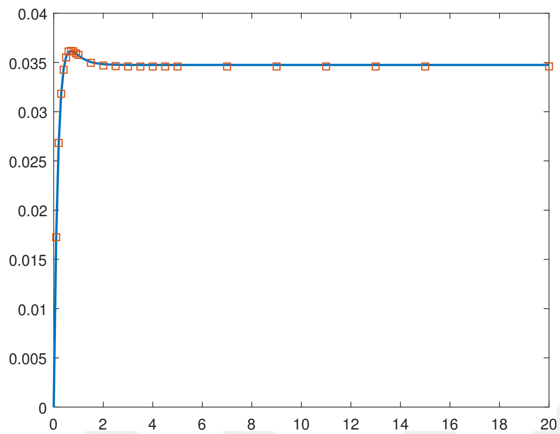
Second, we can observe how for a large (but fixed) n, the data strongly suggests that the process $V^{(n)}(t)$, also admits a distributional limit as $t \to \infty$. As opposed to the similar observation for V, this one should come off as no surprise, since we know that, $\mathbb{P}$-a.s.

$$\lim_{t\to\infty} V^{(n)}(t) = \lim_{t\to\infty} \left(n^{-1/2}Q^{(n)}(t) - \sqrt{n}Q(t)\right) = n^{-1/2}Q^{(n)}(T_n) - Q(\infty).$$

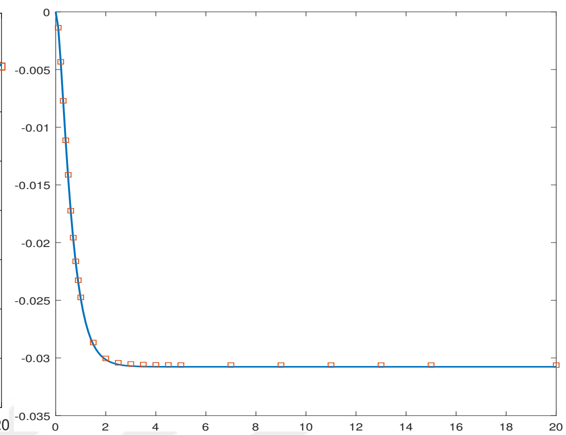
What we find more interesting, however, is that for large n, the covariance functions of V and $V^{(n)}$ do not deviate from each other as t approaches infinity. Taking into perspective our discussion regarding the LLN for the mating pattern in Section 5.1, this observation suggests that

$$V^{(n)}(T_n) := \lim_{t\to\infty} V^{(n)}(t) \Rightarrow V(\infty), \quad \text{as} \quad n \to \infty,$$

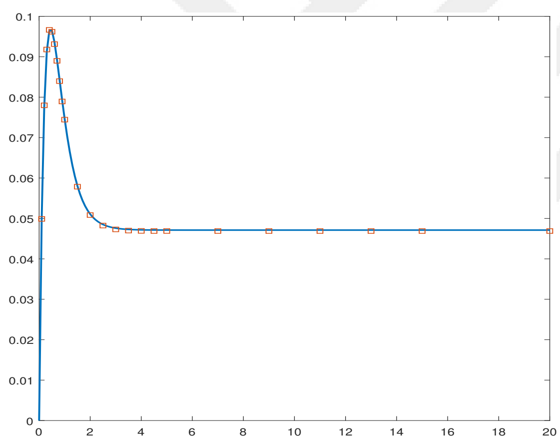
which if true, confirms the existence of a CLT for the mating pattern. This is another open problem, which builds on the previous one, and whose proof we leave for future projects. Here, instead of a proof, we provide results from our simulation data as positive evidence to support this conjecture. To obtain this new set of data, we simulated $V^{(n)}(T_n)$ a total of 10^5 times. While doing so, we used the same initial parameters given in (5.6)-(5.8), and conveniently approximated $Q(\infty)$ with $Q(20)$. Similarly, $T = 20$ was taken to be high enough to approximate $\mathcal{C}(\infty)$, the covariance matrix of $V(\infty)$. The histograms (see Figures 5.9 and 5.10) show that for large n, the components of $V^{(n)}(T_n)$ are approximately normally distributed, as expected, with mean and variance corresponding to what $\mathcal{C}(\infty)$ was calculated to be.



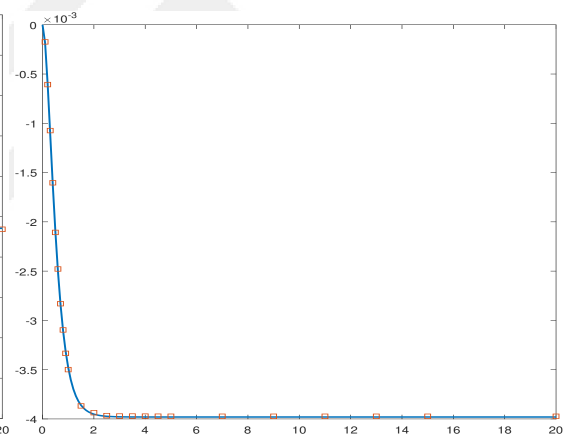
(a) $\mathcal{C}_{11}(t)$ vs. $\text{Var}\big(V_1^{(n)}(t)\big)$



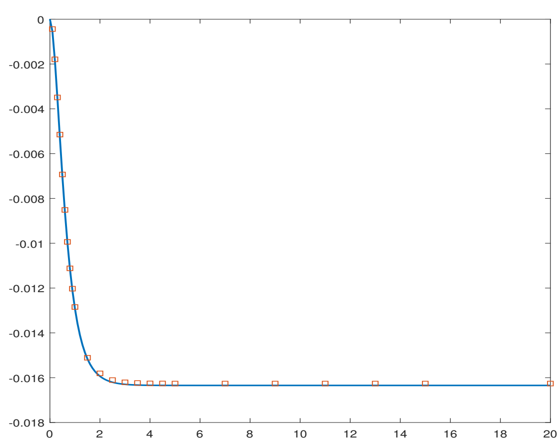
(b) $\mathcal{C}_{21}(t)$ vs. $\text{Cov}[V_2^{(n)}(t), V_1^{(n)}(t)]$



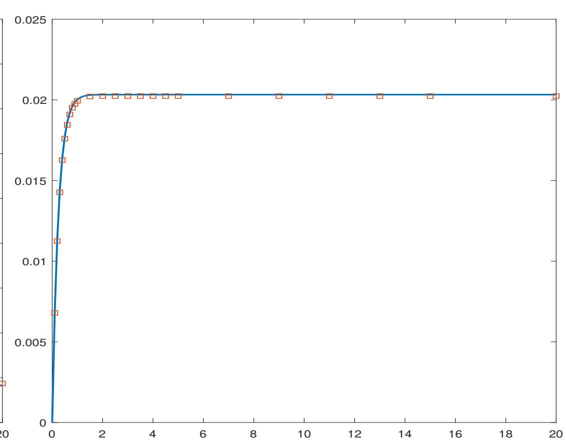
(c) $\mathcal{C}_{22}(t)$ vs. $\text{Var}\big(V_2^{(n)}(t)\big)$



(d) $\mathcal{C}_{31}(t)$ vs. $\text{Cov}[V_3^{(n)}(t), V_1^{(n)}(t)]$

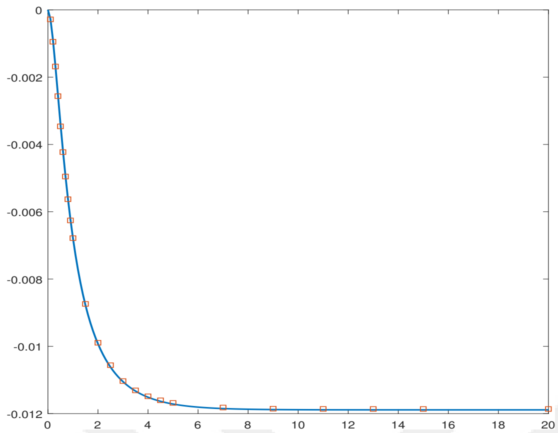


(e) $\mathcal{C}_{32}(t)$ vs. $\text{Cov}[V_3^{(n)}(t), V_2^{(n)}(t)]$

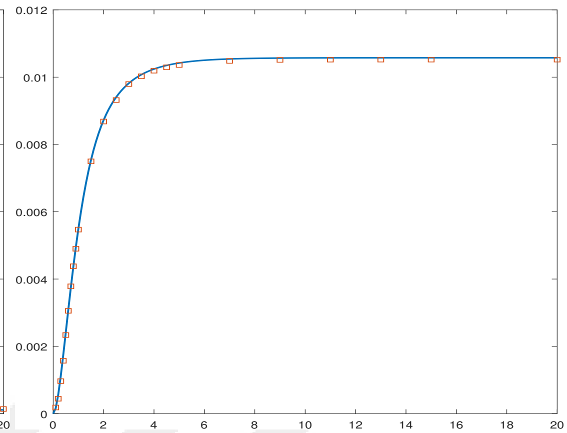


(f) $\mathcal{C}_{33}(t)$ vs. $\text{Var}\big(V_3^{(n)}(t)\big)$

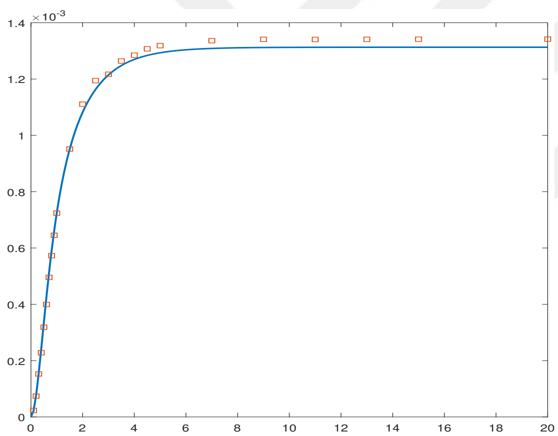
Figure 5.1: $\mathcal{C}(t) := \mathbf{Cov}[V(t), V(t)]$ (blue line) is plotted against the corresponding sample covariances for $V^{(n)}(t)$, drawn from the simulation data (red squares).



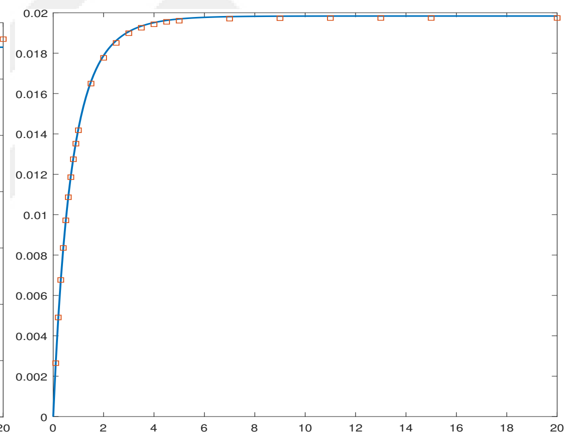
(a) $\mathcal{C}_{41}(t)$ vs. $\mathrm{Cov}[V_4^{(n)}(t), V_1^{(n)}(t)]$



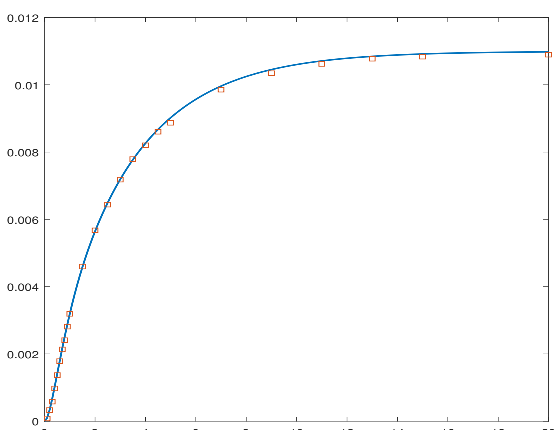
(b) $\mathcal{C}_{42}(t)$ vs. $\mathrm{Cov}[V_4^{(n)}(t), V_2^{(n)}(t)]$



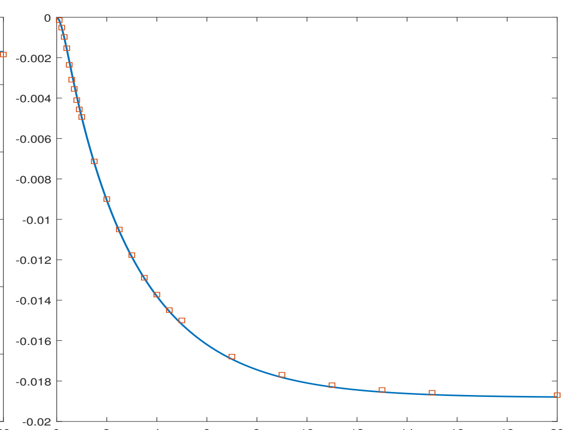
(c) $\mathcal{C}_{43}(t)$ vs. $\mathrm{Cov}[V_4^{(n)}(t), V_3^{(n)}(t)]$



(d) $\mathcal{C}_{44}(t)$ vs. $\mathrm{Var}\big(V_4^{(n)}(t)\big)$

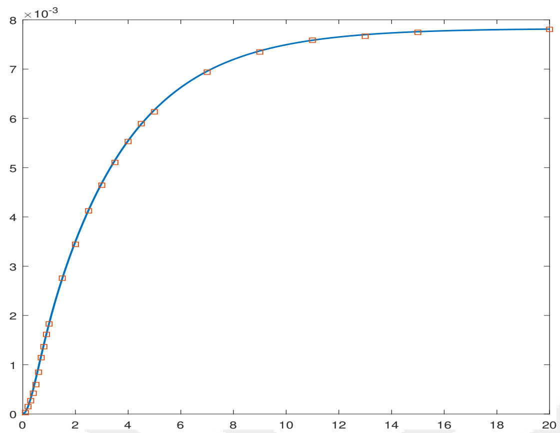


(e) $\mathcal{C}_{51}(t)$ vs. $\mathrm{Cov}[V_5^{(n)}(t), V_1^{(n)}(t)]$

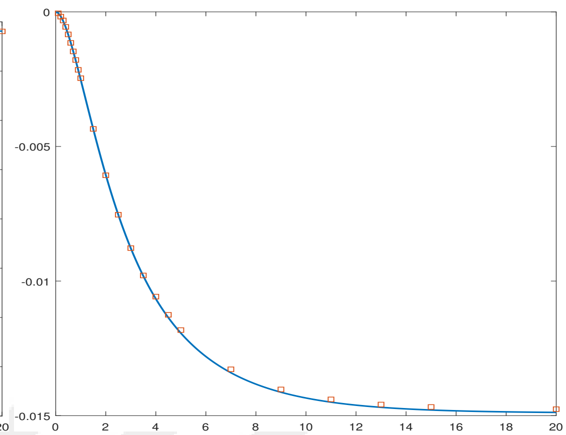


(f) $\mathcal{C}_{52}(t)$ vs. $\mathrm{Cov}[V_5^{(n)}(t), V_2^{(n)}(t)]$

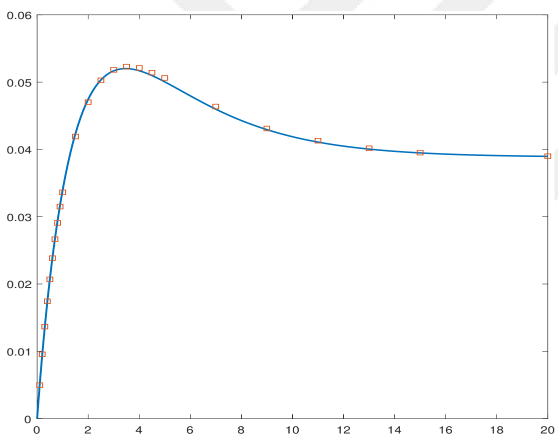
Figure 5.2: $\mathcal{C}(t) := \mathbf{Cov}[V(t), V(t)]$ (blue line) is plotted against the corresponding sample covariances for $V^{(n)}(t)$, drawn from the simulation data (red squares).



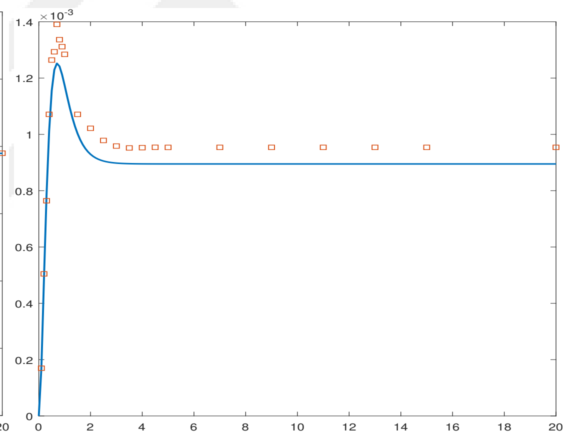
(a) $\mathcal{C}_{53}(t)$ vs. $\mathrm{Cov}[V_5^{(n)}(t), V_3^{(n)}(t)]$



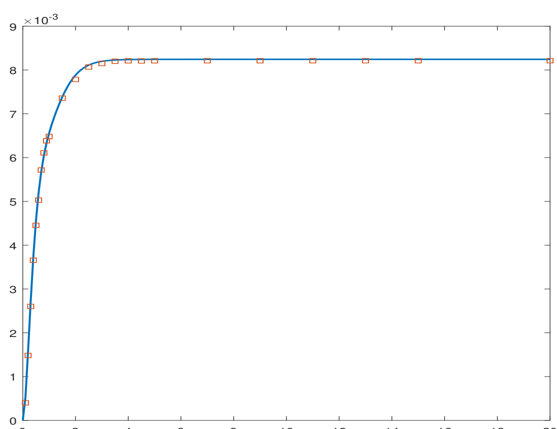
(b) $\mathcal{C}_{54}(t)$ vs. $\mathrm{Cov}[V_5^{(n)}(t), V_4^{(n)}(t)]$



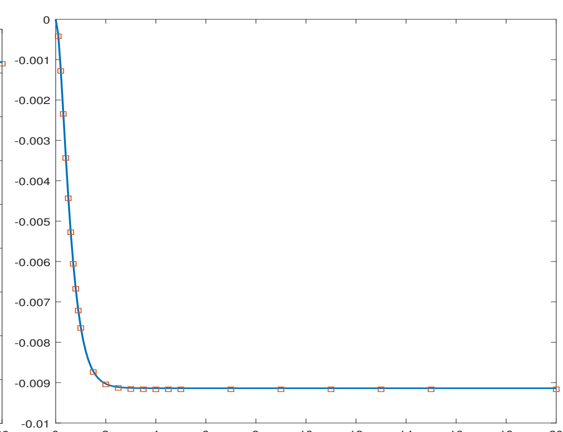
(c) $\mathcal{C}_{55}(t)$ vs. $\mathrm{Var}\big(V_5^{(n)}(t)\big)$



(d) $\mathcal{C}_{61}(t)$ vs. $\mathrm{Cov}[V_6^{(n)}(t), V_1^{(n)}(t)]$

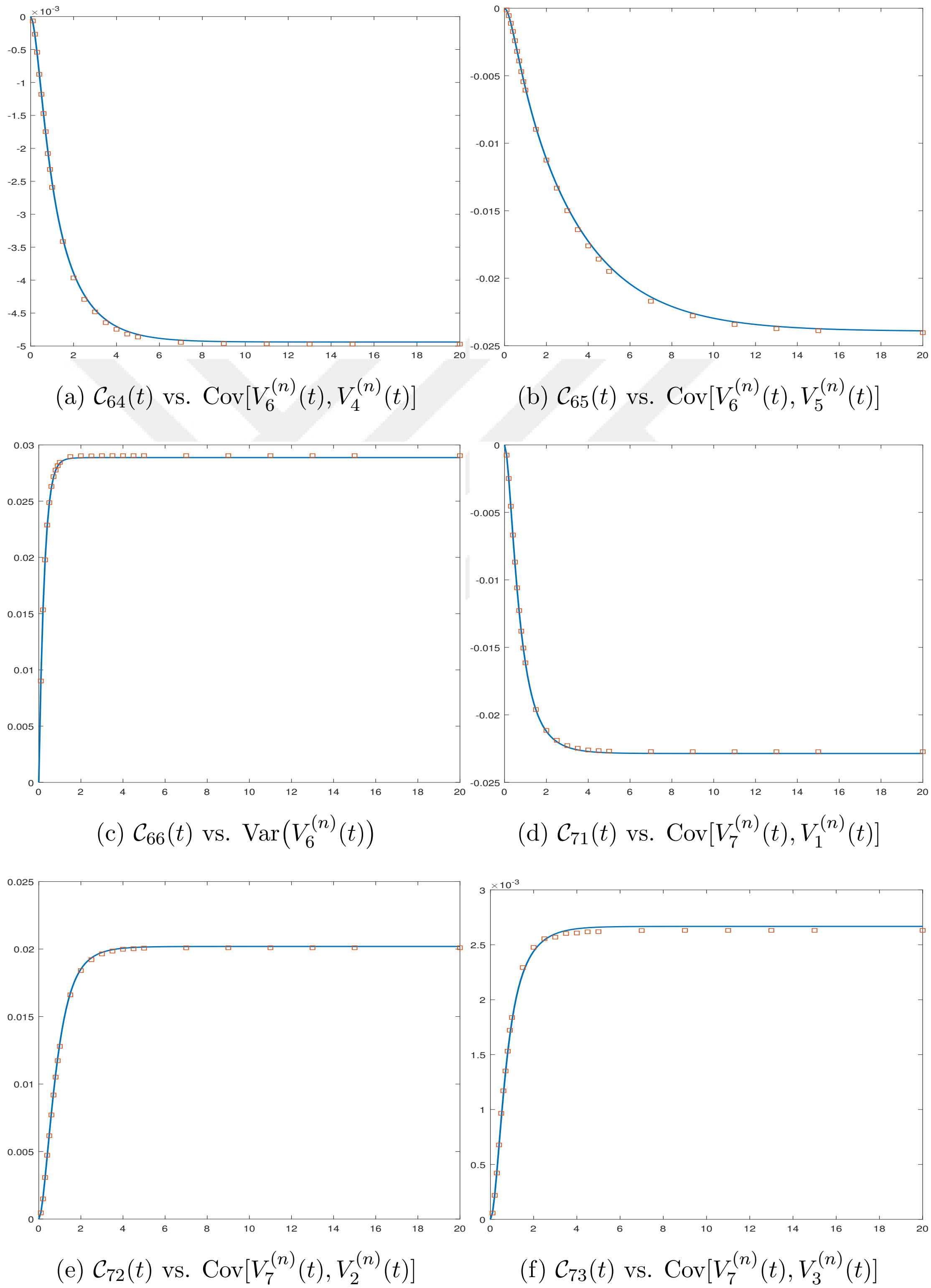


(e) $\mathcal{C}_{62}(t)$ vs. $\mathrm{Cov}[V_6^{(n)}(t), V_2^{(n)}(t)]$



(f) $\mathcal{C}_{63}(t)$ vs. $\mathrm{Cov}[V_6^{(n)}(t), V_3^{(n)}(t)]$

Figure 5.3: $\mathcal{C}(t) := \mathbf{Cov}[V(t), V(t)]$ (blue line) is plotted against the corresponding sample covariances for $V^{(n)}(t)$, drawn from the simulation data (red squares).



(a) $\mathcal{C}_{64}(t)$ vs. $\mathrm{Cov}[V_6^{(n)}(t), V_4^{(n)}(t)]$

(b) $\mathcal{C}_{65}(t)$ vs. $\mathrm{Cov}[V_6^{(n)}(t), V_5^{(n)}(t)]$

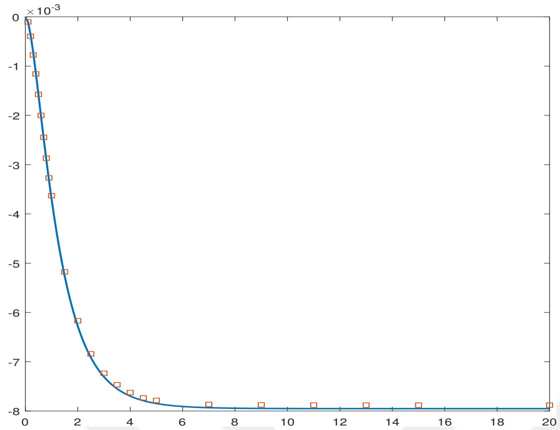
(c) $\mathcal{C}_{66}(t)$ vs. $\mathrm{Var}\big(V_6^{(n)}(t)\big)$

(d) $\mathcal{C}_{71}(t)$ vs. $\mathrm{Cov}[V_7^{(n)}(t), V_1^{(n)}(t)]$

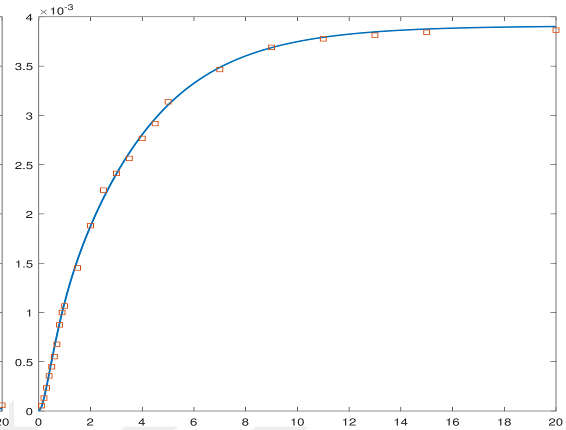
(e) $\mathcal{C}_{72}(t)$ vs. $\mathrm{Cov}[V_7^{(n)}(t), V_2^{(n)}(t)]$

(f) $\mathcal{C}_{73}(t)$ vs. $\mathrm{Cov}[V_7^{(n)}(t), V_3^{(n)}(t)]$

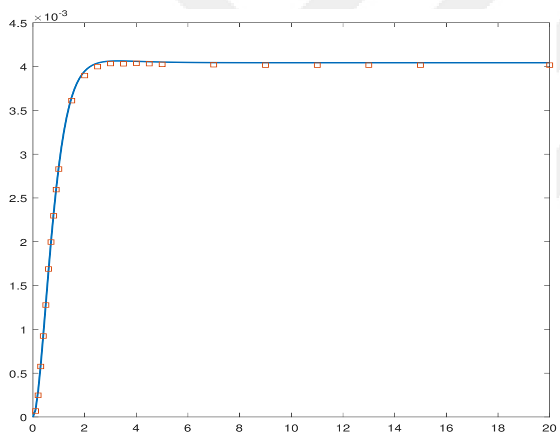
Figure 5.4: $\mathcal{C}(t) := \mathbf{Cov}[V(t), V(t)]$ (blue line) is plotted against the corresponding sample covariances for $V^{(n)}(t)$, drawn from the simulation data (red squares).



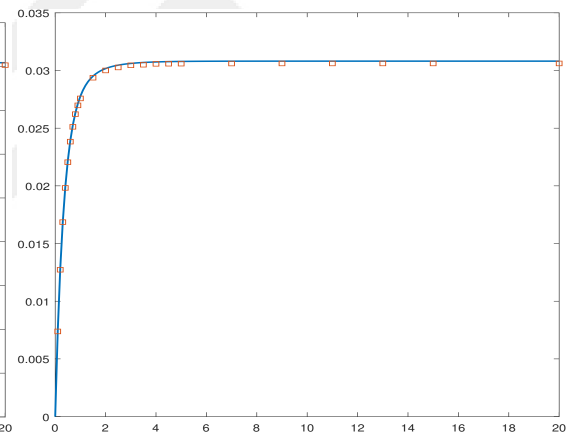
(a) $\mathcal{C}_{74}(t)$ vs. $\text{Cov}[V_7^{(n)}(t), V_4^{(n)}(t)]$



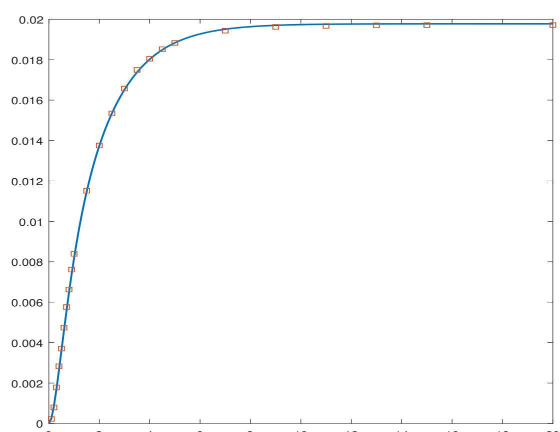
(b) $\mathcal{C}_{75}(t)$ vs. $\text{Cov}[V_7^{(n)}(t), V_5^{(n)}(t)]$



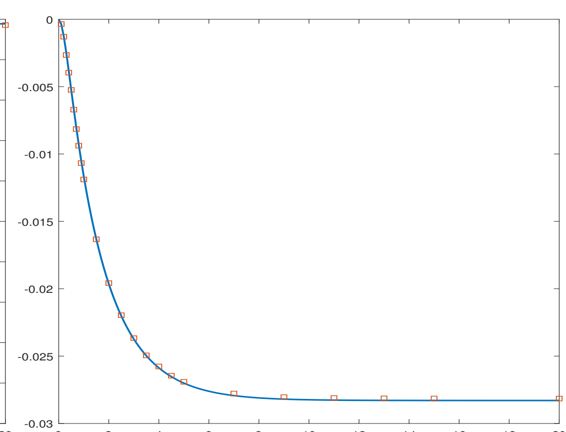
(c) $\mathcal{C}_{76}(t)$ vs. $\text{Cov}[V_7^{(n)}(t), V_6^{(n)}(t)]$



(d) $\mathcal{C}_{77}(t)$ vs. $\text{Var}\big(V_7^{(n)}(t)\big)$

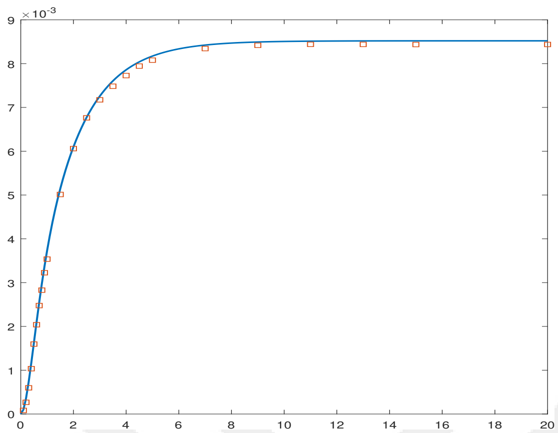


(e) $\mathcal{C}_{81}(t)$ vs. $\text{Cov}[V_8^{(n)}(t), V_1^{(n)}(t)]$

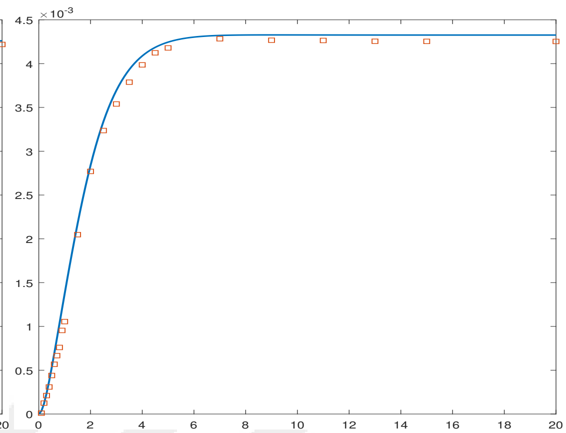


(f) $\mathcal{C}_{82}(t)$ vs. $\text{Cov}[V_8^{(n)}(t), V_2^{(n)}(t)]$

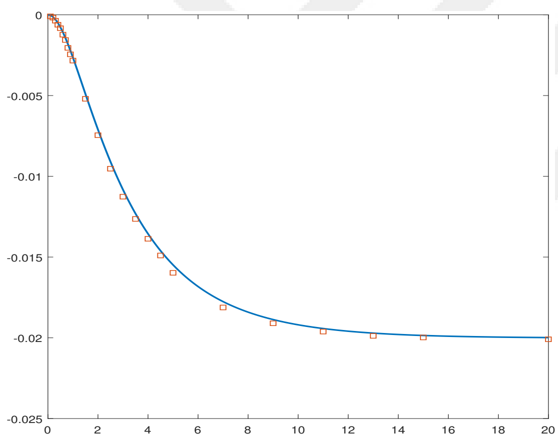
Figure 5.5: $\mathcal{C}(t) := \mathbf{Cov}[V(t), V(t)]$ (blue line) is plotted against the corresponding sample covariances for $V^{(n)}(t)$, drawn from the simulation data (red squares).



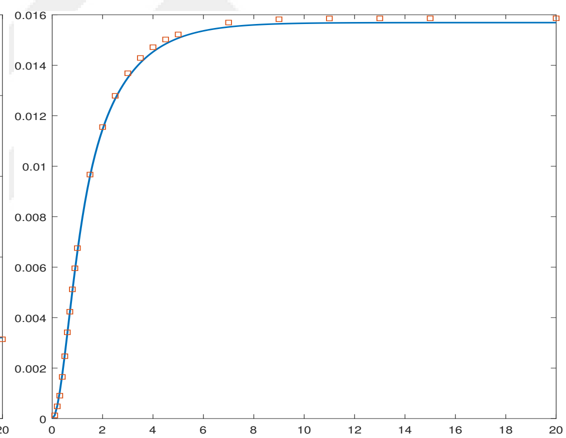
(a) $\mathcal{C}_{83}(t)$ vs. $\mathrm{Cov}[V_8^{(n)}(t), V_3^{(n)}(t)]$



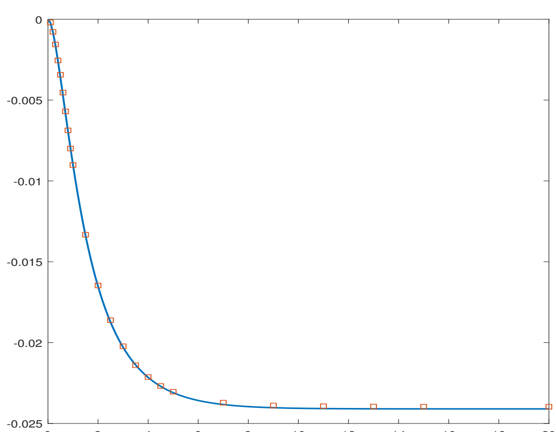
(b) $\mathcal{C}_{84}(t)$ vs. $\mathrm{Cov}[V_8^{(n)}(t), V_4^{(n)}(t)]$



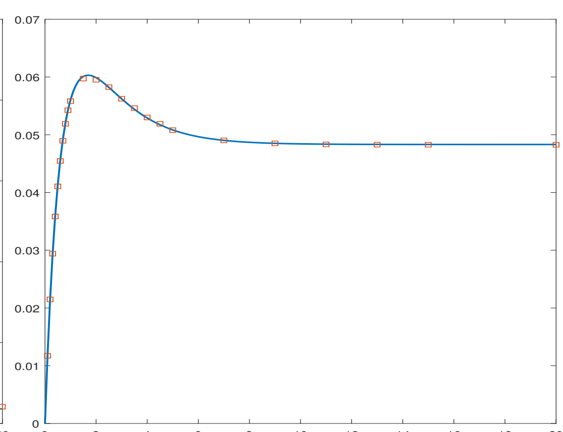
(c) $\mathcal{C}_{85}(t)$ vs. $\mathrm{Cov}[V_8^{(n)}(t), V_5^{(n)}(t)]$



(d) $\mathcal{C}_{86}(t)$ vs. $\mathrm{Cov}[V_8^{(n)}(t), V_6^{(n)}(t)]$

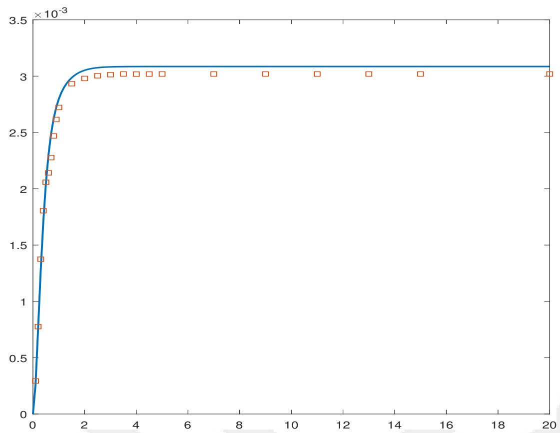


(e) $\mathcal{C}_{87}(t)$ vs. $\mathrm{Cov}[V_8^{(n)}(t), V_7^{(n)}(t)]$

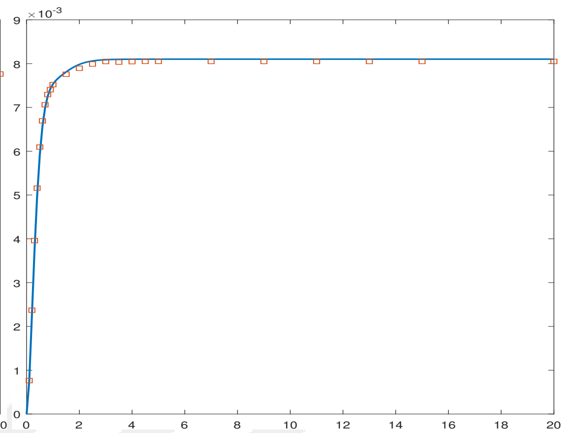


(f) $\mathcal{C}_{88}(t)$ vs. $\mathrm{Var}\big(V_8^{(n)}(t)\big)$

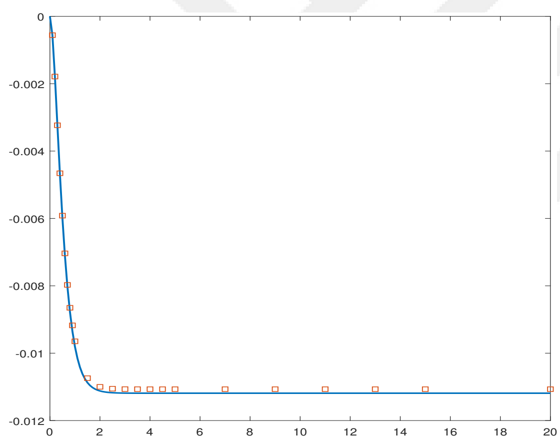
Figure 5.6: $\mathcal{C}(t) := \mathbf{Cov}[V(t), V(t)]$ (blue line) is plotted against the corresponding sample covariances for $V^{(n)}(t)$, drawn from the simulation data (red squares).



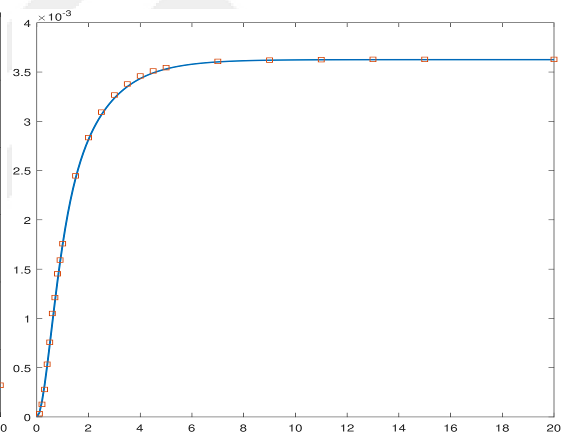
(a) $\mathcal{C}_{91}(t)$ vs. $\mathrm{Cov}[V_9^{(n)}(t), V_1^{(n)}(t)]$



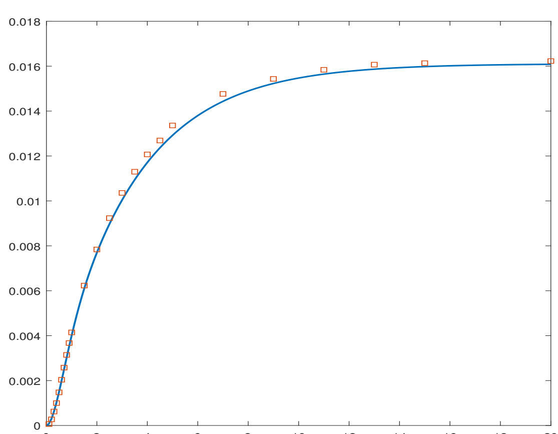
(b) $\mathcal{C}_{92}(t)$ vs. $\mathrm{Cov}[V_9^{(n)}(t), V_2^{(n)}(t)]$



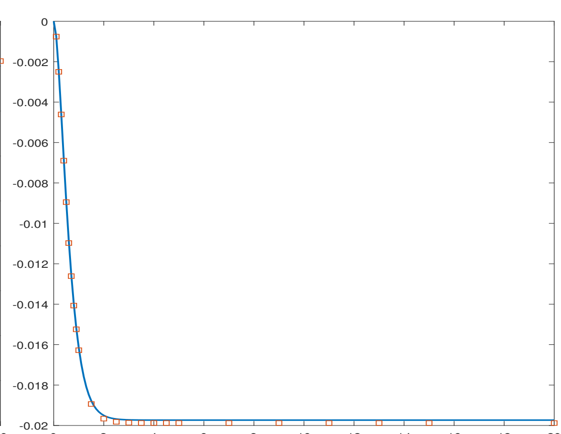
(c) $\mathcal{C}_{93}(t)$ vs. $\mathrm{Cov}[V_9^{(n)}(t), V_3^{(n)}(t)]$



(d) $\mathcal{C}_{94}(t)$ vs. $\mathrm{Cov}[V_9^{(n)}(t), V_4^{(n)}(t)]$

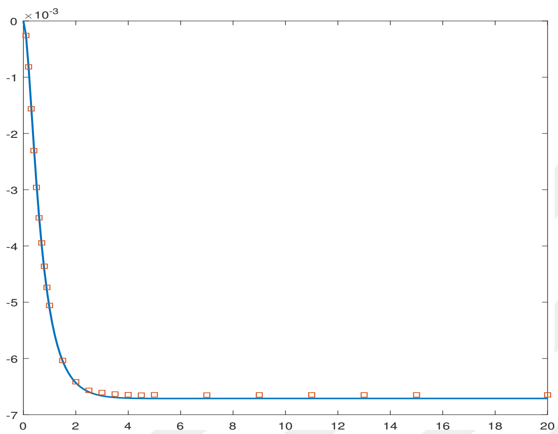


(e) $\mathcal{C}_{95}(t)$ vs. $\mathrm{Cov}[V_9^{(n)}(t), V_5^{(n)}(t)]$

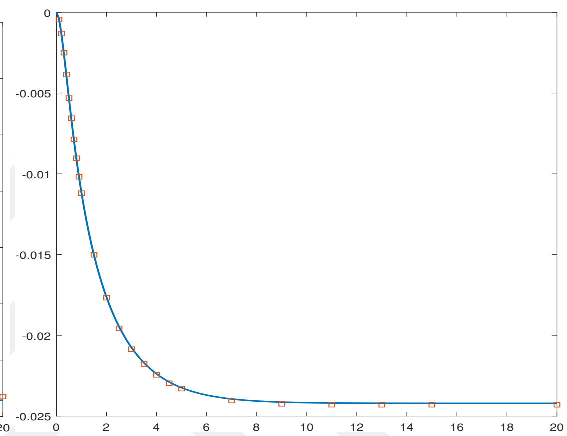


(f) $\mathcal{C}_{96}(t)$ vs. $\mathrm{Cov}[V_9^{(n)}(t), V_6^{(n)}(t)]$

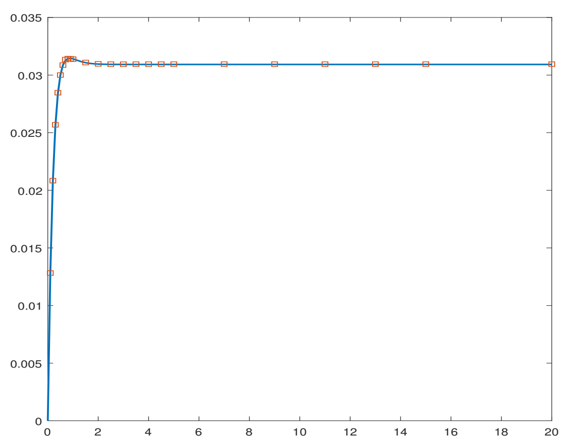
Figure 5.7: $\mathcal{C}(t) := \mathbf{Cov}[V(t), V(t)]$ (blue line) is plotted against the corresponding sample covariances for $V^{(n)}(t)$, drawn from the simulation data (red squares).



(a) $\mathcal{C}_{97}(t)$ vs. $\mathrm{Cov}[V_9^{(n)}(t), V_7^{(n)}(t)]$

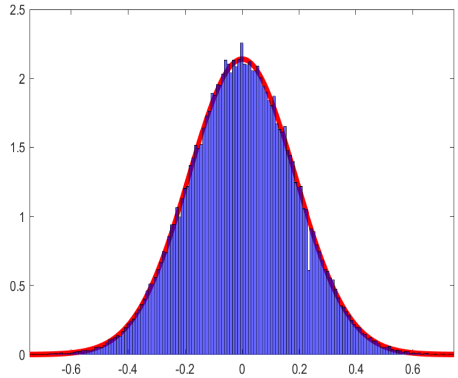


(b) $\mathcal{C}_{98}(t)$ vs. $\mathrm{Cov}[V_9^{(n)}(t), V_8^{(n)}(t)]$

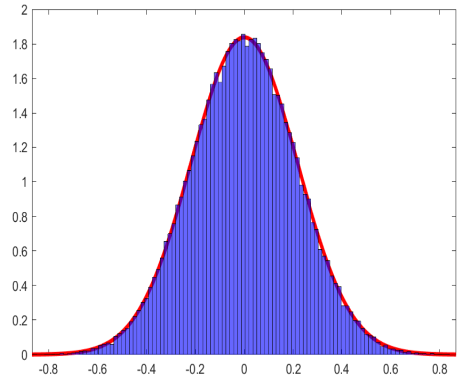


(c) $\mathcal{C}_{99}(t)$ vs. $\mathrm{Var}\big(V_9^{(n)}(t)\big)$

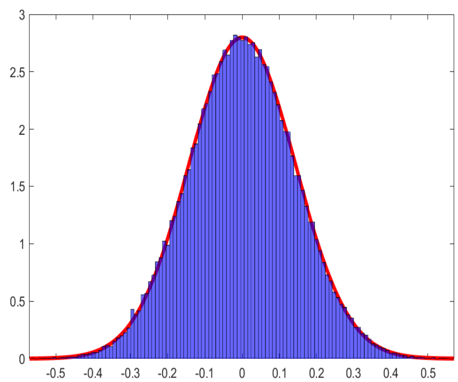
Figure 5.8: $\mathcal{C}(t) := \mathbf{Cov}[V(t), V(t)]$ (blue line) is plotted against the corresponding sample covariances for $V^{(n)}(t)$, drawn from the simulation data (red squares).



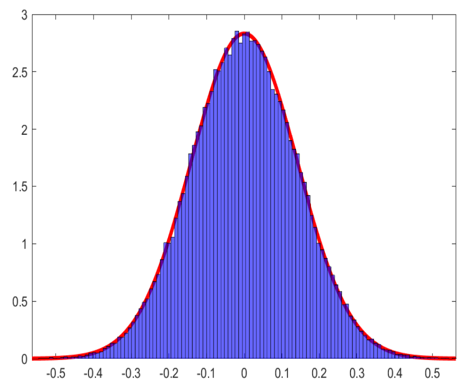
(a) $V_1(\infty)$ vs. $V_1^{(n)}(T_n)$



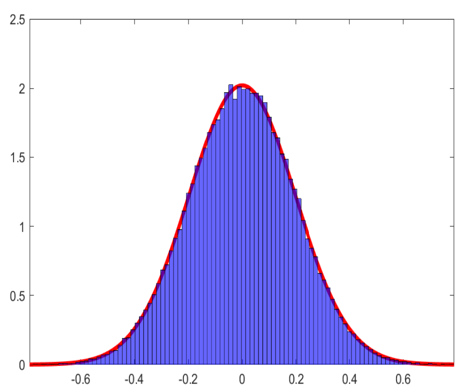
(b) $V_2(\infty)$ vs. $V_2^{(n)}(T_n)$



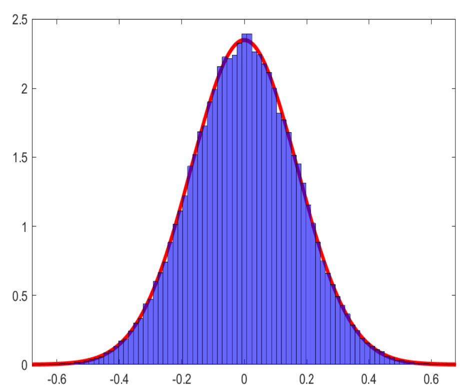
(c) $V_3(\infty)$ vs. $V_3^{(n)}(T_n)$



(d) $V_4(\infty)$ vs. $V_4^{(n)}(T_n)$

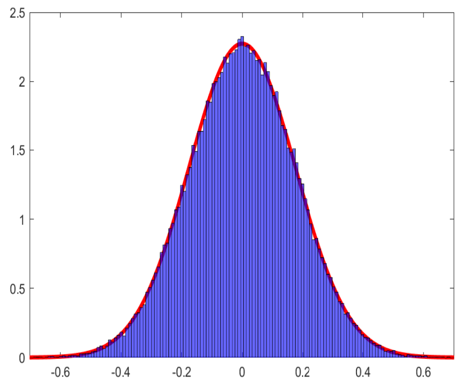


(e) $V_5(\infty)$ vs. $V_5^{(n)}(T_n)$

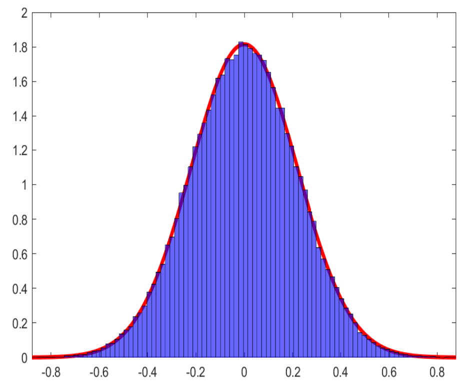


(f) $V_6(\infty)$ vs. $V_6^{(n)}(T_n)$

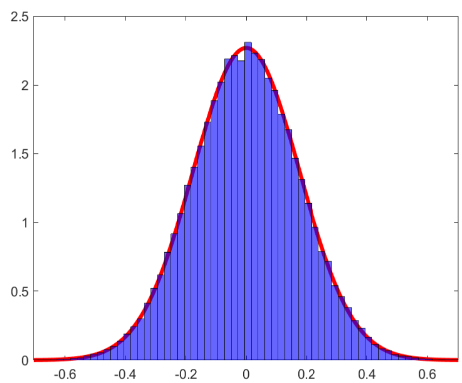
Figure 5.9: Density histograms (blue) retrieved from $V^{(n)}(T_n)$'s sample data are plotted against the density curves (red) of $V(\infty)$'s vector components.



(a) $V_7(\infty)$ vs. $V_7^{(n)}(T_n)$



(b) $V_8(\infty)$ vs. $V_8^{(n)}(T_n)$



(c) $V_9(\infty)$ vs. $V_9^{(n)}(T_n)$

Figure 5.10: Density histograms (blue) retrieved from $V^{(n)}(T_n)$'s sample data are plotted against the density curves (red) of $V(\infty)$'s vector components.

BIBLIOGRAPHY

[1] Andersson, H. k., and Britton, T. *Stochastic epidemic models and their statistical analysis*, vol. 151 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 2000.

[2] Billingsley, P. *Convergence of probability measures*, second ed. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, 1999. A Wiley-Interscience Publication.

[3] Burley, N. The meaning of assortative mating. *Ethology and Sociobiology 4*, 4 (1983), 191–203.

[4] Ethier, S. N., and Kurtz, T. G. *Markov processes: characterization and convergence*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986.

[5] Gimelfarb, A. Processes of pair formation leading to assortative mating in biological populations: encounter-mating model. *The American Naturalist 131*, 6 (1988), 865–884.

[6] Gün, O., and Yilmaz, A. Fluid limit for the Poisson encounter-mating model. *Adv. in Appl. Probab. 49*, 4 (2017), 1201–1229.

[7] Gün, O., and Yilmaz, A. The stochastic encounter-mating model. *Acta Appl. Math. 148* (2017), 71–102.

[8] Kloeden, P. E., and Platen, E. *Numerical solution of stochastic differential equations*, vol. 23 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1992.

[9] Kurtz, T. G. Solutions of ordinary differential equations as limits of pure jump Markov processes. *J. Appl. Probability 7* (1970), 49–58.

[10] Kurtz, T. G. Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *J. Appl. Probability 8* (1971), 344–356.

[11] Kurtz, T. G. Strong approximation theorems for density dependent Markov chains. *Stochastic Processes Appl. 6*, 3 (1977/78), 223–240.

[12] Lande, R. Models of speciation by sexual selection on polygenic traits. *Proceedings of the National Academy of Sciences 78*, 6 (1981), 3721–3725.

[13] Mosteller, F. Association and estimation in contingency tables. *Journal of the American Statistical Association 63*, 321 (1968), 1–28.

[14] O'Donald, P. *Genetic models of sexual selection.* Cambridge University Press, 1980.

[15] Romney, A. K. Measuring endogamy. *Explorations in mathematical anthropology* (1971), 191–213.

[16] Scudo, F. M., and Karlin, S. Assortative mating based on phenotype: I. two alleles with dominance. *Genetics 63*, 2 (1969), 479.