STOCHASTIC MODELS ASSOCIATED WITH THE TWO-PARAMETER POISSON-DIRICHLET DISTRIBUTION

STOCHASTIC MODELS ASSOCIATED WITH THE TWO-PARAMETER POISSON-DIRICHLET DISTRIBUTION

By Fang Xu, B.Sc

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree Doctor of Philosophy

McMaster University ©Copyright by Fang Xu, November 2011 DOCTOR OF PHILOSOPHY (2011) (Mathematics)

McMaster University Hamilton, Ontario

TITLE: STOCHASTIC MODELS ASSOCIATED WITH THE TWO-PARAMETER POISSON-DIRICHLET DISTRIBUTION AUTHOR: Fang Xu, B.Sc. (Beijing Normal University) SUPERVISOR: Professor Shui Feng NUMBER OF PAGES: vi, 100

Abstract

In this thesis, we explore several stochastic models associated with the two-parameter Poisson-Dirichlet distribution and population genetics. The impacts of mutation, selection and time on the population evolutionary process will be studied by focusing on two aspects of the model: equilibrium and non-equilibrium.

In the first chapter, we introduce relevant background on stochastic genetic models, and summarize our main results and their motivations.

In the second chapter, the two-parameter GEM distribution is constructed from a linear birth process with immigration. The derivation relies on the limiting behavior of the age-ordered family frequencies.

In the third chapter, to show the robustness of the sampling formula we derive the Laplace transform of the two-parameter Poisson-Dirichlet distribution from Pitman sampling formula. The correlation measure of the two-parameter point process is obtained in our proof. We also reverse this derivation by getting the sampling formula from the Laplace transform. Then, we establish a central limit theorem for the infinitely-many-neutral-alleles model at a fixed time as the mutation rate goes to infinity. Lastly, we get the Laplace transform for the selection model from its sampling formula.

In the fourth chapter, we establish a central limit theorem for the homozygosity functions under overdominant selection with mutation approaching infinity. The selection intensity is given by a multiple of certain power of the mutation rate. This result shows an asymptotic normality for the properly scaled homozygosities, resembling the neutral model without selection. This implies that the influence of selection can hardly be observed with large mutation.

In the fifth chapter, the stochastic dynamics of the two-parameter extension of the infinitely-many-neutral-alleles model is characterized by the derivation of its transition function, which is absolutely continuous with respect to the stationary distribution being the two-parameter Poisson-Dirichlet distribution. The transition density is obtained by the expansion of eigenfunctions. Combining this result with the correlation measure in Chapter 3, we obtain the probability generating function of a random sampling from the two-parameter model at a fixed time.

Finally, we obtain two results based on the quasi-invariance of the Gamma process with respect to the multiplication transformation group. One is the quasi-invariance property of the two-parameter Poisson-Dirichlet distribution with respect to Markovian transformation group. The other one is the equivalence between the quasiinvariance of the stationary distributions of a class of branching processes and their reversibility.

Acknowledgements

I wish to express my deepest and heartfelt thanks to my supervisor, Professor Shui Feng, for his guidance and encouragement without which the completion of this thesis would have been impossible. My thanks also go to Professor Fred M. Hoppe and Professor Roman Viveros-Aguilera for their advice and help through the course of my graduate study. In addition, I would like to thank McMaster University for financial support during these years. Last but not the least, I am indebted to my parents for their love and steadfast support.

Declaration of Academic Achievement

This "sandwich thesis" includes three papers completed during my PhD study from 2008 to 2011.

The first paper "A central limit theorem associated with the transformed twoparameter Poisson–Dirichlet distribution" was published in the *Journal of Applied Probability* [67]. This paper demonstrates that the limiting behavior of the twoparameter Poisson-Dirichlet distribution with selection looks like the neutral case as the mutation rate goes to infinity. Specifically we establish a central limit theorem for the homozygosity functions with selection intensity given by a multiple of certain power of the mutation rate. Selection intensities governed by different powers of mutation rate give rise to a phase transition in the limiting distribution.

The second paper "Functional inequalities for the two-parameter extension of the infinitely-many-neutral-alleles diffusion" was published in the *Journal of Functional Analysis* [27]. This paper indicates that the two-parameter extension of the infinitely-many-neutral-alleles diffusion process converges exponentially in entropy to its reversible measure the two-parameter Poisson-Dirichlet distribution by establishing the super-Poincaré and super-log-Sobolev inequalities, which in particular implies the Gross log-Sobolev inequality. My coauthors Professor Feng, Professor Sun, Professor Wang and I are all equal contributors of this paper. More specifically, I made the following contributions: explicitly identifying the transition density function of the two-parameter process and identifying the upper bound of the transition density.

The third paper "The sampling formula and Laplace transform associated with the two-parameter Poisson-Dirichlet distribution" has been accepted by the *Advances in Applied Probability* [68]. We derive the Laplace transform of the two-parameter Poisson-Dirichlet distribution from the Pitman sampling formula. We also use the resulting Laplace transform to obtain the sampling formula. Using the Laplace transform method we further establish the central limit theorem for the infinitemany-neutral-alleles model at a fixed time as mutation goes to infinity. In the end we derive the Laplace transform for the selection model from its sampling formula.

Contents

1	Bac	kground and Statement of Results	1
	1.1	Background	1
	1.2	Summary of Main Results	8
		1.2.1 The Two-parameter GEM Distribution	8
		1.2.2 The Sampling Formula and Laplace Transform	10
		1.2.3 The Transformed Two-parameter Poisson-Dirichlet Distribution	16
		1.2.4 Transition Function of the Two-parameter Model	18
		1.2.5 The Quasi-invariance Property	20
	1.3	Future Work	23
2	The	e Genealogical Structure of the Linear Birth Process with Immi-	
	grat	tion	25
	2.1	Abstract	25
	2.2	Introduction	25
	2.3	Linear Birth Process	27
	2.4	Age-ordered Infinite Alleles Model	29
3	The	e Sampling Formula and Laplace Transform	33
	3.1	Abstract	33
	3.2	Introduction	34
	3.3	Preliminaries and Notations	36
	3.4	The Sampling Formula and Laplace Transform	37
	3.5	Non-equilibrium of the IMNA model	42
		3.5.1 The Laplace Transform of the IMNA Model	42
		3.5.2 CLT for the IMNA Model	47
	3.6	Symmetric Selection	50

4	$Th\epsilon$	Transformed Two-Parameter Poisson-Dirichlet Distribution	54	
	4.1	Abstract	54	
	4.2	Introduction	54	
	4.3	Uniform Integrability	56	
	4.4	The Proof of Theorem	62	
5	Tra	nsition Function of the Two-Parameter Diffusion Process	64	
	5.1	Abstract	64	
	5.2	Introduction	65	
	5.3	Preliminaries	68	
	5.4	Transition Density Function	70	
	5.5	Random Sampling at Time t	76	
6	The	Quasi-invariant Property and Its Application	79	
	6.1	Abstract	79	
	6.2	Introduction	79	
	6.3	Quasi-invariance of the $PD(\alpha, \theta)$ distribution	83	
	6.4	Reversibility of Branching Processes	85	
7	Sun	nmary	93	
R	References			

Chapter 1

Background and Statement of Results

In this chapter, we provide an introduction to the background of stochastic genetic models, and then summarize main results and their motivations. The future work is briefly discussed in the end.

1.1 Background

Population genetics is the study about changes of allele frequency distribution over time under the influences of various forces. It was primarily founded by Fisher, Haldane and Wright in the 1920's and 1930's based on the Mendelian hereditary mechanism. Over the years mathematical modeling has been extensively applied during the development of population genetics. Even though these mathematical models are highly idealized, many of their theoretical predictions on the patterns of genetic variations in actual populations turn out to be consistent with empirical data.

The following terminologies are frequently used in population genetics. Alleles refer to distinct types of each gene. Different alleles may lead to different biological traits. Mutation means a sudden change in the DNA sequence. Selection is described as certain alleles having stronger ability to live and to produce offspring; this ability is called the *fitness* of individual. A mutation that brings no change to the fitness of an individual is called *neutral*. The random change in allele frequencies caused by random sampling of genes from one generation to the next is referred as a *random genetic drift*.

A probability measure-valued stochastic process is commonly used to model the

variation of the frequency distribution of different alleles in the population over time under the influences of mutation, selection, and random genetic drift. Therefore, we first introduce some basic concepts from stochastic process theory.

A discrete time homogeneous Markov chain $\{X_n, n = 0, 1, ...\}$ with finite states $S = \{S_1, \ldots, S_r\}$ is a stochastic process given by transition matrix $P = (p_{ij})_{r \times r}$, whose entry p_{ij} denotes the transition probability from state *i* to state *j* in one step. An initial probability distribution defined on *S* specifies the starting state of the process. Suppose that the law of initial state X_0 is given by a row probability vector *w* with the property wP = w. Then, the probability of being in the various states after *n* steps is $wP^n = w$, and is the same on all steps. The process starting with this method is called a stationary process and *w* is called the stationary distribution or the equilibrium state of the process.

By changing the scale of time and space, many Markov chains can be approximated by diffusion processes $\{X_t, t \ge 0\}$, which is a continuous time Markov process with continuous sample paths and state space S. Given the past history up to time s, the conditional probability

$$P(X_t \in dy | X_s = x, X_{u_i} = x_{u_i}, 0 \le u_i < s, x_{u_i} \in S, \forall i = 1, 2, \ldots)$$

of the process at a future time t is given by the transition function

$$P(X_t \in dy | X_s = x) = P(s, x, t, dy), \quad t > s.$$

This thesis only considers the homogeneous case, where

$$P(s, x, t, dy) = P(0, x, t - s, dy).$$

Hereafter, we always denote the transition probability P(0, x, t, dy) by P(t, x, dy). Generally, diffusion process can be characterized by several tools, such as the Markov semigroup, infinitesimal generator, transition function, martingale problem, etc. The Markov semigroup of the process $\{X_t, t \ge 0\}$ starting from $X_0 = x$ is a family of linear operators $T_t, t \ge 0$ on C(S) the space of continuous functions on S such that

$$T_t f(x) = E_x[f(X_t)] = \int_S f(y) P(t, x, dy).$$

The infinitesimal generator L of the semigroup T_t is given by

$$Lf = \lim_{t \to 0} \frac{T_t f - f}{t}$$

with a properly defined domain. Assume that the law of X_0 is given by a probability measure μ on S satisfying

$$\int_S P(t,x,dy)\mu(dx)=\mu(dy)$$

for any t > 0. Then, the process X_t starting with μ has the same distribution μ at any time t > 0. In this case, we call X_t a stationary process, and μ is the stationary distribution. Equivalently, for the semigroup T_t of the process the stationary measure μ satisfies

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

for $f \in C(S)$. Finally, the process with Markov semigroup T_t is reversible with a probability measure μ such that

$$\int_{S} f(x)T_{t}g(x)\,\mu(dx) = \int_{S} g(x)T_{t}f(x)\,\mu(dx)$$

for all $f, g \in C(S)$. Abundant literature on Markov processes can be found in such as [16] and references therein.

The classic discrete time Markov chain models include the Wright-Fisher, Moran's and Cannings models. Furthermore, we can employ the diffusion approximation technique to the above models and generate continuous time diffusion models. Refer to [16] and [20] for a comprehensive survey of genetic models.

The Wright-Fisher model is considered as the simplest genetic model, in which the randomness comes from random genetic drift during reproduction. It assumes that a population of fixed size N with two alleles evolves according to the binomial sampling between non-overlapping generations. Thus, the frequency of one allele among the nth generation with $n = 0, 1, \ldots$ can be modeled by a Markov chain X_n with state space $\{0, \frac{1}{N}, \ldots, \frac{N-1}{N}, 1\}$ and specific transition matrix. However, the population under this framework will eventually be fixed with one allele. If there is two-way mutation no such fixation occurs. A more general model can be constructed by introducing more than two alleles, including mutation factor and applying diffusion approximation. As a result, we obtain the K-allele Wright-Fisher diffusion model $\{X_t = (X_1(t), X_2(t), \ldots, X_K(t)), t \ge 0\}$ for $K = 2, 3, \ldots$, where $X_i(t)$ denotes the frequency of the *i*th allele in the population at time *t*. Therefore, it is a diffusion process on the K-dimensional simplex

$$\Delta_K = \left\{ x = (x_1, \dots, x_K) : \sum_{i=1}^K x_i = 1, 0 < x_i < 1, i = 1, \dots, K \right\}$$

with infinitesimal generator given by

$$L_K f(x) = \frac{1}{2} \sum_{i,j=1}^K a_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_{i=1}^K b_i(x) \frac{\partial f}{\partial x_i},$$
(1.1)

where $a_{ij}(x) = x_i(\delta_{ij} - x_j)$, $b_i(x) = \sum_{j \neq i} x_j u_{ji} - \sum_{j \neq i} x_i u_{ij}$ and $u_{ij} \geq 0$ is the scaled mutation rate from type *i* to type *j*. With symmetric mutation where u_{ij} is independent of *i* and *j* and equals to u/(K-1) for some $u \geq 0$, the *K*-allele Wright-Fisher diffusion is shown in [66] to be stationary with the *K*-dimensional symmetric Dirichlet distribution with parameter $\epsilon = 2u/(K-1)$. Denote this distribution as $D(\epsilon, \ldots, \epsilon)$. The probability density is given by

$$f(x_1,\ldots,x_K) = \frac{\Gamma(K\epsilon)}{\Gamma(\epsilon)^K} (x_1\cdots x_K)^{\epsilon-1},$$

where $x = (x_1, \ldots, x_K) \in \Delta_K$ and $\Gamma(\cdot)$ is the gamma function given by

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt, \quad z > 0.$$

If selection is introduced into the model (c.f.[17] for the construction of the process), the stationary distribution in a special case is given by

$$f(x_1,\ldots,x_K) = Ce^{sH_2(x)}(x_1\cdots x_K)^{\epsilon-1},$$

where $H_2(x) = x_1^2 + x_2^2 + \cdots + x_K^2$ is called the homozygosity function, $s \in \mathbb{R}$ is the selection intensity and C is the appropriate normalization constant. This special selection assumes that in a diploid population, all homozygotes (pair of alleles of the same type) have equal fitness 1+s and all heterozygotes (pair of alleles with different types) have fitness 1. When s > 0, the homozygote has selective advantage over heterozygote and this kind of selection is called underdominant selection. Otherwise, when s < 0 it is called overdominant selection.

Consider the order statistics $(x_{(1)}, x_{(2)}, \dots, x_{(K)})$ of K-dimensional Dirichlet distribution $D(\epsilon, \dots, \epsilon)$, where we arrange the alleles by the order of their own family sizes. The joint distribution of $(x_{(1)}, x_{(2)}, \dots, x_{(K)})$ is given by the density

$$f(x_{(1)}, x_{(2)}, \dots, x_{(K)}) = \frac{K! \Gamma(K\epsilon)}{\Gamma(\epsilon)^K} (x_{(1)} \cdots x_{(K)})^{\epsilon - 1},$$

where $x_{(1)} \ge x_{(2)} \ge \ldots \ge x_{(K)} \ge 0$ and $\sum x_{(i)} = 1$. Letting $K \to \infty$ in a way such that $K\epsilon \to \theta > 0$, Kingman [43] showed that the distribution of the above order

statistics under the projection onto the infinite-dimensional ordered simplex

$$\nabla_{\infty} = \left\{ (x_1, x_2, \ldots) : x_1 \ge x_2 \ge \cdots \ge 0, \sum_{i=1}^{\infty} x_i = 1 \right\}$$

converges, and named the limiting distribution as the Poisson-Dirichlet distribution with parameter θ , and is denoted by $PD(\theta)$. In population genetics, the parameter θ denotes the suitably normalized mutation rate.

Ethier and Kurtz [15] apply the same limitation procedure to the K-allele diffusion model and obtain the infinitely-many-neutral-alleles (IMNA) model

$$\{X_t = (X_1(t), X_2(t), \ldots), t \ge 0\},\$$

where $X_i(t)$ denotes the frequency of the *i*th largest family at time *t* in the population with infinite alleles. The model is a diffusion process on ∇_{∞} with generator

$$L = \frac{1}{2} \sum_{i,j=1}^{\infty} x_i (\delta_{ij} - x_j) \frac{\partial^2}{\partial x_i \partial x_j} - \frac{\theta}{2} \sum_{i=1}^{\infty} x_i \frac{\partial}{\partial x_i}.$$
 (1.2)

The domain of L is $\mathcal{D}(L) = span\{1, \varphi_2, \varphi_3, \ldots\} \subset C(\nabla_{\infty})$, where $\varphi_n(x) = \sum_{i=1}^{\infty} x_i^n$ is defined on ∇_{∞} and extends continuously to $\overline{\nabla}_{\infty}$, the closure of ∇_{∞} in the product space $[0, 1]^{\infty}$. Hereafter, the topologies of ∇_{∞} and $\overline{\nabla}_{\infty}$ are inherited from $[0, 1]^{\infty}$. The parameter θ denotes the scaled mutation rate. Ethier and Kurtz [15] also prove the convergence of the corresponding invariant measures. This can be seen as a new proof of Kingman's derivation of $PD(\theta)$ in [43]. Equivalently speaking, the Poisson-Dirichlet distribution is the stationary distribution of the IMNA model. Hereafter, IMNA will be used as an acronym for infinitely-many-neutral-alleles.

Consider a population composed of countable types labeled by $\{1, 2, \ldots\}$. Suppose that the proportion of each type is denoted by a random vector (p_1, p_2, \ldots) such that $p_i > 0$ and $\sum p_i = 1$. A size-biased permutation of (p_1, p_2, \ldots) is a random vector $(\tilde{p}_1, \tilde{p}_2, \ldots)$ on the infinite simplex Δ_{∞} with conditional probabilities $P(\tilde{p}_1 = p_i | (p_i)_{i=1}^{\infty}) = p_i$ and for $j \geq 2, n = 1, 2, \ldots$,

$$P\left(\tilde{p}_{j} = p_{n} | (p_{i})_{i=1}^{\infty}, (\tilde{p}_{i})_{i=1}^{j-1}\right) = \frac{p_{n}}{1 - \sum_{i=1}^{j-1} \tilde{p}_{i}} \mathbf{1}_{\{p_{n} \notin \{\tilde{p}_{i}\}_{i=1}^{j-1}\}},$$

where **1** is the indicator function. The above conditional probabilities can be interpreted as follows. The first individual is picked randomly in the population. \tilde{p}_1 represents the proportion of individuals belonging to the type of the first sample. Then all the individuals with the same type of previous sample are removed. Iterations of these two steps generate an infinite sequence $(\tilde{p}_1, \tilde{p}_2, ...)$ corresponding to the frequencies of alleles in the order of their appearances in the sample.

The size-biased permutation of the Poisson-Dirichlet distribution, known as the GEM distribution, has a very simple structure which readily lends itself to calculation. The GEM is defined as below.

Definition 1.1. Consider a vector (V_1, V_2, \ldots) defined by

$$V_1 = U_1, V_n = (1 - U_1) \cdots (1 - U_{n-1})U_n, \ n \ge 2,$$
 (1.3)

where $(U_i)_{i=1}^{\infty}$ are independent identically distributed $Beta(1,\theta)$ random variables with density function $\theta(1-x)^{\theta-1}, 0 < x < 1$. The distribution of $(V_1, V_2, ...)$ is called the GEM distribution with parameter θ denoted by $GEM(\theta)$.

GEM was termed by Ewens [19] after Griffiths, Engen and McCloskey, who contributed to its development and application in the fields of genetics and ecology.

It is easy to verify that the GEM distribution concentrates on the unordered infinite simplex

$$\Delta_{\infty} = \left\{ (x_1, x_2, \ldots) : x_i \ge 0 \ \forall i, \ \sum_{i=1}^{\infty} x_i = 1 \right\}$$

equipped with the topology inherited from $[0, 1]^{\infty}$. The GEM distribution is the sizebiased permutation of the Poisson-Dirichlet, and the Poisson-Dirichlet is the ranked permutation of GEM. This result, due originally to Patil and Taillie [50], is used continuously throughout the literature. Proofs of this result have been provided by Hoppe [40], and more explicitly by Donnelly and Joyce [9].

The scheme (1.3) used to generate a random discrete distribution (V_i) from independent (U_i) is known as a residual allocation model. As a generalization, Perman, Pitman and Yor [51] introduce the two-parameter Poisson-Dirichlet distribution by bringing another parameter $\alpha \in [0, 1)$ into the model.

Definition 1.2. For $0 \le \alpha < 1$ and $\theta > -\alpha$, let $U_k, k = 1, 2, ...$ be a sequence of independent random variables such that U_k has $Beta(1-\alpha, \theta+k\alpha)$ distribution with density function given by

$$f(x) = \frac{\Gamma(\theta + 1 + (k - 1)\alpha)}{\Gamma(1 - \alpha)\Gamma(\theta + k\alpha)} x^{-\alpha} (1 - x)^{\theta + k\alpha - 1}, \ 0 < x < 1.$$

Set (V_1, V_2, \ldots) in the same way as in (1.3). Arrange (V_1, V_2, \ldots) in descending order and denote it as $P = (P_1, P_2, \ldots)$. Then the law of P is called the two-parameter Poisson-Dirichlet distribution denoted by $PD(\alpha, \theta)$. The law of $(V_1, V_2, ...)$ is called the two-parameter GEM distribution denoted by $GEM(\alpha, \theta)$.

Since $\sum_{i=1}^{\infty} V_i = 1$ with probability one, the two-parameter Poisson-Dirichlet distribution is still concentrated on the ordered infinite simplex ∇_{∞} . Pitman [55] showed that $GEM(\alpha, \theta)$ is the only non-degenerate residual allocation model which is invariant under size-biased permutation.

Overall, the Poisson-Dirichlet distribution and its two-parameter generalization have similar structures, including the urn construction in [39] and [25], GEM representation, sampling formula, etc. Carlton [4] obtained a general moments formula for the two-parameter Poisson-Dirichlet distribution and provided several methods for estimating the two parameters θ and α . Readers can also refer to [22] and references therein for a comprehensive study of this family of distributions.

The study of the Poisson-Dirichlet distribution is closely related to population genetics. The IMNA model is often used to serve as the null-hypothesis model in testing the neutral theory of molecular evolution. The literature of this model can be found in [65], [32], [31], [15], etc. In practice, one cannot observe the entire population. Instead, one must make inferences based on information in a finite sample. A basic tool in this statistical analysis is Ewens sampling formula, which assumes stationarity of the IMNA model. It describes the probabilities associated with counts of how many different alleles are observed a given number of times in the sample. The formula was first obtained by Ewens [18].

Recently stochastic dynamics associated with the two-parameter model was discussed in several papers. For example, Bertoin [3] constructed a continuous time Markov chain through an exchangeable fragmentation-coalescence process. The twoparameter Poisson-Dirichlet distribution was shown to be the unique reversible measure of the process. In [28], a class of diffusion processes was constructed from a sequence of independent Wright-Fisher diffusion processes on [0, 1]. The process is reversible with the two-parameter GEM distribution. [52] and [26] used different methods to construct a two-parameter extension of the IMNA model. It is an infinite dimensional symmetric diffusion process with state space ∇_{∞} and generator

$$L_{\theta,\alpha} = \frac{1}{2} \left\{ \sum_{i,j=1}^{\infty} x_i (\delta_{ij} - x_j) \frac{\partial^2}{\partial x_i \partial x_j} - \sum_{i=1}^{\infty} (\theta x_i + \alpha) \frac{\partial}{\partial x_i} \right\}$$
(1.4)

defined on the same domain with (1.2). Its unique reversible measure is given by the $PD(\alpha, \theta)$ distribution.

1.2 Summary of Main Results

In this thesis, we mainly study the impacts of mutation, selection and time on the evolutionary process of the population by focusing on two major aspects of the model: equilibrium and non-equilibrium.

Regarding to the equilibrium, we study the characterization and asymptotic behavior of the stationary distribution associated with several models. We derive the Laplace transform of the two-parameter Poisson-Dirichlet distribution from the Pitman sampling formula, and vice versa. We also study the Laplace transform and asymptotic behavior of the stationary distribution of the selection model as the mutation rate goes to infinity. In the last chapter, we study the quasi-invariance of the $PD(\alpha, \theta)$ distribution. The study of the non-equilibrium aspect is mainly based on the transition function of the process. With the transition density of the IMNA model, we obtain a Gaussian limit theorem for the homozygosites at a fixed time when mutation goes to infinity. We also give the transition density of the twoparameter extension model, and use it to derive the probability generating function of random sampling from the two-parameter model at each time point. In the end, the quasi-invariance of the stationary distribution of a class of branching process with immigration is employed to characterize the reversibility of the process.

Generally, our results illustrate the impact of large mutation and selection on the model at equilibrium and non-equilibrium states. The time effect can be studied from the transition function and further inferred from the model's limiting behaviors at different periods.

In the following sections, we will proceed to summarize these main results and the corresponding motivations.

1.2.1 The Two-parameter GEM Distribution

Our first result presents a construction of the two-parameter GEM distribution from the linear growth model given in Feng and Hoppe [25] by studying the asymptotic behavior of the population at infinite time. The urn construction of the model is used to derive the Pitman sampling formula. This intuitively implies that the model is associated with the GEM distribution. In the second chapter, we give an explicit derivation of this result.

The GEM distribution was previously found to be associated with several genetic models. For instance, Donnelly and Tavaré [10] proved that for a sample of size n whose genealogy is described by a coalescent process with ages, the age-ordered sample frequencies converge in distribution to the GEM as $n \to \infty$. Donnelly and

Tavaré [11] constructed an infinite coalescent with ages in which the frequencies of age-ordered alleles at equilibrium are distributed as $GEM(\theta)$. Hoppe [40] showed that the frequencies of age-ordered alleles are distributed as GEM in any population with the property that the genealogy of a sample of arbitrary size n is described by a coalescent with mutation. Ethier [12] showed that the distribution of the frequencies of the oldest, second-oldest, third-oldest, ... alleles in the stationary IMNA model is given by the GEM distribution. Tavaré [61] showed that the GEM gives the limiting distribution of the age-ordered family proportion of a linear birth process with immigration when time goes to infinity. Our conclusion can be seen as a two-parameter generalization of Tavaré's result, which implies the interconnection between the age-ordered family size process and the genealogical structure of the IMNA model.

The process we consider here was constructed in [25] as follows. Let I(t) be the immigration process such that I(t) is a pure birth process with I(0) = 0 and birth rate

$$\lambda_k = \lim_{h \to 0} \frac{1}{h} P(I(t+h) - I(t) = 1 | I(t) = k),$$

given by $\lambda_k = k\alpha + \theta$, $k \ge 0$. Upon their arrivals, immigrants independently initiate their own families according to a common pure birth process $\{X(t), t \ge 0\}$ with birth rate $b_n = n - \alpha$ starting at X(0) = 1. Therefore, the population is composed of families initiated by immigrants. At time t, denote the size of the *i*th oldest family by $\eta_i(t)$ and N(t) as the total number of individuals in the population with N(0) = 0. Finally we find that the limiting behavior of genealogy structure of the model coincides with the distribution of $\text{GEM}(\alpha, \theta)$, as the following theorem states.

Theorem 1.3.

$$N(t)^{-1}(\eta_1(t), \eta_2(t), \ldots) \to (P_1, P_2, \ldots) \quad a.s. \ as \ t \to \infty,$$

where P_i is the asymptotic fraction of the *i*th oldest family size in the population. Furthermore, $\{P_i\}_{i=1}^{\infty}$ has a joint distribution the same as a size-biased permutation of the two-parameter Poisson-Dirichlet distribution. That is, let $\{U_i\}_{i=1}^{\infty}$ be a sequence of independent random variables such that U_i has $Beta(1 - \alpha, \theta + i\alpha)$ distribution. Then

$$P_i \stackrel{D}{=} (1 - U_1)(1 - U_2) \cdots (1 - U_{i-1})U_i, \ i \ge 2 \quad and \quad P_1 = U_1.$$
(1.5)

Our derivation is through the description of asymptotic behaviors of family sizes and total population size. Since they are all pure linear birth processes, we first show that they achieve infinity at exponential rate as time increases up. Furthermore, we obtain the limiting distribution by properly scaling. Based on these results, we calculate the joint density function for the asymptotic family frequencies, which gives the final conclusion. Lacking of identical distribution property in the two-parameter GEM adds more complexity of the calculation than the one parameter case.

The key factor leading to the conclusion of Theorem 1.3 is the embedded jumping chain, which has the structure of a two-parameter Pólya-like urn structure described in [54] and [56]. The urn scheme, originated from Bayesian statistics, is of significance in deriving the two-parameter sampling formula.

1.2.2 The Sampling Formula and Laplace Transform

In practice, one cannot observe the entire population. Instead, one must make inferences based on information in a finite sample. A basic tool in the statistical analysis is Ewens' sampling formula, which describes the probabilities associated with counts of how many different alleles are observed a given number of times in the sample.

The composition of a sample can be represented in the following way. For a random sample of size n, a vector $\mathbf{A}_n = (a_1, \ldots, a_n)$ called the allele frequency spectrum denotes its allelic partition, in which a_i is the number of alleles appearing in the sample exactly i times. Thus, $a_i \geq 0$ and $\sum_{i=1}^n ia_i = n$. The number of distinct alleles in the sample is given by $k = \sum_{i=1}^n a_i$.

However, statistics inference based on population genetics models is complicated by the randomness of both population and sample. The composition of the current population is just one of many possibilities that arise under the given evolutionary scenario. Given a sample drawn from the population, there are two probability distributions of interest: the conditional probability and unconditional probability of observing the sample. Suppose that all allele frequencies are given by vector $\mathbf{x} =$ (x_1, x_2, \ldots) . In [44] the conditional sampling probability $P(\mathbf{A}_n = (a_1, a_2, \ldots, a_n) | \mathbf{X} =$ \mathbf{x}) is given as multinomial sampling function:

$$P(\mathbf{A}_n = (a_1, a_2, \dots, a_n) | \mathbf{X} = \mathbf{x}) = \frac{n!}{\prod_{i=1}^n (i!)^{a_i}} \sum_{\mathbf{n}} x_1^{n_1} x_2^{n_2} \cdots,$$

where the summation is over the set $\{\mathbf{n} = (n_1, n_2, \ldots) \in \mathbb{Z}_+ \times \mathbb{Z}_+ \times \cdots : \sharp (i : n_i = j) = a_j\}$. It represents the set of allele counts which are consistent with the partition \mathbf{A}_n . When all the allele frequencies are distributed as μ , the unconditional sampling probability is given by

$$P(\mathbf{A}_n = (a_1, a_2, \dots, a_n)) = \int_{\nabla_{\infty}} P(\mathbf{A}_n = (a_1, a_2, \dots, a_n) | \mathbf{X} = \mathbf{x}) \, \mu(d\mathbf{x})$$

When the alleles frequencies distribution μ is the $PD(\theta)$ distribution, the above unconditional probability is given by the Ewens sampling formula

$$P(\mathbf{A}_n = (a_1, a_2, \dots, a_n)) = \frac{n!}{\theta_{(n)}} \prod_{j=1}^n (\frac{\theta}{j})^{a_j} \frac{1}{a_j!},$$
(1.6)

where $\theta_{(n)} = \theta(\theta + 1) \cdots (\theta + n - 1)$. This formula was first obtained by Ewens [18]. It motivates the constructions of several well-known combinatoric structures. For example, Antoniak [1] derived the Ewens sampling formula from the Blackwell-MacQueen description of sampling from a Dirichlet prior distribution. Hoppe [39] devised an urn scheme to obtain (1.6). Furthermore, a genealogical interpretation of Hoppe's Urn model was given by Donnelly [8]. He equates the urn scheme to the construction using the Poisson-Dirichlet paintbox. That is, one can think of colouring balls by dipping a paintbrush at random into a paintbox of which a fraction x_1 of the paint is of one colour, x_2 of a second colour, and so on. Assume that $x = (x_1, x_2, \ldots) \in \nabla_{\infty}$ has the law of the Poisson-Dirichlet distribution. The distribution of the partition induced by the colours of the first n balls can be given by the Ewens sampling formula. Hoppe [40] further explained the relationship between the urn model and the age distribution of the infinitely many alleles diffusion model as well as Kingman's coalescent with mutation.

As a two-parameter generalization of Ewens sampling formula, Pitman [53] derived the Pitman sampling formula corresponding to the two-parameter Poisson-Dirichlet distribution. That is, when $\mathbf{X} \sim PD(\alpha, \theta)$ the sampling probability is given by

$$P(\mathbf{A}_n = (a_1, a_2, \dots, a_n)) = \frac{n!}{\theta_{(n)}} \prod_{l=0}^{k-1} (\theta + l\alpha) \prod_{j=1}^n \frac{(1-\alpha)_{(j-1)}^{a_j}}{(j!)^{a_j} (a_j!)}.$$
 (1.7)

We also find similar urn model for the two-parameter sampling formula in [25] and Blackwell-MacQueen urn scheme in [53]. Under the framework of Kingman's random partitions theory, $PD(\alpha, \theta)$ is the distribution of random mass partition and the Pitman sampling formula (1.7) is the corresponding partition function. Kingman [45] sets up a one-to-one correspondence between the partition function and random mass partition.

In our thesis, we apply the sampling formula (1.7) to derive the Laplace transform of the two-parameter distribution. The Laplace transform as a probability generating functional of random vector X with the $PD(\alpha, \theta)$ distribution is first given by Pitman, Yor [57] by using the α -stable subordinator. It is given as

$$\frac{\lambda^{\theta}}{\Gamma(\theta+1)} \int_0^\infty s^{\theta-1} \mathrm{e}^{-\lambda s} \mathrm{E}[\prod_{i=1}^\infty (1+g(sX_i))] ds$$

for appropriate function g, where $(X_1, X_2, ...)$ has the $PD(\alpha, \theta)$ distribution. The special form of Laplace transform can be seen as a result of Campbell's theorem from the Poisson point process theory (e.g. see [46]). Handa [38] adopts a different approach to get the same Laplace transform through the theory of point process. In addition, as shown in [38] the Laplace transform can be used as a tool to derive several essential results for $PD(\alpha, \theta)$, including the joint probability density, the moment formula, the limiting theorem, etc. In Chapter 3, we use this tool to derive the sampling formula.

Before giving the main result, we first introduce the two-parameter Poisson-Dirichlet point process defined as $\xi := \sum_{i=1}^{\infty} \delta_{X_i}$ in [38] where $(X_i) \sim PD(\alpha, \theta)$. For any positive integer *n*, the *n*th correlation measure of ξ , if it exists, is defined to be a σ -finite Borel measure μ_n such that for any nonnegative measurable function *f* on \mathbf{R}^n

$$\operatorname{E}\left[\sum_{i_1,\ldots,i_n\neq}f(X_{i_1},\cdots,X_{i_n})\right] = \int_{\mathbf{R}^n}f(x_1,\ldots,x_n)\,\mu_n(dx_1\cdots dx_n),$$

where the subscript on the left side denotes that the sum is taken over n-tuple of distinct indices (this notation will be used continuously throughout the thesis). If μ_n has a density with respect to the *n*-dimensional Lebesgue measure, the density is called the *n*th correlation function of ξ . When $\alpha = 0$ the correlation measure of ξ was obtained by Watterson [65] and referred as the multivariate frequency spectra. Using this result Griffiths computed the probability density and marginal distribution for the $PD(\theta)$ in [34]. Regarding the two-parameter point process, the correlation function was first obtained by Handa [38], and was used to obtain the Laplace transform. His derivation is through the size-biased permutation of the two-parameter Poisson-Dirichlet distribution. Here we obtain the correlation functions only based on the Pitman sampling formula (1.7).

Theorem 1.4. Suppose that the sampling formula (1.7), which is equivalent to the following equation

$$\mathbf{E}\left[\sum_{i_1,\dots,i_k\neq} X_{i_1}^{n_1}\cdots X_{i_k}^{n_k}\right] = \frac{\prod_{l=0}^{k-1} (\theta+l\alpha)}{\theta_{(n)}} \prod_{i=1}^k (1-\alpha)_{(n_i-1)},$$
(1.8)

holds for a random element $(X_i)_{i=1}^{\infty} \in \nabla_{\infty}$. Then the kth correlation function of the point process $\xi = \sum_{i=1}^{\infty} \delta_{X_i}$ for each $k = 1, 2, \ldots$ is as follows:

$$q_k(x_1,\ldots,x_k) = c_{k,\alpha,\theta} \prod_{i=1}^k x_i^{-(\alpha+1)} (1 - \sum_{j=1}^k x_j)^{\theta + \alpha k - 1} \mathbf{1}_{\Delta_k}(x_1,\ldots,x_k),$$
(1.9)

where

$$\Delta_k = \left\{ (x_1, \dots, x_k) : x_1 \ge 0, \dots, x_k \ge 0, \sum_{i=1}^k x_i \le 1 \right\}$$

and

$$c_{k,\alpha,\theta} = \prod_{i=1}^{k} \frac{\Gamma(\theta + 1 + (i-1)\alpha)}{\Gamma(1-\alpha)\Gamma(\theta + i\alpha)}.$$

It is known that correlation functions appear in the expansion of the "probability generating function" of a random point process $\sum \delta_{X_i}$. Therefore, from the above correlation functions we derive the Laplace transform of the two-parameter Poisson-Dirichlet point process, which is the same as Handa's result [38].

Theorem 1.5. Suppose that $g: (0, \infty) \to \mathbb{C}$ is a measurable function such that

$$\lambda_{\alpha}(g) := \inf \left\{ \lambda > 0 : \int_0^\infty \frac{\mathrm{e}^{-\lambda z}}{z^{\alpha+1}} |g(z)| \, dz < \infty \right\} < \infty.$$

Put

$$\lambda_{\alpha}^{*}(g) = \inf \left\{ \lambda > \lambda_{\alpha}(g) : \frac{C_{\alpha}}{\lambda^{\alpha}} \int_{0}^{\infty} \frac{\mathrm{e}^{-\lambda z}}{z^{\alpha+1}} |g(z)| \, dz \in \mathrm{C} \setminus [1,\infty) \right\},\,$$

so that in particular $\lambda_0^*(g) = \lambda_0(g)$. The correlation measure of $\xi = \sum_{i=1}^{\infty} \delta_{X_i}$ is given by (1.9). Then

$$\operatorname{E}[\prod_{i=1}^{\infty} (1 + |g(sX_i)|)] < \infty$$

a.e. s > 0, and for $\lambda > \lambda^*_{\alpha}(g)$

$$\frac{\lambda^{\theta}}{\Gamma(\theta+1)} \int_0^\infty s^{\theta-1} e^{-\lambda s} \left(\operatorname{E}\left[\prod_{i=1}^\infty (1+g(sX_i)) - 1\right] \right) ds \\ = \begin{cases} \frac{1}{\theta} \exp(\theta \int_0^\infty \frac{e^{-\lambda z}}{z} g(z) \, dz) - \frac{1}{\theta}, & \alpha = 0, \theta > 0, \\ \frac{1}{\theta} (1 - \frac{C_\alpha}{\lambda^\alpha} \int_0^\infty \frac{e^{-\lambda z}}{z^{\alpha+1}} g(z) \, dz)^{-\frac{\theta}{\alpha}} - \frac{1}{\theta}, & 0 < \alpha < 1, \theta \neq 0 \\ -\frac{1}{\alpha} \log(1 - \frac{C_\alpha}{\lambda^\alpha} \int_0^\infty \frac{e^{-\lambda z}}{z^{\alpha+1}} g(z)) \, dz), & 0 < \alpha < 1, \theta = 0. \end{cases}$$

Inversely, we utilize the Laplace transform of the $PD(\alpha, \theta)$ point process to obtain the sampling formula (1.7).

Corollary 1.6. Suppose that the above Laplace transform exists for sufficiently large $\lambda > 0$, then the sampling formula corresponding to $\mathbf{X} = (X_1, X_2, \ldots) \in \nabla_{\infty}$ coincides with the Pitman sampling formula (1.7).

Given the alleles frequencies with the $PD(\theta)$ distribution, Joyce et al. [41] obtain a Gaussian limit distribution of properly scaled homozygosity functions as the mutation rate θ goes to infinity. Their result further shows that the conditional probability distribution can be approximated by the unconditional probability distribution if θ is large. Here we study the limiting behavior of homozygosity functions associated with the IMNA model (1.2) at a fixed time as the mutation rate θ goes to infinity. We find a similar Gaussian limit behavior of properly scaled homozygosity functions comparing with Joyce et al.'s result at equilibrium state. This implies that the distribution of the IMNA model at any finite time is similar to the equilibrium state when mutation is quite strong.

Theorem 1.7. Consider the IMNA model

$$X(t) = (X_1(t), X_2(t), \ldots) \in \nabla_{\infty}$$

with generator given by (1.2) and starting point $x = (x_1, x_2, ...) \in \nabla_{\infty}$. For any time t > 0, set

$$W_p(t) = \sqrt{\theta} \left(\frac{\theta^{p-1}}{(p-1)!} H_p(t) - 1\right),$$

where $H_p(t) = \sum_{i=1}^{\infty} X_i(t)^p$ denotes the pth homozygosity at time t for p = 2, 3, ...Then as $\theta \to \infty$, $W_p(t)$ converges in law to a normal random variable with mean 0 and variance $\sigma_p^2 = \Gamma(2p)/\Gamma(p)^2 - p^2$, which is strictly positive.

The result is proved through the convergence of the characteristic function. The key element in the proof is the explicit expression of the Laplace transform of X(t), which is obtained from the transition density of X(t) and the correlation measure.

The other model considered in the third chapter is a non-neutral case involving a special case of selection. As shown previously in the finite alleles model with selection, the stationary distribution is absolutely continuous with respect to the neutral one. In [17] a similar density is also found for an infinite alleles model with selection. The density is given under the Poisson-Dirichlet distribution by

$$e^{sH_2(\mathbf{X})}/C, \qquad \mathbf{X} = (X_1, X_2, \ldots) \in \nabla_{\infty},$$

where $H_2(\mathbf{X}) = \sum X_i^2$ denotes the homozygosity function and the selection intensity s is any real number. The Radon-Nikodym derivative can be generalized to

$$\frac{PD^{s,q}(\theta)}{PD(\theta)}(d\mathbf{X}) = e^{sH_q(\mathbf{X})}/C, \qquad \mathbf{X} = (X_1, X_2, \ldots) \in \nabla_{\infty},$$

where $H_q(\mathbf{X}) = \sum X_i^q$, q = 2, 3, ... denotes the *q*th order homozygosity, *s* is the selection intensity and *C* is the normalization constant. Under the resulting distribution $PD^{s,q}(\theta)$, Handa [37] obtains the corresponding sampling formula as

$$P_{\theta,s,q}(\mathbf{A}_n = \mathbf{a}) = \mathbf{E}[e^{sH_q(\mathbf{X})}\psi_{\mathbf{n}}(\mathbf{X})]/\mathbf{E}[e^{sH_q(\mathbf{X})}], \qquad (1.10)$$

where

$$\psi_{\mathbf{n}}(\mathbf{x}) = \frac{n!}{n_1! \cdots n_k! a_1! \cdots a_n!} \sum_{i_1, \dots, i_k \neq} x_{i_1}^{n_1} \cdots x_{i_k}^{n_k}, \qquad (1.11)$$

 $\mathbf{n} = (n_1, \ldots, n_k)$ is a partition of $n \ge 1$, $a_i = \sharp\{j : n_j = i\}$ and the sum is taken over *n*-tuple of distinct indices. The expectation is taken under the $PD(\theta)$ distribution. The Laplace transform in this case is computed as follows.

Theorem 1.8. Suppose that g(z) is a measurable function such that

$$\lambda(g) := \inf\{\lambda > 0 : \int_0^\infty \frac{e^{-\lambda z}}{z} |g(z)| \, dz < \infty\} < \infty,$$

and the random element $(X_i)_{i=1}^{\infty}$ has the $PD^{s,q}(\theta)$ distribution. Then

$$\int_0^\infty e^{-\lambda t} t^{\theta-1} \mathbf{E}[\prod_{i=1}^\infty (1+g(tX_i))] dt = \frac{\Gamma(\theta)}{\lambda^\theta} + C^{-1} \sum_{k=1}^\infty \frac{\theta^k}{k!} M_k,$$

where

$$J_{l}(t) = \int_{\Delta_{l}} \prod_{j=1}^{l} (e^{st^{q}x_{j}^{q}} - 1)(1 - \sum_{\beta=1}^{l} x_{\beta})^{\theta-1} \frac{dx_{1} \cdots dx_{l}}{x_{1} \cdots x_{l}},$$
$$F_{s}(t) = \sum_{l=0}^{\infty} \frac{\theta^{l}}{l!} J_{l}(t),$$

and

$$M_{k} = \int_{(0,\infty)^{k}} \prod_{j=1}^{k} (g(x_{j})x_{j}^{-1}e^{-\lambda x_{j}}) dx_{1} \cdots dx_{k} \int_{0}^{\infty} e^{-\lambda t} t^{\theta-1} \prod_{j=1}^{k} e^{\frac{sx_{j}^{q}}{(t+\sum_{j=1}^{k} x_{j})^{q}}}$$
$$F_{s}\left(\frac{t}{t+\sum_{j=1}^{k} x_{j}}\right) dt.$$

1.2.3 The Transformed Two-parameter Poisson-Dirichlet Distribution

Motivated by the above selection density, we apply it to generate a transformed two-parameter Poisson-Dirichlet Distribution. Define the two-parameter Poisson-Dirichlet distribution with selection denoted by $PD^{\sigma}(\alpha, \theta)$ with the Radon-Nikodym derivative

$$\frac{PD^{\sigma}(\alpha,\theta)}{PD(\alpha,\theta)}(d\mathbf{x}) = \frac{e^{-\sigma H_m(\mathbf{x})}}{C(\sigma,\theta,\alpha)},$$

where $H_m(\mathbf{x}) = \sum_{i=1}^{\infty} x_i^m$, $\mathbf{x} = (x_1, x_2, \ldots) \in \nabla_{\infty}$, $m = 2, 3, \ldots$ denotes the *m*th order homozygosity, $\sigma > 0$ and $C(\sigma, \theta, \alpha) = \int_{\nabla_{\infty}} e^{-\sigma H_m(\mathbf{x})} PD(\alpha, \theta)(d\mathbf{x})$ is the normalization constant. Though the application of the two-parameter Poisson-Dirichlet distribution in population genetics is not found yet, for convenience we still call $PD^{\sigma}(\alpha, \theta)$ the two-parameter Poisson-Dirichlet distribution with selection.

The property we want to investigate is the limiting behavior of $PD^{\sigma}(\alpha, \theta)$ as the parameter θ goes to infinity with fixed α . The motivation of studying asymptotic behavior for large θ is from Gillespie's work [30], where he considered various models with fixed mutation rate and increased population size. This corresponds to the scaled mutation rate $\theta = 4Nu$ going to infinity in the IMNA model, where N is the effective population size and u is the individual mutation rate. For the Poisson-Dirichlet distribution with parameter θ , there are numerous results concerning the limiting behavior for large θ such as a limit theorem of Griffiths [33], the associated large deviation estimates shown by Dawson and Feng [7], the moderate principle by Feng and Gao [23] and the central limit theorem obtained by Joyce, Krone and Kurtz [41].

Due to the similarity between the Poisson-Dirichlet distribution and its twoparameter counterpart, it is natural to expect similar limiting behaviors between these two models. Such results include the large deviation principle in [21], the moderate principle in [24] and central limit theorem for the homozygosity functions in [38].

On the other hand, Gillespie [30] presented simulations for several models including an infinite alleles model with selective overdominance. He observed that the asymptotic behavior in non-neutral case looks like that in neutral model as the selection intensity and mutation rate get large with the same rate. Therefore, in [42] a rigorous theoretical study was carried out for the model with overdominant selection when the selection intensity and mutation rate get large together through Gaussian fluctuations. As a result, the central limit theorem obtained in [41] can be seen as a special case of the conclusion in [42] with selection intensity being zero. To study the impact of selection in the two-parameter model, we establish a central limit theorem for the transformed two-parameter Poisson-Dirichlet distribution $PD^{\sigma}(\alpha, \theta)$. Our result indicates similar limiting behaviors between the selection model and the neutral one $PD(\alpha, \theta)$ for large mutation.

The neutral case was considered by Handa [38]. He studied the asymptotic behavior of $PD(\alpha, \theta)$ for large θ and fixed α as follows. For p = 2, 3, ... and $X = (X_1, X_2, ...) \sim PD(\alpha, \theta)$, define

$$Z_{p,\theta} = \sqrt{\theta} \left(\frac{\Gamma(1-\alpha)}{\Gamma(p-\alpha)} \theta^{p-1} H_p(X) - 1 \right), \qquad (1.12)$$

where $H_p(X)$ denotes the *p*th homozygosity of X. A central limit theorem for $(Z_{p,\theta})_{p=2}^{\infty}$ was established in [38].

Lemma 1.9 ([38]). Let $(Z_{2,\alpha}, Z_{3,\alpha}, ...)$ be multivariate normal with mean 0 and the covariance of $Z_{i,\alpha}$ and $Z_{j,\alpha}$ given by

$$\frac{\Gamma(1-\alpha)\Gamma(i+j-\alpha)}{\Gamma(i-\alpha)\Gamma(j-\alpha)} + \alpha - ij.$$
(1.13)

Then, as $\theta \to \infty$, we have

$$(Z_{2,\theta}, Z_{3,\theta}, \ldots) \Rightarrow (Z_{2,\alpha}, Z_{3,\alpha}, \ldots),$$

where, here and hereafter \Rightarrow denotes convergence in distribution.

Similarly, for $Y = (Y_1, Y_2, \ldots) \sim PD^{\sigma}(\alpha, \theta)$, where $\sigma = c\theta^{m-\frac{1}{2}-\beta}$, $m = 2, 3, \ldots, \beta \geq 0$ and c > 0 is a constant, we define

$$Z_{p,\theta}^{\beta} = \sqrt{\theta} \left(\frac{\Gamma(1-\alpha)}{\Gamma(p-\alpha)} \theta^{p-1} H_p(Y) - 1 \right).$$
(1.14)

This is our final conclusion.

Theorem 1.10. Suppose that $Y = (Y_1, Y_2, ...) \sim PD^{\sigma}(\alpha, \theta)$ and the selection intensity σ is given by $c\theta^{m-\frac{1}{2}-\beta}$, $\beta \geq 0$, where c > 0 is a constant and m = 2, 3, ... Let $(Z_{2,\theta}^{\beta}, Z_{3,\theta}^{\beta}, ...)$ be defined as (1.14) and $(Z_{2,\alpha}, Z_{3,\alpha}, ...)$ given as in the above lemma. As $\theta \to \infty$,

$$(Z_{2,\theta}^{\beta}, Z_{3,\theta}^{\beta}, \ldots) \Rightarrow \begin{cases} (Z_{2,\alpha}^{*}, Z_{3,\alpha}^{*}, \ldots), & \text{if } \beta = 0, \\ (Z_{2,\alpha}, Z_{3,\alpha}, \ldots), & \text{if } \beta > 0, \end{cases}$$

where $Z_{p,\alpha}^* = Z_{p,\alpha} - cb_p$, $b_p = (1 - \alpha)_{(p-1)} \operatorname{cov}(Z_{p,\alpha}, Z_{m,\alpha})$ and for p = 2, 3, ...

$$(1-\alpha)_{(p-1)} = \frac{\Gamma(p-\alpha)}{\Gamma(1-\alpha)} = (1-\alpha)\cdots(p-1-\alpha).$$

The limiting distribution shows a phase transition depending on the relative strength of the selection intensity and the mutation rate. The proof is based on a fluctuation argument through the previous result in Lemma 1.9. The main part of proof is devoted to the uniform integrability of $\exp\{-tZ_{p,\theta}\}$. Instead of the common tool the two-parameter GEM, we employ a representation of the $PD(\alpha, \theta)$ distribution in [53] and [57] through $PD(\theta)$ and an i.i.d. sequence of $PD(\alpha, 0)$ random variables.

1.2.4 Transition Function of the Two-parameter Model

Besides the semigroup and infinitesimal generator, the transition function is one of the common tools to characterize a stochastic process. For the IMNA model (1.2), Griffiths [32] first obtained the transition function as a limit of transition densities of finite-dimensional Wright-Fisher diffusions, based on the symmetric expansion with a series of orthogonal polynomials. Ethier [13] rederived the same transition density by an expansion of eigenfunctions, which makes the derivation more transparent.

Even recently, Petrov [52] and Feng and Sun [26] have used different approaches to develop the two-parameter extension of the IMNA diffusion model $X_{\theta,\alpha}(t)$ with generator (1.4). It is reversible with respect to the two-parameter Poisson-Dirichlet distribution. Petrov [52] identified the set of eigenvalues for the generator (1.4). It is the same as that for the IMNA model (1.2). But we are still unable to find a proper finite-dimensional approximation of the two-parameter model like the IMNA model does. Therefore, we derive the transition density of the two-parameter model (1.4) via the eigenfunction expansion as [13]. The result is given as follows. Set $\lambda_m = m(m + \theta - 1)/2, m = 2, 3, \ldots$ and

$$\psi_{\mathbf{n}}(x) = \frac{n!}{n_1! \cdots n_k! a_1! \cdots a_n!} \sum_{i_1, \dots, i_k \neq} x_{i_1}^{n_1} \cdots x_{i_k}^{n_k},$$

where $\mathbf{n} = (n_1, \ldots, n_k)$ is a partition of $n \ge 1$, $a_i = \sharp\{j : n_j = i\}$ and the sum is taken over *n*-tuple of distinct indices.

Theorem 1.11. The transition function of the process $X_{\theta,\alpha}(t) \in \nabla_{\infty}$ with the generator (1.4) has a density function p(t, x, y) with respect to its invariant measure $PD(\alpha, \theta)$ given by

$$p(t, x, y) = 1 + \sum_{m=2}^{\infty} e^{-\lambda_m t} q_m(x, y) \qquad x, y \in \nabla_{\infty},$$
(1.15)

where

$$q_m(x,y) = \frac{2m-1+\theta}{m!} \sum_{n=0}^m (-1)^{m-n} \binom{m}{n} (n+\theta)_{(m-1)} P_n(x,y),$$

where $P_n(x, y)$ is defined as follows.

$$P_n(x,y) = \sum_{\vec{n} \in \mathbb{N}_{\infty} : |\vec{n}| = n} \frac{\psi_{\vec{n}}(x)\psi_{\vec{n}}(y)}{PSF(\vec{n})},$$

where \vec{n} is taken over all the partitions of integer n, and $PSF(\vec{n})$ is the Pitman sampling formula (1.7). Moreover, there exists a constant c > 1 such that

$$p(t, x, y) \le ct^{c(\log t)/t}, \quad t > 0, x, y \in \nabla_{\infty}.$$
(1.16)

Since the main idea of proof is similar to that in [13], we focus only on the derivations of results that require additional efforts due to the introduction of the additional parameter α . We also obtain a uniform upper bound for the transition density function, which leads to explicit super-Poincar—/log-Sobolev inequalities for the two-parameter diffusion as [27]. Recall the work in [59] where it was shown that a log-Sobolev inequality holds for the Fleming-Viot process (FV) with parent independent mutation if and only if the number of alleles or type space S is finite. The invalidity of the log-Sobolev inequality indicates the extreme singularity in the Fleming-Viot process when there are infinite number of alleles. The Fleming-Viot process we monitor both the type and frequency of alleles. The IMNA model can be derived from the Fleming-Viot process by dropping off the type of alleles and looking at atom masses which are in decreasing order. It can be inferred that the singularity is reduced by this mapping from the validity of a super log-Sobolev inequality for the two-parameter unlabeled infinite-alleles diffusion.

With the correlation function obtained previously and the transition function, we compute the probability generating function of random sampling from $X_{\alpha,\theta}(t)$. This result can be seen as a two-parameter generalization of Theorem 1 in [31]. There, the author's approach is based on the fact that the IMNA diffusion process can be approximated by the finite-alleles Wright-Fisher diffusion model. However, a finitedimensional approximation of the two-parameter model has not yet been found. That is why we turn to the correlation measure of the two-parameter point process. The difference between the sampling formula at a fixed time and at the equilibrium state can be found from our result. **Theorem 1.12.** Consider the two-parameter diffusion process $X_{\alpha,\theta}(t)$ with generator $L_{\alpha,\theta}$ (1.4) and initial value $x = (x_1, x_2, \ldots) \in \nabla_{\infty}$. The probability generating function

$$\mathbb{E}\left[\prod_{i=1}^{n} u_i^{a_i} | X(0) = x\right]$$

of a sample $\mathbf{a} = (a_1, a_2, \dots, a_r)$ of size r at time t is the coefficient of φ^r in

$$G_0 + \sum_{m=2}^{\infty} \frac{2m - 1 + \theta}{m!} e^{-\lambda_m t} \sum_{n=0}^m (-1)^{m-n} \binom{m}{n} (n+\theta)_{(m-1)} G_n,$$

where

$$G_0 = G_1 = \frac{r!}{\theta_{(r)}} \left(1 - \alpha \sum_{l=1}^r (1 - \alpha)_{(l-1)} u_l \varphi^l / l! \right)^{-\theta/\alpha},$$

$$G_{m} = r! \frac{\theta_{n}}{\theta_{(n+r)}} \sum_{\boldsymbol{n}=(n_{1},\dots,n_{k})} \psi_{\boldsymbol{n}}(x) \left(1 - \alpha \sum_{l=1}^{r} (1-\alpha)_{(l-1)} u_{l} \varphi^{l} / l!\right)^{(-k+\frac{\theta}{\alpha})}$$
$$\prod_{i=1}^{k} \left(1 + \sum_{l=1}^{r} (n_{i} - \alpha)_{(l)} \varphi^{l} u_{l} / l!\right).$$

1.2.5 The Quasi-invariance Property

On a probability space S with measure μ , the quasi-invariance of μ with respect to a measurable transformation group G on S is defined as follows: given any mapping $T \in G$, the image measure $T\mu$ and the original measure μ are mutually absolutely continuous. A classical example is the Girsanov formula for additive perturbations of Brownian motion.

Our study is motivated by the quasi-invariance of the gamma process, which is a random discrete measure $\eta = \sum X_i \delta_{S_i}$, where S_i are i.i.d. random variables on a compact Polish space S with common distribution ν and $(X_i)_{i=1}^{\infty}$ are distinct points in a Poisson point process with mean measure $\theta x^{-1} e^{-x} dx$, x > 0. Denote by P_{ν} the law of gamma process.

The gamma process has been studied extensively by several authors, including Tsilevich el al. [62], [63], Handa [35] and Yor [69]. A strong parallelism is shown between the gamma process and Brownian motion from various aspects. Tsilevich el al. show that stable process can be considered as a deformation from the Brownian motion to the gamma process. Both of them correspond to the extreme values of the parameter $\alpha \in [0, 2]$. They also prove the quasi-invariance of the gamma process with respect to the linear transformation group defined by

$$M_a\eta(ds) = a(s)\eta(ds),$$

where the multiplicator a is nonnegative measurable function on S. The quasiinvariance of P_{ν} means that the induced measure $M_a P_{\nu}$ and the original measure P_{ν} are equivalent.

The normalized gamma process $\bar{\eta}$ is the so-called Dirichlet process defined as

$$\bar{\eta} = \eta(ds)/\eta(S) = \sum \frac{X_i}{\sum X_j} \delta_{S_i}$$

where $\eta = \sum X_i \delta_{S_i}$ is the gamma process. The normalized process $\bar{\eta}$ is independent of the total mass $\eta(S)$. Using this property, Handa [35] shows that the quasi-invariance is preserved for the Dirichlet process under the induced transformation group on the space of probability measure. The image of probability measure $\bar{\eta}$ under the induced mapping S_a is given by

$$S_a\bar{\eta}(ds) = a(s)\eta(ds)/\langle a,\eta\rangle,$$

where $\langle a, \eta \rangle$ denotes the integration of a Borel measurable function a with respect to η .

M.-K. von Renesse et al. [64] unify and extend the quasi-invariance properties of the gamma and Dirichlet processes to a class of subordinators with respect to a large class of non-linear sample path transformations. The quasi-invariance property can also yield easily an integration by parts formula on the path space.

The $PD(\theta)$ distribution is the law of the ordered atom masses of Dirichlet process denoted by

$$P = (P_1, P_2, \ldots) = \Phi(\frac{X_1}{\sum X_j}, \frac{X_2}{\sum X_j}, \ldots),$$

where Φ is a map that arranges the coordinates in non-increasing order and (X_i) are the atom masses of the gamma process.

Tsilevich et al. [63] use this property to show the quasi-invariance of the Poisson-Dirichlet distribution based on the induced Markovian transformations S_a on the ordered atom masses $P = (P_1, P_2, \ldots) \in \nabla_{\infty}$ of the Dirichlet process. The mapping is as follows

$$S_a P = \Phi(a(S_1)P_1/\rho, a(S_2)P_2/\rho, \ldots),$$

where $\rho = \sum a(S_i)P_i$.

We note that the $PD(\alpha, 0)$ distribution is invariant under the above transformation S_a for certain multiplicator a in [2]. In Chapter 6, we generalize the quasiinvariance property to the two-parameter Poisson-Dirichlet distribution under the transformation group S_a . This also allows us to treat the previous results about $PD(\theta)$ and $PD(\alpha, 0)$ within a unified framework.

Theorem 1.13. Suppose that $\{S_i, i = 1, 2, ...\}$ is a sequence of i.i.d. ν_0 distributed random variables on S and a is is a bounded nonnegative Borel measurable functions on S with $\inf_s a(s) > 0$. Independently, set $P = (P_i)_{i=1}^{\infty} \sim PD(\alpha, \theta)$ and $S_a P = \Phi(a(S_1)P_1/\rho, a(S_2)P_2/\rho, ...),$ where $\rho = \sum a(S_i)P_i, \Phi$ is a map that arranges the coordinates in non-increasing order. Then the two-parameter Poisson-Dirichlet distribution $PD(\alpha, \theta)$ with $\theta \neq 0$ is quasi-invariant under the operator S_a and

$$\frac{S_a P_{\alpha,\theta}}{P_{\alpha,\theta}}(dp) = \langle a^{\alpha}, \nu_0 \rangle^{-\theta/\alpha} \frac{1}{\Gamma(\theta)} \int_0^\infty \sigma^{\theta-1} \prod_{i=1}^\infty E_{\nu_0^{\alpha}} [e^{-\sigma a^{-1}(X_i)p_i}] d\sigma.$$

Our argument is based on the α -stable subordinator whose atom masses after normalization distribute as $PD(\alpha, 0)$. Under the transformation M_a , we found that the distribution of normalized atom masses is invariant, while the location of atom is shifted with a density function. The proof also utilizes the absolute continuity of $PD(\alpha, \theta)$ with respect to $PD(\alpha, 0)$. The density we get covers the previous two cases with α or θ being zero. This also provides an alternative perspective to look at the result obtained by Arguin [2] for $PD(\alpha, 0)$.

Another application of the quasi-invariance property is to show the reversibility for a class of measure-valued branching process with immigration. By reversibility it means that

$$\int \Phi L \Psi d\Pi = \int \Psi L \Phi d\Pi, \qquad \Phi, \Psi \in \mathcal{D}(L),$$

where L is the generator of process and Π is the stationary distribution.

For finite-dimensional diffusion processes, conditions for the process to be reversible have been studied intensively such as [47] and [49] through integration by parts. The reversible distributions are identified by potential function whose gradient coincides with the negative of the drift. Previously Handa [36] considered the same kind of problem for an infinite-dimensional diffusion process, the Fleming-Viot process. From his result, we know that the condition for the Fleming-Viot process with parent-independent mutation to be reversible can be given by the quasi-invariance of Dirichlet process, which is the stationary distribution of the process.

We also note that for the gamma process, the density given by the quasi-invariance property corresponds to the exponential martingale in [69]. Stannat [60], associated

the quasi-invariance of the gamma process with an analogue of the Cameron-Martin-Girsanov theorem for the classical Wiener measure.

Here, we use this quasi-invariance property with respect to the multiplication group to characterize the reversibility of a class of measure-valued branching processes. The quasi-invariance property is the behavior of the stationary distribution under the mapping over the state space. On the other hand, reversibility can be seen the behavior under the semigroup. Thus, our conclusion implicitly associates these two mappings for the stationary distribution.

Theorem 1.14. Suppose the branching process with immigration $\{\nu_t, t \ge 0\}$ is defined by generator L,

$$LF(\mu) = \int_{S} \mu(dx)a(x)\frac{\delta^{2}F(\mu)}{\delta\mu(x)^{2}} + \int_{S} (\nu(dx) - \mu(dx)b(x))\frac{\delta F(\mu)}{\delta\mu(x)},$$
(1.17)

for suitable F. Let Π be a Borel probability measure on $\mathcal{M}(S)$ the set of finite positive Borel measures on S. Then L is reversible with respect to Π if and only if Π is quasi-invariant w.r.t. the transformation S_f where $S_f\mu(ds) = e^f\mu(ds)$, where f is a bounded measurable function on S. It means that the image measure $S_f\Pi$ is absolutely continuous with respect to Π . The density is given by $e^{\Lambda(f,\mu)}$ where

$$\Lambda(f,\mu) = \int_0^1 \langle f/a, c(S_{uf}(\mu)) \rangle du,$$

and

$$\langle f, c(\mu) \rangle = \langle f, \nu \rangle - \langle bf, \mu \rangle.$$

This result is similar to Handa's [36] for the Fleming-Viot process. When a and b are both constants in generator L (1.17), the gamma process is the reversible measure of the process. Since the quasi-invariance of Dirichlet process is resulted from the gamma process, our conclusion can be seen as an analogy relationship of the reversibility between the branching process and Fleming-Viot process. These two processes are closely connected in several ways.

1.3 Future Work

The study of the two-parameter Poisson-Dirichlet distribution is in the middle of rapid development over the last few years. There are still a number of problems remaining unsolved for the two-parameter model. A particular example is the construction of the two-parameter measure-valued process generalizing the Fleming-Viot process with parent independent mutation. The immediate projects will include the construction of this model and finding more applications of the two-parameter related models in other areas.

Chapter 2

The Genealogical Structure of the Linear Birth Process with Immigration

2.1 Abstract

In this chapter, we give an explicit construction of the two-parameter GEM from a linear birth process with immigration given in Feng and Hoppe [25]. Immigrants arrive according a linear birth process and produce their own offsprings. Based on some preliminary results on linear birth process, we obtain the limiting distribution of scaled family and total population sizes as time goes to infinity. In the end, we calculate the joint probability density function for the limiting age-ordered family proportions, which coincides with the two-parameter GEM distribution.

2.2 Introduction

A continuous time Markov chain is a process $\{X(t), t \ge 0\}$ with values in $\mathbb{N} = \{0, 1, 2, ...\}$ such that transition function satisfying the following Markov property:

$$P(X(t+s) = i | X(s) = j, X(u) = x(u), 0 < u < s) = P(X(t+s) = i | X(s) = j)$$

for all $t, s \ge 0$ and all states i, j, x(u). The process is homogeneous such that the transition matrix P(t) consists of

$$P_{ij}(t) = P(X(t+s) = j | X(s) = i),$$

and P(0) = I. The process is usually characterized by a matrix $Q = (q_{ij}), i, j \in \mathbb{N}$ consisting of transition rates q_{ij} defined by

$$P(X(t+h) = j|X(t) = i) = q_{ij}h + o(h), \ i \neq j.$$

The row sum of a conservative Q-matrix is zero, yielding

$$q_{ii} = -\sum_{j \neq i} q_{ij}.$$

A pure birth process with birth rate ℓ_i is a continuous time homogenous Markov chain with the conservative *Q*-matrix consisting of elements $q_{i,i+1} = \ell_i$ and other off-diagonal elements being 0 for i = 0, 1, 2, ...

Here, we associate a pure birth process with the two-parameter GEM distribution defined as follows. For $0 \leq \alpha < 1$ and $\theta > -\alpha$, let $U_k, k = 1, 2, ...$ be a sequence of independent random variables such that U_k has $Beta(1 - \alpha, \theta + k\alpha)$ distribution with density function given by

$$f(x) = \frac{\Gamma(\theta + 1 + (k - 1)\alpha)}{\Gamma(1 - \alpha)\Gamma(\theta + k\alpha)} x^{-\alpha} (1 - x)^{\theta + k\alpha - 1}, \ 0 < x < 1.$$

Set $(V_1, V_2, ...)$ as

$$V_1 = U_1, V_n = (1 - U_1) \cdots (1 - U_{n-1})U_n, \ n \ge 2.$$
 (2.1)

The random points (V_1, V_2, \ldots) are in the unordered infinite simplex

$$\Delta_{\infty} = \left\{ (x_1, x_2, \ldots) : x_i \ge 0 \ \forall i, \sum_{i=1}^{\infty} x_i = 1 \right\}$$

equipped with topology inherited from $[0, 1]^{\infty}$. Arrange $(V_1, V_2, ...)$ in descending order and denote it as $P = (P_1, P_2, ...)$. Then, the law of P is called the two-parameter Poisson-Dirichlet distribution denoted by $PD(\alpha, \theta)$. The law of $(V_1, V_2, ...)$ is called the two-parameter GEM distribution denoted by $GEM(\alpha, \theta)$. Since the infinitedimensional ordered simplex

$$\nabla_{\infty} = \left\{ (x_1, x_2, \ldots) : x_1 \ge x_2 \ge \cdots \ge 0, \sum_{i=1}^{\infty} x_i = 1 \right\}$$

is contained in Δ_{∞} , $PD(\alpha, \theta)$ can be thought of as a probability on Δ_{∞} which concentrates on ∇_{∞} .
The two-parameter GEM distribution is the size-biased permutation of the twoparameter Poisson-Dirichlet distribution, and the two-parameter PD distribution is the ranked permutation of GEM.

Tavaré [61] shows that the GEM distribution where $\alpha = 0$ gives the asymptotic family frequencies of an age-ordered infinite alleles model, which is a pure birth process with immigration. In particular, we notice that in Tavaré's model the jump chain embedded corresponds to the Hoppe's Urn model [39], which leads to Ewens sampling formula (1.6), while the two-parameter urn construction [25] implied the two-parameter sampling formula, i.e. the Pitman sampling formula (1.7). Therefore, as a two-parameter extension of Tavaré's model, we shown the two-parameter GEM distribution corresponds to the equilibrium state of a linear birth model constructed by Feng and Hoppe [25].

In the next section, we give some results on linear birth processes. Following that, we introduce an age-ordered infinite alleles model constructed from linear birth process with immigration. The main results are laid out in the last section.

2.3 Linear Birth Process

The linear birth process is a time-homogeneous Markov process taking values in \mathbb{N} . Suppose that Y(t) is a pure birth process with birth rate $\ell_i = \rho i + c, \rho, c > 0$. Thus, the process has the conservative Q-matrix consisting of elements $q_{i,i+1} = \rho i + c$ with other off-diagonal elements being 0 for $i = 0, 1, 2, \ldots$ It followed that the forward equation for the transition probability $P_{ij}(t)$ has the form

$$P'_{ij}(t) = \ell_{j-1} P_{i,j-1}(t) - \ell_j P_{i,j}(t), \ t \ge 0.$$
(2.2)

Since this is a pure birth process, $P_{ij}(t) = 0$ for j < i and t > 0. For j = i, we have

$$P'_{ii}(t) = -\ell_i P_{ii}(t), \ t \ge 0.$$
(2.3)

Solving the ODE (2.3) with initial condition $P_{ii}(0) = 1$, we conclude that

$$P_{ii}(t) = e^{-\ell_i t} = e^{-(\rho i + c)}, \ t \ge 0.$$

Following (2.2), we get an iteration equation:

$$P_{ij}(t) = \ell_{j-1} \int_0^t e^{-\ell_j(t-s)} P_{i,j-1}(s) \, ds \tag{2.4}$$

,

for j > i and t > 0. Particularly for j = i + 1, we have

$$P_{i,i+1}(t) = \ell_i e^{-\ell_{i+1}t} \frac{e^{t(\ell_{i+1}-\ell_i)} - 1}{\ell_{i+1} - \ell_i}$$
$$= (i + \frac{c}{\rho}) e^{-t(\rho i + c)} (1 - e^{-t\rho}).$$

By induction we have for general $j \ge i \ge 0$ and t > 0

$$P_{ij}(t) = {\binom{c}{\rho} + j - 1}{j - i} e^{-t(\rho i + c)} (1 - e^{-\rho t})^{j - i}.$$

In our case with Y(0) = 0 it follows that

$$P(Y(t) = n | Y(0) = 0) = {\binom{c}{\rho} + n - 1}{n} e^{-ct} (1 - e^{-\rho t})^n, \ n = 0, 1, \dots$$
(2.5)

In addition, the processes $e^{-\rho t}Y(t)$ are nonnegative submartingales with respect to their respective natural filtration. Thus the limit as t goes to infinity almost surely exists. Furthermore, the limit random variable has a gamma distribution. This is because of the convergence of the moment generating functions. By (2.5), letting $p = e^{-\rho t}$ we have for $\lambda < 1$

$$E[e^{\lambda pY(t)}] = \sum_{n=0}^{\infty} e^{np\lambda} {\binom{\frac{c}{\rho} + n - 1}{n}} e^{-ct} (1 - e^{-\rho t})^n$$

= $e^{-ct} (1 - e^{p\lambda} (1 - e^{-\rho t}))^{-c/\rho}$
= $(e^{\rho t} (1 - e^{p\lambda} (1 - e^{-\rho t})))^{-c/\rho} \rightarrow \left(\frac{1}{1 - \lambda}\right)^{-c/\rho}$

as $t \to \infty$. Therefore, we conclude that

$$\lim_{t \to \infty} e^{-\rho t} Y(t) = \Gamma_{c/\rho} \quad \text{almost surely,}$$

where Γ_d denotes the gamma(d,1) random variable with density $f(x) = \frac{x^{d-1}e^{-x}}{\Gamma(d)}, x > 0$. Applying similar arguments to Y(t) starting at Y(0) = 1, we have

$$P(Y(t) = n | Y(0) = 1) = {\binom{c}{\rho} + n - 1}{n - 1} e^{-(\rho + c)t} (1 - e^{-\rho t})^{n - 1}, \ n = 1, 2, \dots$$

and

$$e^{-\rho t}Y(t) \to \Gamma_{1+\frac{c}{\rho}} \quad a.s.,$$

as $t \to \infty$.

2.4 Age-ordered Infinite Alleles Model

The evolution of the total population is modeled by a linear birth process $\{N(t), t \ge 0\}$ with N(0) = 0. The immigration process I(t) is a pure birth process with I(0) = 0 and birth rate

$$\lambda_k = \lim_{h \to 0} \frac{1}{h} P(I(t+h) - I(t)) = 1 | I(t) = k),$$

given by $\lambda_k = k\alpha + \theta$, $k \ge 0$. Each family is initiated by an immigrant according to a pure birth process $\{X(t), t \ge 0\}$ starting at X(0) = 1 with birth rate $b_n = n - \alpha$. For i = 1, 2, ... let $\{X_i(t)\}$ be independent copies of X(t) and T_i be the arrival times of the *i*th immigrant. Then $X_i(t - T_i)$ would be the population size at time *t* of the *i*th family and $N(t) = \sum_{i=1}^{I(t)} X_i(t - T_i)$. The birth rate of N(t) is

$$r_n = \lim_{h \to 0} \frac{1}{h} P(N(t+h) - N(t)) = 1 | N(t) = n) = \theta + n,$$

since

$$P(N(t+h) - N(t) = 1 | N(t) = n)$$

= $E\left[(\theta + I(t)\alpha + \sum_{i=1}^{I} (t)(X_i(t-T_i) - \alpha))h + o(h)|N(t) = n\right]$
= $(\theta + n)h + o(h).$

Therefore, we have

$$e^{-t}N(t) \to \Gamma_{\theta} \quad a.s.$$
 (2.6)

and

$$e^{-t}X(t) \to \Gamma_{1-\alpha} \quad a.s.$$
 (2.7)

as $t \to \infty$.

Let the stochastic process $\{\eta(t), t \geq 0\}$ record the sizes of families in the order of their appearances in the population. Thus, define $\eta_i(t) = X_i(t - T_i) \mathbb{1}_{\{t \geq T_i\}}$. Then $\eta(t) = (\eta_1(t), \eta_2(t), \ldots), t > 0$ with $\eta(0) = (0, 0, \ldots)$. From the result of (2.7), let E_1, E_2, \ldots be independent copies of $\Gamma_{1-\alpha}$ satisfying $e^{-t}X_i(t) \to E_i$, a.s. as $t \to \infty$. Consider the *i*th oldest family size $\eta_i(t)$ at time t, then

$$e^{-t}\eta_{i}(t) = e^{-t}X_{i}(t-T_{i})1_{\{t\geq T_{i}\}}$$

= $e^{-T_{i}}e^{-(t-T_{i})}X_{i}(t-T_{i})1_{\{t\geq T_{i}\}}$
 $\rightarrow e^{-T_{i}}E_{i}, \quad a.s.$

as $t \to \infty$. We get the following general result.

Theorem 2.1. The age-ordered family size $\{\eta(t) = (\eta_1(t), \eta_2(t), \ldots)\}$ has asymptotic structure given by

$$e^{-t}(\eta_1(t), \eta_2(t), \ldots) \to (e^{-T_1}E_1, e^{-T_2}E_2, \ldots)$$
 (2.8)

almost surely as $t \to \infty$.

Proof. Fix $n \in \mathbb{N}$. We have

$$e^{-t}(\eta_1(t), \dots, \eta_n(t)) = \sum_{i=1}^{n-1} e^{-t}(X_1(t-T_1), \dots, X_i(t-T_i), 0, \dots, 0) \mathbf{1}_{\{T_i \le t < T_{i+1}\}} + e^{-t}(X_1(t-T_1), \dots, X_n(t-T_n)) \mathbf{1}_{T_n \le t}$$

 $\to (e^{-T_1}E_1, \dots, e^{-T_n}E_n), \quad a.s. \quad \text{as } t \to \infty$

by an argument analogous to the above. Intersecting the sets on which the a.s. convergence occurs for each n gives a set of probability one on which the convergence holds for each n, and the theorem is proved.

We will be interested in the asymptotic behavior of the fraction of the population that belong to the oldest, next oldest,...families in the process. We need the following result.

Theorem 2.2. Let $\sigma_i = e^{-T_i}E_i$, i = 1, 2, ... be the limiting random variables in the above theorem. Then

$$e^{-t}N(t) = \sum_{j\ge 1} e^{-t}\eta_j(t) \to \sum_{j\ge 1} \sigma_j =: \sigma, \ a.s. \ as \ t \to \infty,$$

where σ has the gamma(θ , 1) distribution.

Proof. Following the result of (2.6) we know that $e^{-t}N(t)$ converges almost surely to Y with the gamma(θ ,1) distribution. By Fatou's lemma we also have

$$e^{-t}N(t) = \sum_{i=1}^{I(t)} e^{-(t-T_i)} X_i(t-T_i) e^{-T_i}$$

$$\geq 1_{\{I(t)\geq r\}} \sum_{i=1}^r e^{-(t-T_i)} X_i(t-T_i) e^{-T_i}$$

Letting t goes to infinity on both sides we have $Y \ge \sum_{i=1}^{r} \sigma_i$ for any $r \in \mathbb{N}$. Thus, $Y \ge \sum_{i=1}^{\infty} \sigma_i$.

On the other side, we have $E[\sum_{i=1}^{\infty} \sigma_i] = \sum_{i=1}^{\infty} E[e^{-T_i}E_i] = (1-\alpha)\sum_{i=1}^{\infty} E[e^{-T_i}].$ We know that $T_i = \sum_{j=1}^{i} \tau_j$, where τ_j 's are independent random variables with the exponential distribution whose mean is $\frac{1}{\theta + (j-1)\alpha}$. Thus,

$$\begin{split} \sum_{i=1}^{\infty} E[\mathrm{e}^{-T_i}] &= \sum_{i=1}^{\infty} \prod_{j=1}^{i} E\mathrm{e}^{-\tau_j} \\ &= \sum_{i=1}^{\infty} \prod_{j=1}^{i} \frac{\theta + (j-1)\alpha}{\theta + (j-1)\alpha + 1} \\ &= \frac{\Gamma((\theta+1)/\alpha)}{\Gamma(\theta/\alpha)} \sum_{i=1}^{\infty} \prod_{j=1}^{i} \frac{\theta/\alpha + j - 1}{(\theta+1)/\alpha + j - 1} \cdot \frac{\Gamma(\theta/\alpha)}{\Gamma((\theta+1)/\alpha)} \\ &= \frac{\Gamma((\theta+1)/\alpha)}{\Gamma(\theta/\alpha)} \sum_{i=1}^{\infty} \frac{\Gamma(\theta/\alpha + i)}{\Gamma((\theta+1)/\alpha + i)} \\ &= \frac{\Gamma((\theta+1)/\alpha)}{\Gamma(\theta/\alpha)\Gamma(1/\alpha)} \sum_{i=1}^{\infty} Beta(\theta/\alpha + i, 1/\alpha) \\ &= \frac{\Gamma((\theta+1)/\alpha)}{\Gamma(\theta/\alpha)\Gamma(1/\alpha)} Beta(\theta/\alpha + 1, 1/\alpha - 1) \\ &= \frac{\theta}{1-\alpha}. \end{split}$$

Therefore, $E[\sum_{i=1}^{\infty} \sigma_i] = \theta$. Since $EY = \theta$, we have $Y = \sum_{i=1}^{\infty} \sigma_i$ a.s..

Combining Theorem 2.1 with Theorem 2.2, we immediately get the following result.

Corollary 2.3.

$$N(t)^{-1}(\eta_1(t), \eta_2(t), \ldots) \to (P_1, P_2, \ldots) \quad a.s. \ as \ t \to \infty,$$

where

$$P_i = \sigma_i / \sigma = \mathrm{e}^{-T_i} E_i / \left(\sum_{i \ge 1} \mathrm{e}^{-T_i} E_i \right), \quad i = 1, 2, \dots$$

The random variable P_i gives the asymptotic proportion of the *i*th oldest family size. The next theorem describes the structure of $\{P_i\}$.

$$P_i \stackrel{D}{=} (1 - U_1)(1 - U_2) \cdots (1 - U_{i-1})U_i, \ i \ge 2 \ and \ P_1 = U_1.$$
(2.9)

Proof. For any $r \in \mathbb{N}$, we calculate the joint distribution of $(\sigma_1, \ldots, \sigma_r, \sum_{j>r} \sigma_j)$. Since

$$(\sigma_1, \dots, \sigma_r, \sum_{j>r} \sigma_j) = (e^{-T_1} E_1, \dots, e^{-T_r} E_r, e^{-T_r} \sum_{j>r} e^{-(T_j - T_r)} E_j)$$

= $(e^{-T_1} E_1, \dots, e^{-T_r} E_r, e^{-T_r} \sigma^*),$

where σ^* is independent of $E_1, \ldots, E_r, T_1, \ldots, T_r$ with the gamma $(\theta + r\alpha, 1)$ distribution. Then the joint probability of $(e^{-T_1}E_1, \ldots, e^{-T_r}E_r, e^{-T_r}\sigma^*)$:

$$P(e^{-\tau_1}E_1 \le x_1, e^{-(\tau_1 + \tau_2)}E_2 \le x_2, \dots, e^{-(\tau_1 + \dots + \tau_r)}E_r \le x_r, e^{-(\tau_1 + \dots + \tau_r)}\sigma^* \le x_{r+1})$$

$$= \frac{\prod_{i=1}^r (\theta + (i-1)\alpha)}{\Gamma(1-\alpha)^r \Gamma(\theta + r\alpha)} \int_A \prod_{i=1}^r e^{-(\theta + (i-1)\alpha)\tau_i} \prod_{i=1}^r z_i^{-\alpha} e^{-z_i} z_{r+1}^{\theta + r\alpha - 1} e^{-z_{r+1}} d\tau dz,$$

where A denotes the set of $\{(z, \tau)\}$ satisfying the following inequalities

$$\begin{cases} e^{-\tau_1} z_1 \leq x_1 \\ e^{-(\tau_1 + \tau_2)} z_2 \leq x_2 \\ \vdots \\ e^{-(\tau_1 + \dots + \tau_r)} z_r \leq x_r \\ e^{-(\tau_1 + \dots + \tau_r)} z_{r+1} \leq x_{r+1}. \end{cases}$$

After a change of variables by letting $\nu_i = \sum_{j=1}^i \tau_j$, $y_i = e^{-\nu_i} z_i$ for $i = 1, \ldots, r+1$, we have the Jacobian matrix $|J| = e^{\sum_{i=1}^r \nu_i + \nu_r}$ and do the integration on the region of $\{0 \le \nu_1 \le \cdots \le \nu_r < \infty\}$ we get the density function for $(\sigma_1, \ldots, \sigma_r, \sum_{j>r} \sigma_j)$ is

$$f(x_1, \dots, x_{r+1}) = \frac{\prod_{i=1}^r (\theta + (i-1)\alpha) \prod_{i=1}^r x_i^{-\alpha} x_{r+1}^{\theta + r\alpha - 1} e^{-(\sum_{i=1}^{r+1} x_i)}}{\Gamma(1-\alpha)^r \Gamma(\theta + r\alpha) (x_{r+1} + x_r) \cdots (x_{r+1} + \dots + x_1)}.$$

Hence, the joint density of $(\sigma_1/\sigma, \ldots, \sigma_r/\sigma, \sigma)$ is

$$g(y_1, \dots, y_{r+1}) = \frac{\Gamma(\theta) \prod_{i=1}^r (\theta + (i-1)\alpha) \prod_{i=1}^r y_i^{-\alpha} (1 - y_1 - \dots - y_r)^{\theta + r\alpha - 1}}{\Gamma(1 - \alpha)^r \Gamma(\theta + r\alpha) (1 - y_1 - \dots - y_{r-1}) \cdots (1 - y_1)} \frac{y_{r+1}^{\theta - 1} e^{-y_{r+1}}}{\Gamma(\theta)}$$

Thus, σ is independent of $(\sigma_1/\sigma, \ldots, \sigma_r/\sigma)$. After a standard calculation it is shown that the above density is exactly the joint density of (P_1, \ldots, P_r) defined in (2.9). \Box

Chapter 3

The Sampling Formula and Laplace Transform

3.1 Abstract

A basic tool in the statistical inference of population genetics is Ewens sampling formula, which describes the probabilities associated with counts of how many different alleles are observed a given number of times in the sample. In this chapter, we show the robustness of the sampling formula by using it to derive the Laplace transform of the two-parameter Poisson-Dirichlet distribution. Then, we reverse this derivation by using the Laplace transform to get the sampling formula. Our result can be seen as an interpretation of Kingman's conclusion on the equivalence between the sampling formula and random mass partition.

Then, we proceed with two examples: the IMNA model and selection model.

In Section 3.5, we first obtain the Laplace transform of the IMNA model at a fixed time. Then, we use it to derive a Gaussian limit theorem for the homozygosity functions at each fixed time as the mutation rate goes to infinity.

In Section 3.6, the Laplace transform associated with a non-neutral infinite alleles model with a special selection is obtained from its sampling formula.

It should be noted that this chapter is originated from the accepted paper [68]. The references of the chapter are indexed to adapt to the thesis, therefore differ from those in the original paper.

3.2 Introduction

We start with the random partition structure associated with the two-parameter Poisson-Dirichlet distribution with parameter $0 \le \alpha < 1, \theta > -\alpha$, which concentrates on the infinite dimensional ordered simplex

$$\nabla_{\infty} := \left\{ \mathbf{x} = (x_1, x_2, \ldots) : x_1 \ge x_2 \ge \cdots \ge 0, \sum_{i=1}^{\infty} x_i = 1 \right\}$$

with the topology inherited from $[0.1]^{\infty}$. For $0 \leq \alpha < 1$ and $\theta > -\alpha$, let $U_k, k = 1, 2, \ldots$ be a sequence of independent random variables such that U_k has $Beta(1 - \alpha, \theta + k\alpha)$ distribution with density function given by

$$f(x) = \frac{\Gamma(\theta + 1 + (k - 1)\alpha)}{\Gamma(1 - \alpha)\Gamma(\theta + k\alpha)} x^{-\alpha} (1 - x)^{\theta + k\alpha - 1}, \ 0 < x < 1.$$

Set $(V_1, V_2, ...)$ as

$$V_1 = U_1, V_n = (1 - U_1) \cdots (1 - U_{n-1})U_n, \ n \ge 2.$$
(3.1)

Arrange $(V_1, V_2, ...)$ in descending order and denote it as $P = (P_1, P_2, ...)$. Then, the law of P is called the two-parameter Poisson-Dirichlet distribution denoted by $PD(\alpha, \theta)$. The law of $(V_1, V_2, ...)$ is called the two-parameter GEM distribution denoted by $GEM(\alpha, \theta)$. $PD(0, \theta)$ is the Poisson-Dirichlet distribution with parameter θ usually denoted as $PD(\theta)$.

The exchangeable random partition functions associated with $PD(\theta)$ and $PD(\alpha, \theta)$ are known as Ewens sampling formula and Pitman sampling formula. Consider a random sample of n individuals. The composition of the sample can be given by a partition vector $\mathbf{n} = (n_1, \ldots, n_k)$ of n, which indicates there are k different types of alleles in the sample and each family consists of n_i individuals. The partition structure of sample \mathbf{n} is often denoted by a different but equivalent vector $\mathbf{a} = (a_1, \ldots, a_n)$, where a_i is the number of families containing i individuals. Therefore $a_i \ge 0$ and $\sum_{i=1}^n ia_i = n$. Given $\mathbf{x} = (x_1, x_2, \ldots) \in \nabla_{\infty}$ where x_i denotes the frequency of the ith largest family in the population, the conditional probability of a random sample of size n with partition structure $A_n = \mathbf{a}$ is expressed as

$$P(A_n = \mathbf{a} | \mathbf{x} = (x_1, x_2, \ldots)) = \frac{n!}{\prod_{i=1}^n i!^{a_i}} \sum_{\mathbf{n}} x_1^{n_1} x_2^{n_2} \cdots ,$$

where the summation is over the set $\{\mathbf{n} = (n_1, n_2, ...) : \sharp\{i : n_i = j\} = a_j, n_i \ge 0\}$. When $\mathbf{x} \sim PD(\theta)$ the probability

$$P(A_n = \mathbf{a}) = \int_{\nabla_{\infty}} P(A_n = \mathbf{a} | \mathbf{x} = (x_1, x_2, \ldots)) PD(\theta)(d\mathbf{x})$$

is given by Ewens sampling formula (ESF)

$$P(A_n = \mathbf{a}) = \frac{n!}{\theta_{(n)}} \prod_{j=1}^n (\frac{\theta}{j})^{a_j} \frac{1}{a_j!},$$
(3.2)

where $\theta_{(n)} = \theta \cdots (\theta + n - 1)$. In the two-parameter case, i.e. $\mathbf{x} \sim PD(\alpha, \theta)$, the corresponding formula was obtained by Pitman [54] known as the Pitman sampling formula (PSF)

$$P(A_n = \mathbf{a}) = \frac{n!}{\theta_{(n)}} \prod_{\ell=0}^{k-1} (\theta + \ell\alpha) \prod_{j=1}^n \frac{(1-\alpha)_{(j-1)}^{a_j}}{(j!)^{a_j} (a_j!)}.$$
(3.3)

The sampling formula has numerous applications in population genetics. It includes the estimation of mutation rate θ and comparison between the neutral model and non-neutral one. Here, neutral means we only consider the mutation factor in the evolution. The proof of the sampling formula also leads to recursive constructions of some well-known partition structures, such as Hoppe's Urn model, Blackwell-MacQueen urn and Chinese restaurant process (see e.g. [56]). In this thesis, we give an alternative derivation of the sampling formula from the Laplace transform associated with the $PD(\alpha, \theta)$ distribution. We show that the Laplace transform can also be obtained from the sampling formula.

Pitman, Yor [57] and Handa [38] characterized the two-parameter Poisson-Dirichlet distribution by the Laplace transform of a probability generating functional of the random vector \mathbf{x} with the $PD(\alpha, \theta)$ distribution. The special form of the Laplace transform can be seen as a result of Campbell's theorem from the Poisson point process theory (see [46]). It is also based on the connection of $PD(\theta)$ and $PD(\alpha, \theta)$ with some subordinators in [57]. It can be seen that in [38] the key to derive the Laplace transform is the correlation measure of the point process, which is obtained based on the size-biased permutation of $PD(\alpha, \theta)$ given by the GEM(α, θ) distribution. In this thesis, we derive the correlation measure directly from the sampling formula (3.3), which can be regarded as an interpretation of Kingman's conclusion that the partition function and random mass partition are unique to each other. We gather the above result and the proof in the fourth section followed by introducing several necessary definitions and notations in the third section.

In Section 3.5, we apply the Laplace transform to the IMNA model with mutation rate θ at a fixed time. It leads to the establishment of a central limit theorem associated with the homozygosity functions at time t as θ goes to infinity. Comparing with the limiting behavior corresponding to the equilibrium state studied by Joyce et al, we conclude the similarity of the population at any finite time and the equilibrium state when the mutation is quite strong. If we look at the transition function of the IMNA model obtained in [13], it can be further inferred that the descendants of the ancestors disappear quickly and the population is dominated by the mutants as the mutation rate goes to infinity.

In the end, we consider a non-neutral model involving a special selection. In the finite alleles diffusion model with selection, the stationary distribution is shown to be absolutely continuous with respect to the neutral one with the density $e^{sH_2(\mathbf{X})}/C$, where $H_2 = \sum X_i^2$ is the homozygosity function, $s \in R$ denotes the selection intensity and C is the normalization constant. In an infinite alleles model with selection [17], the stationary distribution $PD^{s,q}(\theta)$ is also absolutely continuous with respect to the Poisson-Dirichlet distribution $PD^{(\theta)}(\theta)$ with the Radon-Nikodym derivative

$$e^{sH_q(\mathbf{X})}/C, \qquad \mathbf{X} = (X_1, X_2, \ldots) \in \nabla_{\infty}$$

where $H_q(\mathbf{X}) = \sum X_i^q$, q = 2, 3, ... denotes the *q*th order homozygosity. Handa [37] considers a more general partition structure under this distribution

$$P_{\theta,s,q}(\mathbf{A}_n = \mathbf{a}) = \mathbb{E}[e^{sH_q(\mathbf{X})}\psi_{\mathbf{n}}(\mathbf{X})]/\mathbb{E}[e^{sH_q(\mathbf{X})}], \qquad (3.4)$$

where

$$\psi_{\mathbf{n}}(\mathbf{x}) = \frac{n!}{n_1! \cdots n_k! a_1! \cdots a_n!} \sum_{i_1, \dots, i_k \neq} x_{i_1}^{n_1} \cdots x_{i_k}^{n_k}$$
(3.5)

and the sum is taken over *n*-tuple of distinct indices. The expectation is taken under the $PD(\theta)$ distribution. In the last section, we derive the correlation function and Laplace transform for this selection model.

3.3 Preliminaries and Notations

First, we briefly introduce some definitions in the theory of point process. Readers can refer to [5] for details. The random element $(X_i)_{i=1}^{\infty}$ governed by the $PD(\alpha, \theta)$ distribution can be viewed as a random point process $\xi := \sum_{i=1}^{\infty} \delta_{X_i}$, which is called the (two-parameter) Poisson-Dirichlet point process with parameters (α, θ) or simply the $PD(\alpha, \theta)$ process discussed in [38]. For any positive integer n, the nth correlation measure of ξ , if it exists, is defined to be a σ -finite Borel measure μ_n such that for any nonnegative measurable function f on \mathbf{R}^n

$$\operatorname{E}\left[\sum_{i_1,\ldots,i_n\neq}f(X_{i_1},\cdots,X_{i_n})\right] = \int_{\mathbf{R}^n}f(x_1,\ldots,x_n)\,\mu_n(dx_1\cdots dx_n),$$

where the subscript on the left side denotes that the sum is taken over n-tuple of distinct indices. If μ_n has a density with respect to the *n*-dimensional Lebesgue measure, the density is called the *n*th correlation function of ξ .

Another thing we need to point out here is that the sampling formula (3.3) is equivalent to the following statement: given an arbitrary partition (n_1, \ldots, n_k) of n, we have k-1

$$E_{\alpha,\theta}\left[\sum_{i_1,\dots,i_k\neq} X_{i_1}^{n_1}\cdots X_{i_k}^{n_k}\right] = \frac{\prod_{l=0}^{n-1}(\theta+l\alpha)}{\theta_{(n)}}\prod_{i=1}^k (1-\alpha)_{(n_i-1)},$$
(3.6)

where $E_{\alpha,\theta}$ means the expectation is taken w.r.t. the $PD(\alpha, \theta)$ distribution. This is the key to show our first result about the correlation function of the $PD(\alpha, \theta)$ process ξ in the following part.

3.4 The Sampling Formula and Laplace Transform

Theorem 3.1. Suppose the sampling formula (3.3), which is equivalent to (3.6), holds for a random element $(X_i)_{i=1}^{\infty} \in \nabla_{\infty}$. Then the kth correlation function of the point process $\xi = \sum_{i=1}^{\infty} \delta_{X_i}$ for each k = 1, 2, ... is as follows:

$$q_k(x_1,\ldots,x_k) = c_{k,\alpha,\theta} \prod_{i=1}^k x_i^{-(\alpha+1)} (1 - \sum_{j=1}^k x_j)^{\theta+\alpha k-1} \mathbf{1}_{\Delta_k}(x_1,\ldots,x_k),$$
(3.7)

where

$$\Delta_k = \left\{ (x_1, \dots, x_k) : x_1 \ge 0, \dots, x_k \ge 0, \sum_{i=1}^k x_i \le 1 \right\}$$

and

$$c_{k,\alpha,\theta} = \prod_{i=1}^{k} \frac{\Gamma(\theta + 1 + (i-1)\alpha)}{\Gamma(1-\alpha)\Gamma(\theta + i\alpha)}.$$

Proof. In fact, we only need to show that for any nonnegative measurable function $f(x_1, \ldots, x_k)$ on Δ_k , the identity

$$\operatorname{E}\left[\sum_{i_1,\dots,i_k\neq} f(X_{i_1},\cdots,X_{i_k})\right] = \int_{\Delta_k} f(x_1,\dots,x_k)q_k(x_1,\dots,x_k)\,dx_1\cdots dx_k \qquad (3.8)$$

holds. Set $f(x_1, \ldots, x_k) = x_1^{n_1} \cdots x_k^{n_k}$. Then, the right side of the above (3.8) equals to

$$c_{k,\alpha,\theta} \int_{\Delta_k} \prod_{i=1}^k x_i^{n_i - \alpha - 1} (1 - \sum_{j=1}^k x_j)^{\theta + \alpha k - 1} dx_1 \cdots dx_k$$

=
$$\prod_{i=1}^k \frac{\Gamma(\theta + 1 + (i - 1)\alpha)}{\Gamma(1 - \alpha)\Gamma(\theta + i\alpha)} \frac{\prod_{i=1}^k \Gamma(n_i - \alpha)\Gamma(\theta + k\alpha)}{\Gamma(\theta + n)}$$

=
$$\frac{\prod_{i=1}^k (1 - \alpha)_{(n_i - 1)} \prod_{i=0}^{k - 1} (\theta + i\alpha)}{\theta_{(n)}},$$

which indicates that (3.8) holds for all the polynomial functions on Δ_k . Let \mathscr{L} be the set of all the nonnegative functions satisfying (3.8). It is easy to check that \mathscr{L} is a λ – system, which means \mathscr{L} satisfies the following three conditions:

- 1. $1 \in \mathscr{L};$
- 2. \mathscr{L} is closed under finite linear combinations;
- 3. If $f_n \in \mathscr{L}$, $f_n \uparrow f$, then $f \in \mathscr{L}$.

If we can show that \mathscr{L} contains all the nonnegative bounded continuous functions on Δ_k , then \mathscr{L} contains all nonnegative measurable functions on Δ_k by the monotone theorem.

Since Δ_k is compact, by the Stone-Weierstrass theorem any continuous function $f \in C(\Delta_k)$ can be approximated by a sequence of polynomial functions uniformly. Thus, for any nonnegative $f \in C(\Delta_k)$, there exists a sequence of polynomials $\{B_m\}$ such that $\sup_{\Delta_k} |B_m(x) - f(x)| \to 0$, as $m \to \infty$. Without loss of generality, we can assume $\{B_m\}$ are also nonnegative. We consider the following two cases.

Case 1: $\int_{\Delta_k} f(x_1, \ldots, x_k) q_k(x_1, \ldots, x_k) dx_1 \cdots dx_k < \infty$. Denote the interior of Δ_k by Δ_k° and set $\mathbb{Q}^k \cap \Delta_k^{\circ} = \{q_1, q_2, \ldots\}$, since it is a countable set. For any $\varepsilon > 0$, and $q_1 = (q_{11}, \ldots, q_{1k})$, since $f(q_1) < f(q_1) + \varepsilon \prod_{j=1}^k q_{1j}$ and $B_m(q_1) \to f(q_1)$, we can find a subsequence of B_m named B_{1m} such that $B_{1m}(q_1) \leq f(q_1) + \varepsilon \prod_{j=1}^k q_{1j}$. For the same reason we can get another subsequence of B_{1m} named B_{2m} for q_2 such that $B_{2m}(q_2) \leq f(q_2) + \varepsilon \prod_{j=1}^k q_{2j}$. Repeating this process for each q_i and using the diagonal method, we can get a sequence which we still call B_m such that $B_m(q_i) \leq f(q_i) + \varepsilon \prod_{j=1}^k q_{ij}$, for any $i = 1, 2, \ldots$ By the continuity of B_m , we have $B_m(x) \leq f(x) + \varepsilon \prod_{j=1}^k x_j, \forall x = (x_1, \ldots, x_k) \in \Delta_k$, since $\int_{\Delta_k^{\circ}} (f(x) + \varepsilon \prod_{j=1}^k x_j) q_k(x_1, \ldots, x_k) dx_1 \cdots dx_k < \infty$. By the dominated theorem, we get

$$\int_{\Delta_k^\circ} f(x)q_k(x) \, dx = \lim_{m \to \infty} \int_{\Delta_k^\circ} B_m(x)q_k(x) \, dx$$

On the other hand, we have $\int_{\Delta_k} f(x)q_k(x) dx = \int_{\Delta_k^\circ} f(x)q_k(x) dx$ and $\int_{\Delta_k} B_m(x)q_k(x) dx = \int_{\Delta_k^\circ} B_m(x)q_k(x) dx$, since the boundary of Δ_k has zero Lebesgue measure. Therefore,

$$\lim_{m \to \infty} \mathbb{E} \left[\sum_{i_1, \dots, i_k \neq} B_m(X_{i_1} \cdots X_{i_k}) \right] = \lim_{m \to \infty} \int_{\Delta_k} B_m(x) q_k(x) \, dx$$
$$= \lim_{m \to \infty} \int_{\Delta_k^\circ} f(x) q_k(x) \, dx$$
$$= \int_{\Delta_k^\circ} f(x) q_k(x) \, dx$$
$$= \int_{\Delta_k} f(x) q_k(x) \, dx.$$

Following Fatou's lemma, we have

$$\mathbb{E}\left[\sum_{i_1,\dots,i_k\neq} f(X_{i_1},\cdots,X_{i_k})\right] \le \lim_{m\to\infty} \mathbb{E}\left[\sum_{i_1,\dots,i_k\neq} B_m(X_{i_1}\cdots X_{i_k})\right].$$
 (3.9)

Since $B_m(X_{i_1}\cdots X_{i_k}) \leq f(X_{i_1},\cdots,X_{i_k}) + \varepsilon \prod_{j=1}^k X_{i_j}, a.s.$ Thus

$$\lim_{m \to \infty} \mathbb{E}\left[\sum_{i_1, \dots, i_k \neq} B_m(X_{i_1} \cdots X_{i_k})\right] \le \mathbb{E}\left[\sum_{i_1, \dots, i_k \neq} f(X_{i_1}, \cdots, X_{i_k})\right] + \varepsilon.$$

Letting $\varepsilon \downarrow 0$ and combining with (3.9), we have

$$\mathbf{E}\left[\sum_{i_1,\dots,i_k\neq}f(X_{i_1},\cdots,X_{i_k})\right] = \lim_{m\to\infty}\mathbf{E}\left[\sum_{i_1,\dots,i_k\neq}B_m(X_{i_1}\cdots X_{i_k})\right],$$

which gives the identity $\int_{\Delta_k} f(x)q_k(x) dx = \mathbb{E}\left[\sum_{i_1,\dots,i_k \neq j} f(X_{i_1},\cdots,X_{i_k})\right].$

Case 2: $\int_{\Delta_n} f(x)q_n(x) dx = \infty$. By Fatou's lemma, we have

$$\lim_{m \to \infty} \int_{\Delta_n} B_m(x) q_n(x) \, dx \ge \int_{\Delta_n} f(x) q_n(x) \, dx$$

and

$$\lim_{m \to \infty} \int_{\Delta_n} B_m(x) q_n(x) \, dx = \lim_{m \to \infty} \mathbb{E}\left[\sum_{i_1, \dots, i_k \neq i_1, \dots, i_k \neq i_k} B_m(X_{i_1} \cdots X_{i_k})\right] = \infty.$$

By the same method as above, we can find a subsequence of B_m which we still denote as B_m such that $B_m(x) \leq f(x) + \prod_{j=1}^n x_j, \forall x = (x_1, \dots, x_n) \in \Delta_n^{\circ}$. So

$$B_m(X_{i_1}\cdots X_{i_k}) \le f(X_{i_1},\cdots,X_{i_n}) + 1 \ a.s.,$$

and

$$\lim_{m \to \infty} \mathbb{E}\left[\sum_{i_1, \dots, i_k \neq} B_m(X_{i_1} \cdots X_{i_n})\right] \le \mathbb{E}\left[\sum_{i_1, \dots, i_n \neq} f(X_{i_1}, \cdots, X_{i_n})\right] a + 1.$$

Thus, $\mathbb{E}\left[\sum_{i_1, \dots, i_n \neq} f(X_{i_1}, \cdots, X_{i_n})\right] = \infty.$

With the correlation function of the point process ξ derived, we can obtain the Laplace transform of the probability generating function which is proved in Theorem 3.1 of [38].

Theorem 3.2 ([38]). Suppose that $g : (0, \infty) \to C$ is a measurable function such that

$$\lambda_{\alpha}(g) := \inf \left\{ \lambda > 0 : \int_0^\infty \frac{\mathrm{e}^{-\lambda z}}{z^{\alpha+1}} |g(z)| \, dz < \infty \right\} < \infty.$$

Put

$$\lambda_{\alpha}^{*}(g) = \inf \left\{ \lambda > \lambda_{\alpha}(g) : \frac{C_{\alpha}}{\lambda^{\alpha}} \int_{0}^{\infty} \frac{\mathrm{e}^{-\lambda z}}{z^{\alpha+1}} |g(z)| \, dz \in \mathrm{C} \setminus [1,\infty) \right\},\,$$

so that in particular $\lambda_0^*(g) = \lambda_0(g)$. The correlation measure of $\xi = \sum_{i=1}^{\infty} \delta_{X_i}$ is given by (3.7). Then

$$\mathbb{E}[\prod_{i=1}^{\infty} (1 + |g(sX_i)|)] < \infty$$

a.e. s > 0, and for $\lambda > \lambda_{\alpha}^{*}(g)$

$$\frac{\lambda^{\theta}}{\Gamma(\theta+1)} \int_0^{\infty} s^{\theta-1} e^{-\lambda s} \left(\mathbf{E} \left[\prod_{i=1}^{\infty} (1+g(sX_i)) - 1 \right] \right) ds \\ = \begin{cases} \frac{1}{\theta} \exp(\theta \int_0^{\infty} \frac{e^{-\lambda z}}{z} g(z) dz) - \frac{1}{\theta}, & \alpha = 0, \theta > 0, \\ \frac{1}{\theta} (1 - \frac{C_{\alpha}}{\lambda^{\alpha}} \int_0^{\infty} \frac{e^{-\lambda z}}{z^{\alpha+1}} g(z) dz)^{-\frac{\theta}{\alpha}} - \frac{1}{\theta}, & 0 < \alpha < 1, \theta \neq 0 \\ -\frac{1}{\alpha} \log(1 - \frac{C_{\alpha}}{\lambda^{\alpha}} \int_0^{\infty} \frac{e^{-\lambda z}}{z^{\alpha+1}} g(z)) dz), & 0 < \alpha < 1, \theta = 0 \end{cases}$$

As a corollary of the above Laplace transform, we have the following result which gives the sampling formula.

Corollary 3.3. Suppose the above Laplace transform exists for sufficiently large $\lambda > 0$, then the sampling formula corresponding to $\mathbf{X} = (X_1, X_2, \ldots) \in \nabla_{\infty}$ coincides with the Pitman sampling formula (3.3) which equals to (3.6).

Proof. Following the uniqueness of Laplace transformation or Lemma 3.1 in [38] and taking the inverse of it we have

$$F(s) = \operatorname{E}\left[\prod_{i=1}^{\infty} (g(sX_i) + 1)\right] - 1$$

= $\sum_{n=1}^{\infty} \frac{c_{n,\alpha,\theta}}{n!} \int_{\Delta_n} \prod_{i=1}^n \frac{g(sx_i)}{x_i^{\alpha+1}} (1 - \sum_{j=1}^n x_j)^{\theta + \alpha n - 1} dx_1 \cdots dx_n,$

such that g(z) satisfies the condition that there exists some $\lambda > 0$, such that $\int_0^\infty \frac{e^{-\lambda z}}{z^{\alpha+1}} |g(z)| dz < \infty$. Set s = 1 and $g(z) = t_1 z^{n_1} + \dots + t_k z^{n_k}$, where $n_1, \dots, n_k \ge 1$. It is obvious that $\int_0^\infty \frac{e^{-\lambda z}}{z^{\alpha+1}} |g(z)| dz < \infty$, $\forall \lambda > 0$. Then, we have

$$\frac{\partial^k}{\partial t_1 \cdots \partial t_k} \Big|_{t_1 = \cdots = t_k = 0} \prod_{i=1}^{\infty} (1 + g(X_i)) = \sum_{i_1, \cdots, i_k \neq} X_{i_1}^{n_1} \cdots X_{i_k}^{n_k}.$$

Therefore,

Furthermore, by interchanging the differentiation and integration processes, we have

$$\sum_{n=1}^{\infty} \frac{c_{n,\alpha,\theta}}{n!} \int_{\Delta_n} \frac{\partial^k}{\partial t_1 \cdots \partial t_k} |_{t_1 = \cdots = t_k = 0} \prod_{i=1}^n \frac{g(sx_i)}{x_i^{\alpha+1}} (1 - \sum_{j=1}^n x_j)^{\theta + \alpha n - 1} dx_1 \cdots dx_n$$
$$= \frac{c_{k,\alpha,\theta}}{k!} \int_{\Delta_k} k! \prod_{i=1}^n x_i^{n_i - \alpha - 1} (1 - \sum_{j=1}^n x_j)^{\theta + \alpha n - 1} dx_1 \cdots dx_n$$
$$= c_{k,\alpha,\theta} \frac{\Gamma(n_1 - \alpha) \cdots \Gamma(n_k - \alpha) \Gamma(\theta + k\alpha)}{\Gamma(\theta + n)}$$
$$= \frac{\prod_{i=1}^k (1 - \alpha)_{(n_i - 1)}}{\theta_{(n)}} \prod_{l=0}^{k-1} (\theta + l\alpha),$$

which implies (3.6).

3.5 Non-equilibrium of the IMNA model

3.5.1 The Laplace Transform of the IMNA Model

As mentioned in the Introduction, the Poisson-Dirichlet distribution $PD(\theta)$ corresponds to the equilibrium of the IMNA process X(t) with the following generator

$$L = \frac{1}{2} \sum_{i,j=1}^{\infty} x_i (\delta_{ij} - x_j) \frac{\partial^2}{\partial x_i \partial x_j} - \frac{\theta}{2} \sum_{i=1}^{\infty} x_i \frac{\partial}{\partial x_i},$$

whose domain is $\mathcal{D}(L) = span\{1, \varphi_2, \varphi_3, \ldots\} \subset C(\nabla_{\infty})$, and $\varphi_n(\mathbf{x}) = \sum_{i=1}^{\infty} x_i^n$ is defined on ∇_{∞} and extends continuously to its closure $\overline{\nabla}_{\infty}$, the closure of ∇_{∞} in $[0, 1]^{\infty}$. In [13] it is shown that the transition probability $P(t, \mathbf{x}, A)$ of X(t) starting from $x = (x_1, x_2, \ldots) \in \nabla_{\infty}$ is absolutely continuous with respect to the stationary distribution $PD(\theta)$. Suppose that p_n is defined as

$$p_n(x,y) = \sum_{\mathbf{n}:|\mathbf{n}|=n} \frac{\psi_{\mathbf{n}}(x)\psi_{\mathbf{n}}(y)}{\mathrm{E}[\psi_{\mathbf{n}}(z)]},$$

where $\psi_{\mathbf{n}}$ is defined as (3.5). The expectation $\mathbb{E}[\psi_{\mathbf{n}}(z)]$ is taken w.r.t. the $PD(\theta)$ distribution, which thus is given by the Ewens sampling formula (3.2). The transition

density function q(t, x, y) is given as

$$q(t, x, y) = 1 + \sum_{m=2}^{\infty} \frac{2m - 1 + \theta}{m!} e^{-\lambda_m t}$$

$$\cdot \sum_{n=0}^{m} (-1)^{m-n} {m \choose n} (n+\theta)_{(m-1)} p_n(x, y), \quad x, y \in \nabla_{\infty},$$
(3.10)

where $\lambda_m = \frac{m(m-1+\theta)}{2}$, $m = 2, 3, \ldots$ The type of expansion (3.10) was first derived by Griffiths [32]. A very similar transition function was also obtained by Ethier and Griffiths [14] for the labeled version of IMNA model, which is a measure-valued Fleming-Viot process with parent-independent mutation. Here, we consider the Laplace transform of $X(t) = (X_1(t), X_2(t), \ldots)$ with initial value x.

Theorem 3.4. Suppose that g(z) is a measurable function such that

$$\lambda(g) := \inf\{\lambda > 0 : \int_0^\infty \frac{e^{-\lambda z}}{z} |g(z)| \, dz < \infty\} < \infty.$$

Then for fixed time t, the Laplace transform of the IMNA process X(t) starting from $x \in \nabla_{\infty}$ is given by

$$\begin{split} &\int_{0}^{\infty} \mathrm{e}^{-\lambda\tau} \tau^{\theta-1} \mathrm{E}_{x} [\prod_{i=1}^{\infty} (1+g(\tau X_{i}(t)))] \, d\tau \\ &= \frac{\Gamma(\theta)}{\lambda^{\theta}} \exp\{\theta \int_{0}^{\infty} \frac{e^{-\lambda z}}{z} g(z) \, dz\} \\ &+ \sum_{m=2}^{\infty} \frac{2m-1+\theta}{m!} \mathrm{e}^{-\lambda_{m} t} \sum_{n=0}^{m} (-1)^{m-n} \binom{m}{n} (n+\theta)_{(m-1)} \sum_{\substack{\mathbf{n}=(n_{1},\ldots,n_{k})\\ \sum_{i=1}^{k} n_{i}=n}} \psi_{\mathbf{n}}(\boldsymbol{x}) \\ &\quad [\frac{\Gamma(\theta)}{\lambda^{\theta}} + ESF(\mathbf{n})^{-1} \sum_{\ell=1}^{\infty} \frac{1}{\ell!} \sum_{r=0}^{\ell \wedge k} \sum_{\substack{B \subset \{1,\ldots,k\}\\ |B|=r}} \binom{\ell}{r} \theta^{k+\ell-r} H_{k,\ell,B}], \end{split}$$

where

$$H_{k,\ell,B} = \int_{(0,\infty)^{k+\ell-r}} y_1^{n_1} \cdots y_k^{n_k} (\prod_{j \in B} g(y_j)) g(y_{k+1}) \cdots g(y_{k+\ell-r}) (\prod_{i=1}^{k+\ell-r} y_i^{-1} e^{-\lambda y_i}) dy_1 \cdots dy_{k+\ell-r} \int_0^\infty e^{-\lambda \tau} (\tau + \sum_{j=1}^{k+\ell-r} y_j)^{-n} \tau^{\theta-1} d\tau,$$

and $ESF(\mathbf{n})$ is given by Ewens sampling formula as

$$ESF(\mathbf{n}) = \frac{\theta^k}{\theta_{(n)}} (n_1 - 1)! \cdots (n_k - 1)!.$$

Proof. For an arbitrary fixed time t > 0,

$$\begin{split} \mathbf{E}_{x} \left[\prod_{i=1}^{\infty} (1 + g(\tau X_{i}(t))) \right] &= \mathbf{E} \left[\prod_{i=1}^{\infty} (1 + g(\tau Y_{i}))q(t, x, Y) \right] \\ &= \mathbf{E} \left[\prod_{i=1}^{\infty} (1 + g(\tau Y_{i})) \right] \\ &+ \sum_{m=2}^{\infty} \frac{2m - 1 + \theta}{m!} e^{-\lambda_{m}t} \sum_{n=0}^{m} (-1)^{m-n} \binom{m}{n} (n + \theta)_{(m-1)} \\ &\qquad \mathbf{E} \left[\prod_{i=1}^{\infty} (1 + g(\tau Y_{i}))p_{n}(x, Y) \right], \end{split}$$

where $Y = (Y_1, Y_2, ...)$ has the Poisson-Dirichlet distribution $PD(\theta)$. Since

$$\prod_{i=1}^{\infty} (1 + g(\tau Y_i)) p_n(x, Y) = \sum_{\substack{\mathbf{n} = (n_1, \dots, n_k) \\ \sum_{i=1}^k n_i = n}} \prod_{i=1}^{\infty} (1 + g(\tau Y_i)) \phi_{\mathbf{n}}(Y) \psi_{\mathbf{n}}(\mathbf{x}) ESF(\mathbf{n})^{-1},$$

where $\phi_{\mathbf{n}}(Y) = \sum_{i_1,\dots,i_k \neq} Y_{i_1}^{n_1} \cdots Y_{i_k}^{n_k}$. Hence, we have

$$\begin{split} &\prod_{i=1}^{\infty} (1+g(\tau Y_i))\phi_{\mathbf{n}}(Y) \\ &= \prod_{i=1}^{\infty} (1+g(\tau Y_i)) \sum_{i_1,\dots,i_k \neq} Y_{i_1}^{n_1} \cdots Y_{i_k}^{n_k} \\ &= \sum_{i_1,\dots,i_k \neq} Y_{i_1}^{n_1} \cdots Y_{i_k}^{n_k} + \sum_{\ell=1}^{\infty} \frac{1}{\ell!} \sum_{\substack{j_1,\dots,j_\ell \neq \\ i_1,\dots,i_k \neq}} g(\tau Y_{j_1}) \cdots g(\tau Y_{j_\ell}) Y_{i_1}^{n_1} \cdots Y_{i_k}^{n_k} \\ &= \sum_{\ell=0}^{\infty} \frac{1}{\ell!} \sum_{r=0}^{\ell \wedge k} \sum_{\substack{B \subset \{1,\dots,k\} \\ |B|=r}} \frac{\ell!}{(\ell-r)!} \sum_{\substack{i_1,\dots,i_{k+\ell-r} \neq}} Y_{i_1}^{n_1} \cdots Y_{i_k}^{n_k} (\prod_{j \in B} g(\tau Y_{i_j})) \\ &g(\tau Y_{i_{k+1}}) \cdots g(\tau Y_{i_{k+\ell-r}}) \end{split}$$

Therefore,

$$E[\prod_{i=1}^{\infty} (1 + g(\tau Y_i))\phi_{\mathbf{n}}(Y)] = ESF(\mathbf{n}) + \sum_{\ell=1}^{\infty} \frac{1}{\ell!} \sum_{r=0}^{\ell \wedge k} \sum_{\substack{B \subset \{1, \dots, k\} \\ |B| = r}} \frac{\ell!}{(\ell - r)!} \int_{\Delta_{k+\ell-r}} y_1^{n_1} \cdots y_k^{n_k} (\prod_{j \in B} g(\tau y_j))g(\tau y_{k+1}) \cdots g(\tau y_{k+\ell-r}) \mu_{k+\ell-r} (dy_1 \cdots dy_{k+\ell-r}).$$

By substituting the correlation measure, the above equals

$$ESF(\mathbf{n}) + \sum_{\ell=1}^{\infty} \frac{1}{\ell!} \sum_{r=0}^{\ell \wedge k} \sum_{\substack{B \subset \{1, \dots, k\} \\ |B| = r}} \frac{\ell!}{(\ell - r)!} \theta^{k+\ell-r} \int_{\Delta_{k+\ell-r}} y_1^{n_1} \cdots y_k^{n_k} (\prod_{j \in B} g(\tau y_j)) g(\tau y_{k+1}) \cdots g(\tau y_{k+\ell-r}) (1 - \sum_{i=1}^{k+\ell-r} y_i)^{\theta-1} \frac{dy_1 \cdots dy_{k+\ell-r}}{y_1 \cdots y_{k+\ell-r}}.$$

Define

$$H_{k,\ell,B} := \int_0^\infty e^{-\lambda \tau} \tau^{\theta-1} d\tau \int_{\Delta_{k+\ell-r}} y_1^{n_1} \cdots y_k^{n_k} \\ (\prod_{j \in B} g(\tau y_j)) g(\tau y_{k+1}) \cdots g(\tau y_{k+\ell-r}) (1 - \sum_{i=1}^{k+\ell-r} y_i)^{\theta-1} \frac{dy_1 \cdots dy_{k+\ell-r}}{y_1 \cdots y_{k+\ell-r}}.$$

Using Fubini's theorem, we have

$$|H_{k,\ell,B}| \le \Gamma(\theta) \left(\int_0^\infty \frac{\mathrm{e}^{-\lambda z} |g(z)|}{z} \, dz\right)^{\ell}.$$

Consequently, the series $\sum_{\ell=1}^{\infty} \frac{1}{\ell!} \sum_{r=0}^{\ell \wedge k} \sum_{\substack{B \subset \{1,\ldots,k\} \\ |B|=r}} \frac{\ell!}{(\ell-r)!} \theta^{k+\ell-r} H_{k,\ell,B}$ is absolutely convergent, with upper bound given by $\Gamma(\theta) e^{\theta M} \theta^k (1+M)^k$, where

$$M = \int_0^\infty e^{-\lambda x} |g(x)| x^{-1} \, dx < \infty.$$

In addition, we have obtained previously

$$\int_0^\infty e^{-\lambda\tau} \tau^{\theta-1} \mathbb{E}\left[\prod_{i=1}^\infty (1+g(\tau Y_i))\right] d\tau = \frac{\Gamma(\theta)}{\lambda^\theta} \exp\{\theta \int_0^\infty \frac{e^{-\lambda z}}{z} g(z) \, dz\}.$$

Therefore,

$$\begin{split} \int_{0}^{\infty} e^{-\lambda\tau} \tau^{\theta-1} E_{x} [\prod_{i=1}^{\infty} (1+g(\tau X_{i}(t)))] d\tau \\ &= \int_{0}^{\infty} e^{-\lambda\tau} \tau^{\theta-1} E[\prod_{i=1}^{\infty} (1+g(\tau Y_{i}))] d\tau + \sum_{m=2}^{\infty} \frac{2m-1+\theta}{m!} e^{-\lambda_{m}t} \sum_{n=0}^{m} (-1)^{m-n} \\ & \left(\binom{m}{n} (n+\theta)_{(m-1)} \int_{0}^{\infty} e^{-\lambda\tau} \tau^{\theta-1} E[\prod_{i=1}^{\infty} (1+g(\tau Y_{i}))p_{n}(x,Y)] d\tau \\ &= \frac{\Gamma(\theta)}{\lambda^{\theta}} \exp\{\theta \int_{0}^{\infty} \frac{e^{-\lambda z}}{z} g(z) dz\} + \sum_{m=2}^{\infty} \frac{2m-1+\theta}{m!} e^{-\lambda_{m}t} \sum_{n=0}^{m} (-1)^{m-n} \\ & \left(\binom{m}{n} (n+\theta)_{(m-1)} \sum_{\substack{\mathbf{n}=(n_{1},\dots,n_{k})\\ \sum_{i=1}^{k} n_{i}=n}} \psi_{\mathbf{n}}(x) [\frac{\Gamma(\theta)}{\lambda^{\theta}} + ESF(\mathbf{n})^{-1} \right] \\ & \sum_{\ell=1}^{\infty} \frac{1}{\ell!} \sum_{r=0}^{\ell \wedge k} \sum_{\substack{B \subset \{1,\dots,k\}\\ |B|=r}} \frac{\ell!}{(\ell-r)!} \theta^{k+\ell-r} H_{k,\ell,B}], \end{split}$$

where

$$H_{k,\ell,B} = \int_{(0,\infty)^{k+\ell-r}} y_1^{n_1} \cdots y_k^{n_k} (\prod_{j \in B} g(y_j)) g(y_{k+1}) \cdots g(y_{k+\ell-r}) \frac{dy_1 \cdots dy_{k+\ell-r}}{y_1 \cdots y_{k+\ell-r}}$$
$$\int_{\sum_{j=1}^{k+\ell-r}}^{\infty} e^{-\lambda \tau} \tau^{-n} (\tau - \sum_{j=1}^{k+\ell-r} y_j)^{\theta-1} d\tau$$
$$= \int_{(0,\infty)^{k+\ell-r}} y_1^{n_1} \cdots y_k^{n_k} (\prod_{j \in B} g(y_j)) g(y_{k+1}) \cdots g(y_{k+\ell-r})$$
$$\prod_{i=1}^{k+\ell-r} (y_i^{-1} e^{-\lambda y_i}) dy_1 \cdots dy_{k+\ell-r} \int_0^{\infty} e^{-\lambda \tau} (\tau + \sum_{j=1}^{k+\ell-r} y_j)^{-n} \tau^{\theta-1} d\tau.$$

The upper bound of $\sum_{\ell=1}^{\infty} \frac{1}{\ell!} \sum_{r=0}^{\ell \wedge k} \sum_{\substack{B \subset \{1,\dots,k\} \ |B|=r}} \frac{\ell!}{(\ell-r)!} \theta^{k+\ell-r} H_{k,\ell,B}$ guarantees the absolute convergence of the above series.

3.5.2 CLT for the IMNA Model

Furthermore, the above Laplace transform enable us to derive the central limit theorem associated with the distribution of the IMNA process X(t) with starting point $x = (x_1, x_2, \ldots) \in \nabla_{\infty}$.

Theorem 3.5. For any time t > 0, set $W_p(t) = \sqrt{\theta}(\frac{\theta^{p-1}}{(p-1)!}H_p(t) - 1)$, where $H_p(t) = \sum_{i=1}^{\infty} X_i(t)^p$ denotes the pth order homozygosity at time t for $p = 2, 3, \ldots$. Then as $\theta \to \infty$, $W_p(t)$ converges in law to a normal random variable with mean θ and variance $\sigma_p^2 = \Gamma(2p)/\Gamma(p)^2 - p^2$, which is strictly positive.

Proof. To prove the central limit theorem for $W_p(t)$, it suffies to show that the characteristic function of $W_p(t)$ which is $\psi_p(x) = \mathbb{E}[\exp(ixW_p(t))]$ converges to $\exp(-\frac{\sigma_p^2}{2}x^2)$ as $\theta \to \infty$.

Recall from Theorem 3.1, for any fixed time t > 0 we have

$$\frac{1}{\Gamma(\theta)} \int_{0}^{\infty} e^{-\tau} \tau^{\theta-1} \mathbb{E}_{x} [\prod_{i=1}^{\infty} (1+g(\tau X_{i}(t)))] d\tau$$

$$= \exp\{\theta \int_{0}^{\infty} \frac{e^{-z}}{z} g(z) dz\}$$

$$+ \sum_{m=2}^{\infty} \frac{2m-1+\theta}{m!} e^{-\lambda_{m}t} \sum_{n=0}^{m} (-1)^{m-n} {m \choose n} (n+\theta)_{(m-1)} \sum_{\substack{\mathbf{n}=(n_{1},\dots,n_{k})\\ \sum_{i=1}^{k} n_{i}=n}}^{\mathbf{n}=(n_{1},\dots,n_{k})} \psi_{\mathbf{n}}(\mathbf{x})$$

$$[1+\frac{1}{\Gamma(\theta)} ESF(\mathbf{n})^{-1} \sum_{\ell=1}^{\infty} \frac{1}{\ell!} \sum_{r=0}^{\ell \wedge k} \sum_{\substack{B \subset \{1,\dots,k\}\\ |B|=r}} {\ell \choose r} \theta^{k+\ell-r} H_{k,\ell,B}], \quad (3.11)$$

where $H_{k,\ell,B}$ and $ESF(\mathbf{n})$ is defined as in Theorem 4.1. Set $g(z) = \exp(c_1 z^p + d_1 z) - 1$, where $c_1 = \frac{ix}{\Gamma(p)\sqrt{\theta}}$ and $d_1 = \frac{-ix}{\sqrt{\theta}}$ for $x \in R$. Then, $\mathbb{E}_x[\prod_{i=1}^{\infty}(1 + g(\tau X_i(t)))] = \mathbb{E}_x[\exp(c_1\tau^p H_p(t) + d_1\tau)]$ and the first term of (3.11) becomes

$$I = \exp\{\theta \int_0^\infty \frac{e^{-z}}{z} \left(\exp(\frac{ix}{\sqrt{\theta}} \frac{z^p}{\Gamma(p)} - \frac{ix}{\sqrt{\theta}} z) - 1\right) dz\}$$

$$= \exp\{\theta \int_0^\infty \frac{e^{-z}}{z} \left[\left(\frac{ix}{\sqrt{\theta}} \frac{z^p}{\Gamma(p)} - \frac{ix}{\sqrt{\theta}} z\right) - \frac{x^2}{2\theta} \left(\frac{z^p}{\Gamma(p)} - z\right)^2\right] dz + o(1)\}$$

$$= \exp\{-\frac{x^2}{2} (\sigma_p^2 + (p-1)^2) + o(1)\}, \quad \text{as } \theta \to \infty.$$

Next, we show that the second term of (3.11) converges to zero as $\theta \to \infty$. Since $|H_{k,\ell,B}| \leq \Gamma(\theta) (\int_0^\infty \frac{e^{-z}|g(z)|}{z} dz)^\ell$, we have

$$|H_{k,\ell,B}| \le \Gamma(\theta) \left(\int_0^\infty \frac{e^{-z}}{z} |\frac{x}{\sqrt{\theta}} (\frac{z^p}{\Gamma(p)} - z)| \, dz\right)^\ell \le \Gamma(\theta) (\frac{2x}{\sqrt{\theta}})^\ell.$$

 $\forall \ \theta \geq 1$ we have

$$\begin{split} \left|\sum_{\ell=1}^{\infty} \frac{1}{\ell!} \sum_{r=0}^{k \wedge k} \sum_{\substack{B \subset \{1,\dots,k\} \\ |B|=r}} \frac{\ell!}{(\ell-r)!} \frac{\theta^{k+\ell-r}}{\Gamma(\theta)} H_{k,\ell,B}\right| &\leq \sum_{\ell=1}^{\infty} \frac{1}{\ell!} \sum_{r=0}^{k \wedge k} \sum_{\substack{B \subset \{1,\dots,k\} \\ |B|=r}} \frac{\ell!}{(\ell-r)!} \theta^{k+\ell-r} (\frac{2x}{\sqrt{\theta}})^{\ell} \\ &= \theta^k \sum_{\ell=1}^{\infty} \sum_{r=0}^{k \wedge k} \sum_{\substack{B \subset \{1,\dots,k\} \\ |B|=r}} \frac{1}{(\ell-r)!} \theta^{\ell-r} (\frac{2x}{\sqrt{\theta}})^{\ell} \\ &= \theta^k \sum_{r=0}^k \binom{k}{r} \sum_{\ell=r \vee 1}^{\infty} \frac{1}{(\ell-r)!} \theta^{\ell-r} (\frac{2x}{\sqrt{\theta}})^{\ell} \\ &\leq \theta^k e^{2x\sqrt{\theta}} \sum_{r=0}^k \binom{k}{r} (\frac{2x}{\sqrt{\theta}})^r \\ &= \theta^k e^{2x\sqrt{\theta}} (1 + \frac{2x}{\sqrt{\theta}})^k \leq \theta^n e^{2x\sqrt{\theta}} (1 + 2x)^n. \end{split}$$

Following $\sum_{\substack{\mathbf{n}=(n_1,\dots,n_k)\\\sum_{i=1}^k n_i=n}} \psi_{\mathbf{n}}(x) = 1$ and $ESF(\mathbf{n}) = (n_1 - 1)! \cdots (n_k - 1)! \frac{\theta^k}{\theta_{(n)}}$, we have $\sum_{\substack{\mathbf{n}=(n_1,\dots,n_k)\\\sum_{i=1}^k n_i=n}} \psi_{\mathbf{n}}(x) ESF(\mathbf{n})^{-1} \leq \theta_{(n)}$. Thus the second term of (3.11) is bounded as

$$|\mathrm{II}| \leq \sum_{m=2}^{\infty} \frac{2m-1+\theta}{m!} e^{-\lambda_m t} \sum_{n=0}^m \binom{m}{n} (n+\theta)_{(m-1)} + \sum_{m=2}^{\infty} \frac{2m-1+\theta}{m!} e^{-\lambda_m t} \sum_{n=0}^m \binom{m}{n} (n+\theta)_{(m-1)} \theta_{(n)} \theta^n \mathrm{e}^{2x\sqrt{\theta}}.$$

If we can show that the above two series are uniformly convergent, then they converge to zero as $\theta \to \infty$ since every term in the series goes to 0.

First of all, we can see that

$$\sum_{m=2}^{\infty} \frac{2m-1+\theta}{m!} e^{-\lambda_m t} \sum_{n=0}^m \binom{m}{n} (n+\theta)_{(m-1)} \leq \sum_{m=2}^{\infty} \frac{(2m-1+\theta)(m+\theta)_{(m-1)}}{m!} e^{-\lambda_m t} 2^m$$
$$= \sum_{m=2}^{\infty} \frac{(m+\theta)_{(m)}}{m!} e^{-\lambda_m t} 2^m.$$

Since

$$(m+\theta)_{(m)}e^{-\lambda_m t} \le (\theta+2m-1)^m e^{-\frac{m(m-1+\theta)}{2}t} \le (\frac{2(\theta+2m-1)}{t(m-1+\theta)}e^{-1})^m \le (\frac{4}{te})^m.$$

Thus,

$$\sum_{m=2}^{\infty} \frac{2m-1+\theta}{m!} e^{-\lambda_m t} \sum_{n=0}^m \binom{m}{n} (n+\theta)_{(m-1)} \le \sum_{m=2}^{\infty} (\frac{4}{te})^m \frac{1}{m!} \le \exp\{\frac{8}{te}\} \le \infty,$$

i.e. the series converges uniformly. Similarly, we have the bound for the second term:

$$\sum_{m=2}^{\infty} \frac{2m-1+\theta}{m!} e^{-\lambda_m t} \sum_{n=0}^m \binom{m}{n} (n+\theta)_{(m-1)} \theta_{(n)} \theta^n e^{2x\sqrt{\theta}}$$

$$\leq \sum_{m=2}^{\infty} \frac{2m-1+\theta}{m!} e^{-\lambda_m t} \sum_{n=0}^m \binom{m}{n} \theta_{(n+m-1)} \theta^n e^{2x\sqrt{\theta}}$$

$$\leq \sum_{m=2}^{\infty} \frac{\theta_{(2m)}}{m!} e^{-\lambda_m t} (\theta+1)^m e^{2x\sqrt{\theta}}.$$

and

$$\theta_{(2m)}(\theta+1)^m e^{-\frac{m(m-1+\theta)}{2}t+2x\sqrt{\theta}} \le ((\theta+2m-1)(\theta+1)e^{-\frac{m-1+\theta}{2}t+2x\sqrt{\theta}/m})^m \le C^m.$$

The second series in II is uniformly convergent. Therefore, the right side of Laplace transform converges to $\exp\{-\frac{x^2}{2}(\sigma_p^2 + (p-1)^2)\}$. On the other hand, the left side becomes

$$\frac{1}{\Gamma(\theta)} \int_0^\infty e^{-\tau} \tau^{\theta-1} \mathbb{E}_x [\exp(c_1 \tau^p H_p(t) + d_1 \tau)] d\tau$$

= $(\int_0^\infty e^{-\theta \tau} \tau^{\theta-1} \mathbb{E}_x [\exp(c_1 \theta^p \tau^p H_p(t) + d_1 \theta \tau)] d\tau) (\int_0^\infty \tau^{\theta-1} e^{-\theta \tau} d\tau)^{-1}$
= $(\int_0^\infty e^{-\theta h(\tau)} \tau^{-1} \psi_p(x\tau^p) \exp(ix\sqrt{\theta}(\tau^p - \tau)) d\tau) (\int_0^\infty \tau^{\theta-1} e^{-\theta h(\tau)} d\tau)^{-1},$

where $h(\tau) = \tau - \log \tau - 1$. So we only need to prove that

$$\frac{\int_0^\infty e^{-\theta h(\tau)} \tau^{-1} \psi_p(x\tau^p) \exp(ix\sqrt{\theta}(\tau^p-\tau)) d\tau}{\int_0^\infty \tau^{\theta-1} e^{-\theta\tau} d\tau} - \psi_p(x) e^{-(p-1)^2 x^2/2} \to 0.$$

Readers can refer the rest proof to that of Theorem 6.2 in [38]. The only thing left is to show $\sup_{\theta>1} \mathbb{E}[|W_p(t)|] < \infty$. Since $q(t, x, y) \to 1$, as $\theta \to \infty$, we have $\sup_{\theta>1} q(t, x, y) \leq K, \forall \theta \geq 1$. Since

$$(\mathbb{E}_x[|W_p(t)|])^2 \le \mathbb{E}_x[|W_p(t)|^2] = \mathbb{E}[W_p^2q(t,x,Y)^2] \le K^2(Var(W_p) + (E[W_p])^2),$$

where W_p is defined as $W_p(t)$ with the replacement of $(X_i(t))_{i=1}^{\infty}$ by $Y = (Y_i)_{i=1}^{\infty}$ with the $PD(\theta)$ distribution. Then, $\operatorname{var}(W_p) + (E[W_p])^2 \to \sigma_p^2$, as $\theta \to \infty$, and $\sup_{\theta>1} \mathbb{E}[|W_p(t)|] < \infty$.

3.6 Symmetric Selection

The Poisson-Dirichlet distribution with selection $PD^{s,q}(\theta)$ is defined as an absolutely continuous distribution with respect to the Poisson-Dirichlet distribution $PD(\theta)$ with density

$$\frac{PD^{s,q}(\theta)(dx)}{PD(\theta)(dx)} = e^{sH_q(x)}/C, \qquad x = (x_1, x_2, \ldots) \in \nabla_{\infty},$$

where s is an arbitrary real number, C is the normalization constant and $H_q(x) = \sum x_i^q, q = 2, 3, \ldots$ The sampling formula with selection was obtained in [37] as

$$P_{\theta,s,q}(n_1,\ldots,n_k) = E_{PD(\theta)}[e^{sH_q(X)}\psi_{\mathbf{n}}(X)]/E_{PD(\theta)}[e^{sH_q(X)}]$$
$$= C^{-1}\theta^k \sum_{l=0}^{\infty} \frac{\theta^l}{l!}I_l(\mathbf{a}),$$

where

$$I_{l}(\mathbf{a}) = \int_{\Delta_{l+k}} \prod_{\alpha=1}^{k} (x_{\alpha}^{n_{\alpha}} e^{sx_{\alpha}^{q}}) \prod_{\alpha=k+1}^{k+l} (e^{sx_{\alpha}^{q}} - 1)(1 - \sum_{\beta=1}^{k+l} x_{\beta})^{\theta-1} \frac{dx_{1} \cdots dx_{k+l}}{x_{1} \cdots x_{k+l}}$$

and

$$C = \mathbb{E}[e^{sH_q(X)}] = 1 + \sum_{l=1}^{\infty} \frac{\theta^l}{l!} \int_{\Delta_l} \prod_{\alpha=1}^l (e^{sx_{\alpha}^q} - 1)(1 - \sum_{\beta=1}^l x_{\beta})^{\theta-1} \frac{dx_1 \cdots dx_l}{x_1 \cdots x_l}.$$

Since the correlation measure and sampling formula are determined by each other, we can obtain the correlation measure in the selection case. The proof is similar to that of in Theorem 3.1. It suffices to replace f by the polynomial functions in (3.8) and the correlation measure is found as follows. Denote

$$J_{l}(t) = \int_{\Delta_{l}} \prod_{j=1}^{l} (e^{st^{q}x_{j}^{q}} - 1)(1 - \sum_{\beta=1}^{l} x_{\beta})^{\theta-1} \frac{dx_{1} \cdots dx_{l}}{x_{1} \cdots x_{l}},$$

and $F_s(t) = \sum_{l=0}^{\infty} \frac{\theta^l}{l!} J_l(t)$. The correlation measure is given by

$$\mu_k(dx_1,\ldots,dx_k) = C^{-1}\theta^k \prod_{j=1}^k (e^{sx_j^q} x_j^{-1})(1-\sum_{j=1}^k x_j)^{\theta-1} F_s(1-\sum_{j=1}^k x_\alpha) \mathbf{1}_{\Delta_k}(x_1,\ldots,x_k) dx_1 \cdots dx_k.$$

We obtain the Laplace transform with selection in the next theorem.

Theorem 3.6. Suppose that g(.) is a measurable function such that

$$\lambda(g) := \inf\{\lambda > 0 : \int_0^\infty \frac{e^{-\lambda z}}{z} |g(z)| \, dz < \infty\} < \infty$$

The random element $(X_i)_{i=1}^{\infty}$ has the $PD^{s,q}(\theta)$ distribution. Then

$$\int_0^\infty e^{-\lambda t} t^{\theta-1} \mathbf{E}[\prod_{i=1}^\infty (1+g(tX_i))] dt = \frac{\Gamma(\theta)}{\lambda^\theta} + C^{-1} \sum_{k=1}^\infty \frac{\theta^k}{k!} M_k,$$

where

~

$$M_{k} = \int_{(0,\infty)^{k}} \prod_{j=1}^{k} (g(x_{j})x_{j}^{-1}e^{-\lambda x_{j}}) dx_{1} \cdots dx_{k}$$
$$\int_{0}^{\infty} e^{-\lambda t} t^{\theta-1} \prod_{j=1}^{k} e^{\frac{sx_{j}^{q}}{(t+\sum_{j=1}^{k} x_{j})^{q}}} F_{s}(\frac{t}{t+\sum_{j=1}^{k} x_{j}}) dt$$

Proof. Since we have got the correlation measure, it is obvious that

$$E[\prod_{i=1}^{\infty} (1+g(tX_i))] = 1 + C^{-1} \sum_{k=1}^{\infty} \frac{\theta^k}{k!} \int_{\Delta_k} \prod_{j=1}^k (e^{sx_j^q} g(tx_j) x_j^{-1}) (1-\sum_{j=1}^k x_j)^{\theta-1} F_s(1-\sum_{j=1}^k x_j) \, dx_1 \cdots dx_k.$$

Thus,

$$\begin{split} &\int_{0}^{\infty} e^{-\lambda t} t^{\theta-1} \mathbb{E}[\prod_{i=1}^{\infty} (1+g(tX_{i}))] dt \\ &= \frac{\Gamma(\theta)}{\lambda^{\theta}} + C^{-1} \sum_{k=1}^{\infty} \frac{\theta^{k}}{k!} \int_{0}^{\infty} e^{-\lambda t} t^{\theta-1} dt \\ &\int_{\Delta_{k}} \prod_{j=1}^{k} (e^{sx_{j}^{q}} g(tx_{j}) x_{j}^{-1}) (1-\sum_{j=1}^{k} x_{j})^{\theta-1} F_{s}(1-\sum_{j=1}^{k} x_{j}) dx_{1} \cdots dx_{k} \\ &= \frac{\Gamma(\theta)}{\lambda^{\theta}} + C^{-1} \sum_{k=1}^{\infty} \frac{\theta^{k}}{k!} \int_{(0,\infty)^{k}} \prod_{j=1}^{k} (g(x_{j}) x_{j}^{-1} e^{-\lambda x_{j}}) dx_{1} \cdots dx_{k} \\ &\int_{0}^{\infty} e^{-\lambda t} t^{\theta-1} \prod_{j=1}^{k} e^{\frac{sx_{j}^{q}}{(t+\sum_{j=1}^{k} x_{j})^{q}}} F_{s}(\frac{t}{t+\sum_{j=1}^{k} x_{j}}) dt. \end{split}$$

Since for 0 < x < 1, we have

$$|e^{sx^{q}} - 1| = x^{q} |\int_{0}^{s} e^{ux^{q}} du| \le (1 \lor e^{s})x^{q} |s|.$$

Then

$$\begin{aligned} |J_l(t)| &\leq \int_{\Delta_l} (1 \vee e^{sl})(t^q |s|)^l \prod_{i=1}^l x_i^q (1 - \sum_{i=1}^l x_i)^{\theta - 1} \frac{dx_1 \cdots dx_l}{x_1 \cdots x_l} \\ &= (1 \vee e^{sl})(t^q |s|)^l \frac{\Gamma(q)^l \Gamma(\theta)}{\Gamma(ql + \theta)}, \\ |F_s(t)| &\leq \sum_{l=0}^\infty \frac{\theta^l}{l!} |J_l(t)| \\ &\leq \sum_{l=0}^\infty \frac{\theta^l (1 \vee e^{sl})(t^q |s|)^l \Gamma(q)^l \Gamma(\theta)}{l! \Gamma(ql + \theta)} \\ &\leq \exp\{K(1 \vee e^{sl})t^q |s| \Gamma(q)\theta^{1 - q}\}, \end{aligned}$$

where K is a constant. Thus, $\prod_{j=1}^{k} e^{\frac{sy_j^q}{(t+\sum_{j=1}^{k} y_j)^q}} |F_s(\frac{t}{t+\sum_{j=1}^{k} y_j})| \le m(\theta, q, s)$, and the second term in the above i.e. the series can be controlled by

$$m(\theta, q, s) \sum_{k=1}^{\infty} \frac{\theta^k}{k!} \left(\int_0^\infty g(z) z^{-1} e^{-\lambda z} \, dz \right)^k \Gamma(\theta) / \lambda^{\theta}.$$

Since $\int_0^\infty e^{-\lambda x} |g(x)| x^{-1} dx < \infty$, the above series is absolutely convergent, which completes the proof.

Chapter 4

The Transformed Two-Parameter Poisson-Dirichlet Distribution

4.1 Abstract

The goal of this chapter is to investigate the impact of selection in the two-parameter model with overdominant selection with mutation going to infinity. To this end, we obtain a Gaussian limiting theorem for the homozygosity functions under the twoparameter Poisson-Dirichlet distribution with selection. The selection intensity is given by a multiple of the mutation rate raised to different powers, which leads to a phase transition in the limiting distribution depending on the relative strength of selection intensity over the mutation rate. This result suggests that the selection model resembles the neutral model when the mutation rate and selection intensity go to infinity together.

It should be noted that this chapter is originated from a published paper [67]. The references of the chapter are indexed to adapt to the thesis, therefore differ from those in the original paper.

4.2 Introduction

For $0 \leq \alpha < 1$ and $\theta > -\alpha$, let $U_k, k = 1, 2, ...$ be a sequence of independent random variables such that U_k has $Beta(1 - \alpha, \theta + k\alpha)$ distribution with density function given by

$$f(x) = \frac{\Gamma(\theta + 1 + (k - 1)\alpha)}{\Gamma(1 - \alpha)\Gamma(\theta + k\alpha)} x^{-\alpha} (1 - x)^{\theta + k\alpha - 1}, \ 0 < x < 1.$$

Set $(V_1, V_2, ...)$ as

$$V_1 = U_1, V_n = (1 - U_1) \cdots (1 - U_{n-1})U_n, \ n \ge 2.$$
(4.1)

Arrange (V_1, V_2, \ldots) in descending order and denote it as $P = (P_1, P_2, \ldots)$. Then the law of P is called the two-parameter Poisson-Dirichlet distribution denoted by $PD(\alpha, \theta)$. The law of (V_1, V_2, \ldots) is called the two-parameter GEM distribution denoted by $GEM(\alpha, \theta)$. $PD(\alpha, \theta)$ is a probability concentrating on the infinitedimensional ordered simplex

$$\nabla_{\infty} = \left\{ (x_1, x_2, \ldots) : x_1 \ge x_2 \ge \cdots \ge 0, \sum_{i=1}^{\infty} x_i = 1 \right\}$$

equipped with the topology inherited from $[0,1]^{\infty}$. $PD(0,\theta)$ corresponds to the Poisson-Dirichlet distribution with parameter θ .

Recall that in the infinite alleles model with overdominant selection intensity $\sigma >$ 0, the stationary distribution denoted by $PD^{\sigma}(\theta)$ is absolutely continuous with the Poisson-Dirichlet distribution $PD(\theta)$. The corresponding Radon-Nikodym derivative is given explicitly in a special selection case by [17] as

$$\frac{\mathrm{e}^{-\sigma H_2(\mathbf{x})}}{C(\sigma,\theta)},$$

where the population homozygosity $H_2(\mathbf{x}) = \sum_{i=1}^{\infty} x_i^2$, σ is the selection intensity, $\mathbf{x} = (x_1, x_2, \ldots) \in \nabla_{\infty}$ and $C(\sigma, \theta) = \int_{\nabla_{\infty}} e^{-\sigma H_2(\mathbf{x})} PD(\theta)(d\mathbf{x})$ is a normalizing constant. Here, we generalize the selection density to the two-parameter case. For convenience, we name it the two-parameter Poisson-Dirichlet distribution with selection though the two-parameter Poisson-Dirichlet distribution does not find a direct application in population genetics. Define the two-parameter Poisson-Dirichlet distribution with selection denoted by $PD^{\sigma}(\alpha, \theta)$ as

$$\frac{PD^{\sigma}(\alpha,\theta)}{PD(\alpha,\theta)}(d\mathbf{x}) = \frac{e^{-\sigma H_m(\mathbf{x})}}{C(\sigma,\theta,\alpha)}$$

where $H_m(\mathbf{x}) = \sum_{i=1}^{\infty} x_i^m$, $\mathbf{x} = (x_1, x_2, \ldots) \in \nabla_{\infty}$, $m = 2, 3, \ldots$ denotes the *m*th order homozygosity and $C(\sigma, \theta, \alpha) = \int_{\nabla_{\infty}} e^{-\sigma H_m(\mathbf{x})} PD(\alpha, \theta)(d\mathbf{x})$. For $p = 2, 3, \ldots$, let $X = (X_1, X_2, \ldots)$ have the distribution $PD(\alpha, \theta)$ and define

$$Z_{p,\theta} = \sqrt{\theta} \left(\frac{\Gamma(1-\alpha)}{\Gamma(p-\alpha)} \theta^{p-1} H_p(X) - 1 \right), \qquad (4.2)$$

where $H_p(X)$ denotes the *p*th homozygosity of X. Similarly for $Y = (Y_1, Y_2, ...) \sim PD^{\sigma}(\alpha, \theta)$, where $\sigma = c\theta^{m-\frac{1}{2}-\beta}$, $m = 2, 3, ..., \beta \geq 0$ and c is a constant, put

$$Z_{p,\theta}^{\beta} = \sqrt{\theta} \left(\frac{\Gamma(1-\alpha)}{\Gamma(p-\alpha)} \theta^{p-1} H_p(Y) - 1 \right).$$
(4.3)

The main ingredient used in the proof of our theorem is the following central limit theorem for $(Z_{p,\theta})_{p=2}^{\infty}$ obtained by Handa [38].

Lemma 4.1. Let $(Z_{2,\alpha}, Z_{3,\alpha}, \ldots) \in \mathbb{R}^{\infty}$ be multivariate normal with mean 0 and the covariance of $Z_{i,\alpha}$ and $Z_{j,\alpha}$ given by

$$\frac{\Gamma(1-\alpha)\Gamma(i+j-\alpha)}{\Gamma(i-\alpha)\Gamma(j-\alpha)} + \alpha - ij.$$
(4.4)

Then, as $\theta \to \infty$, we have

$$(Z_{2,\theta}, Z_{3,\theta}, \ldots) \Rightarrow (Z_{2,\alpha}, Z_{3,\alpha}, \ldots).$$

Here and in what follows, \Rightarrow denotes convergence in distribution. The next theorem is the main result of this chapter.

Theorem 4.2. Suppose that $Y = (Y_1, Y_2, ...) \sim PD^{\sigma}(\alpha, \theta)$ and $\sigma = c\theta^{m-\frac{1}{2}-\beta}, \beta \ge 0$, where c > 0 is a constant and m = 2, 3, ... Let $(Z_{2,\theta}^{\beta}, Z_{3,\theta}^{\beta}, ...) \in \mathbb{R}^{\infty}$ be defined as (4.3) and $(Z_{2,\alpha}, Z_{3,\alpha}, ...)$ given as the above lemma. As $\theta \to \infty$,

$$(Z_{2,\theta}^{\beta}, Z_{3,\theta}^{\beta}, \ldots) \Rightarrow \begin{cases} (Z_{2,\alpha}^{*}, Z_{3,\alpha}^{*}, \ldots), & \text{if } \beta = 0, \\ (Z_{2,\alpha}, Z_{3,\alpha}, \ldots), & \text{if } \beta > 0, \end{cases}$$

where $Z_{p,\alpha}^* = Z_{p,\alpha} - cb_p$, $b_p = (1 - \alpha)_{(p-1)} \operatorname{cov}(Z_{p,\alpha}, Z_{m,\alpha})$ and $(1 - \alpha)_{(p-1)} = \frac{\Gamma(p-\alpha)}{\Gamma(1-\alpha)} = (1 - \alpha) \cdots (p - 1 - \alpha)$ for $p = 2, 3, \ldots$

The uniform integrability for $\exp\{-tZ_{p,\theta}\}$ is essential in the proof of theorem. This is proved in the next section. In Section 4.4 we give the proof of Theorem 4.2 based on the establishing the convergence of the relevant characteristic function.

4.3 Uniform Integrability

To get the uniform integrability of $e^{-tZ_{p,\theta}}$, it suffices to prove the following lemma.

Lemma 4.3. Suppose that for any fixed t > 0 and $p \ge 2$, $Z_{p,\theta}$ is defined as in (4.2). Then there exists $\theta_0 = \theta_0(t) < \infty$ such that

$$\sup_{\theta > \theta_0} \mathcal{E}(e^{-tZ_{p,\theta}}) < \infty$$

We firstly need to introduce an important result of Pitman and Yor [57] as the following lemma.

Lemma 4.4 ([57]). Suppose that (X'_n) is given by the Poisson-Dirichlet distribution with parameter $\theta > 0$. Independent of (X'_n) , let $(U_{mn}), m = 1, 2, ...$ be a sequence of independent copies of (U_n) which has the two-parameter Poisson-Dirichlet distribution with $\theta = 0$ and $\alpha > 0$, i.e. $PD(\alpha, 0)$. Let (X_n) be defined by ranking the collection of products $\{X'_m U_{mn}, m, n \in \mathbb{N}\}$. Then (X_n) has the two-parameter Poisson-Dirichlet distribution with parameter θ, α , i.e. $PD(\alpha, \theta)$.

In virtue of the above representation, we can rewrite the homozygosity corresponding to the two-parameter P-D distribution as well as $Z_{p,\theta}$ in the form that keeps the i.i.d property.

Suppose that W_m , m = 1, 2, ... is a sequence of i.i.d random variables as copies of $\sum_{n=1}^{\infty} U_n^p$ for fixed $p \ge 2$. Therefore, $0 < W_m \le 1$ and its moments depend on α alone.

For any p = 2, 3, ..., by the GEM representation and exchangeability of H_p as a function of $(X_n)_1^{\infty}$ we have

$$H_p = \sum_{n=1}^{\infty} X_n^p = \sum_{m,n=1}^{\infty} (X'_m)^p U_{mn}^p \stackrel{\mathrm{D}}{=} \sum_{m=1}^{\infty} (X'_m)^p W_m = V_1^p W_1 + (1 - V_1)^p \tilde{H}_p,$$

where $V_1 \sim \text{Beta}(1, \theta)$ and \tilde{H}_p is a random variable with the same distribution as H_p and independent of V_1, W_1 . Here $\stackrel{\text{D}}{=}$ denotes the equality in distribution.

and independent of V_1, W_1 . Here $\stackrel{\text{D}}{=}$ denotes the equality in distribution. Since $Z_{p,\theta} = \sqrt{\theta} (\frac{\Gamma(1-\alpha)}{\Gamma(p-\alpha)} \theta^{p-1} H_p - 1) = \sqrt{\theta} (\frac{\theta^{p-1}}{(1-\alpha)_{(p-1)}} H_p - 1)$, we deduce that

$$(1-\alpha)_{(p-1)}Z_{p,\theta} = \sqrt{\theta}(\theta^{p-1}H_p - (1-\alpha)_{(p-1)}) \stackrel{\mathrm{D}}{=} \sqrt{\theta}(\theta^{p-1}V_1^pW_1 + \theta^{p-1}(1-V_1)^p\tilde{H}_p - (1-\alpha)_{(p-1)}) = \sqrt{\theta}(\theta^{p-1}V_1^pW_1 + (1-V_1)^p(1-\alpha)_{(p-1)} - (1-\alpha)_{(p-1)}) + (1-V_1)^p(\theta^{p-1}\tilde{H}_p - (1-\alpha)_{(p-1)})) = f_p(V_1, W_1) + (1-\alpha)_{(p-1)}(1-V_1)^p\tilde{Z}_{p,\theta},$$

where

$$f_p(V_1, W_1) = \sqrt{\theta} (\theta^{p-1} V_1^p W_1 - (1 - \alpha)_{(p-1)} (1 - (1 - V_1)^p)), \qquad (4.5)$$

and $\tilde{Z}_{p,\theta} = \sqrt{\theta} \left(\frac{\theta^{p-1}}{(1-\alpha)_{(p-1)}} \tilde{H}_p - 1 \right)$ with the same distribution as $Z_{p,\theta}$ is independent of V_1, W_1 . Since $0 < 1 - (1 - V_1)^p \le 2^p V_1$ and $0 < W_1 \le 1$,

$$\frac{1}{\sqrt{\theta}}(\theta^{p}V_{1}^{p}W_{1} - (1-\alpha)_{(p-1)}2^{p}\theta V_{1}) \le f_{p} \le \frac{1}{\sqrt{\theta}}(\theta^{p}V_{1}^{p}W_{1} + (1-\alpha)_{(p-1)}2^{p}\theta V_{1}).$$
(4.6)

Also,

$$|f_p(V_1, W_1)| \le \frac{1}{\sqrt{\theta}} (\theta^p V_1^p + (1 - \alpha)_{(p-1)} 2^p \theta V_1).$$

For $k \geq 1$,

$$\begin{split} \mathbf{E}|f_{p}(V_{1},W_{1})|^{k} &\leq \theta^{-\frac{k}{2}} \mathbf{E}(\theta^{p}V_{1}^{p} + (1-\alpha)_{(p-1)}2^{p}\theta V_{1})^{k} \\ &\leq \theta^{-\frac{k}{2}}2^{k}[\mathbf{E}(\theta V_{1})^{pk} + (1-\alpha)_{(p-1)}^{k}2^{pk}\mathbf{E}(\theta V_{1})^{k}] \\ &\leq (2/\sqrt{\theta})^{k}[(pk)! + (1-\alpha)_{(p-1)}^{k}2^{pk}k!] \sim O(\theta^{-\frac{k}{2}}), \end{split}$$

where in the last step we use $0 \le E(\theta V_1)^j = \frac{j!\theta^j}{(\theta+1)\cdots(\theta+j)} \le j!, j \ge 1.$

Lemma 4.5. For $p \ge 2$, let $f_p(V_1, W_1)$ be defined as in (4.5), where $V_1 \sim Beta(1, \theta)$, and W_1 has the same distribution as the pth homozygosity corresponding to alleles frequencies with the distribution $PD(\alpha, 0)$. Then, for $\theta > 1$ and each $j \ge 1$, there is a positive function $g_{p,j}(t, W_1)$ increasing in t such that, for all t > 0,

$$|\sum_{k=j}^{\infty} \frac{(-tf_p)^k}{k!}| \le t^j g_{p,j}(t, W_1) |f_p|^j.$$

Proof. Note that

$$|\sum_{k=j}^{\infty} \frac{(-tf_p)^k}{k!}| = |tf_p|^j |\sum_{k=j}^{\infty} \frac{(-tf_p)^{k-j}}{k!}| = t^j |f_p|^j g_j(tf_p),$$

where

$$0 < g_j(x) = \sum_{k=j}^{\infty} \frac{(-x)^{k-j}}{k!} = \begin{cases} \frac{e^{-x} - \sum_{k=0}^{j-1} (-x)^k / k!}{(-x)^j}, & x \neq 0, \\ \frac{1}{j!}, & x = 0. \end{cases}$$

Define $f(x) = W_1 x^p - (1 - \alpha)_{(p-1)} 2^p x$. Hence f has a lower bound on $[0, \infty)$. Let $x^* = (\frac{(1-\alpha)_{(p-1)} 2^p}{W_1 p})^{\frac{1}{p-1}}$ be the point in $[0, \infty)$ at which f(x) achieves its minimum, that is, $f(x^*) = W_1^{-\frac{1}{p-1}} (\frac{(1-\alpha)_{(p-1)} 2^p}{p})^{\frac{p}{p-1}} (1-p) < 0$. By (4.6), $f_p \ge \frac{f(\theta V_1)}{\sqrt{\theta}} \ge \frac{f(x^*)}{\sqrt{\theta}} > f(x^*)$

for $\theta > 1$. Since $g_j(x)$ is a decreasing function of $x \in (-\infty, \infty)$, and $g_j(tf_p) \leq g_j(tf(x^*)) =: g_{p,j}(t, W_1)$ for all t > 0.

The next lemma shows that the moments of $Z_{p,\theta}$ are uniformly bounded for large θ .

Lemma 4.6. Let $Z_{p,\theta}$ be defined by (4.2). Then, for all integers $r \ge 1$ and $p \ge 2$,

$$\sup_{\theta>1} \mathcal{E}(|Z_{p,\theta}|^r) < \infty.$$
(4.7)

Proof. Suppose that $(X'_1, X'_2, \ldots) \sim \Pi_{\theta}$. Set $H'_p = \sum_n (X'_n)^p$, and

$$Z'_{p,\theta} = \sqrt{\theta} \left(\frac{\theta^{p-1}}{(p-1)!}H'_p - 1\right).$$

By the sampling formulas of the Poisson-Dirichlet distribution and the two-parameter Poisson-Dirichlet distribution we know that $E(Z'_{p,\theta})^{2r}$ and $E(Z_{p,\theta})^{2r}$ have the same highest finite order of θ . Thus,

$$\lim_{\theta \to \infty} \frac{\mathrm{E}(Z_{p,\theta})^{2r}}{\mathrm{E}(Z'_{p,\theta})^{2r}} = k < \infty$$

Following the result obtained in Lemma 5 of [42], we know that

$$\sup_{\theta>1} \mathbf{E} |Z'_{p,\theta}|^r < \infty.$$

Since $\mathrm{E}|Z_{p,\theta}|^r \leq \sqrt{\mathrm{E}(Z_{p,\theta})^{2r}}$, it follows that $\sup_{\theta>1} \mathrm{E}|Z_{p,\theta}|^r < \infty$.

Now we are ready to give the proof of Lemma 2.1. Proof of Lemma 2.1. For each integer r, define

$$\tilde{Z}_{\theta} = (1-\alpha)_{(p-1)}\tilde{Z}_{p,\theta}$$
 and $\mu_{r,\theta} = (1-\alpha)_{(p-1)}^r \mathbb{E}(Z_{p,\theta}^r).$

Since $\sup_{\theta>1} \mathbb{E}(|Z_{p,\theta}|^r) < \infty$ for each r, we only need to show that, for t > 0, there exist finite values l = l(t) and $\theta_0 = \theta_0(l, t)$ such that

$$\sup_{\theta > \theta_0} |\sum_{r=l}^{\infty} \frac{(-t)^r}{r!} \mu_{r,\theta}| < \infty.$$

It is obvious that

$$\begin{split} [(1-\alpha)_{(p-1)}Z_{p,\theta}]^r &\stackrel{\mathrm{D}}{=} (f_p(V_1, W_1) + (1-\alpha)_{(p-1)}(1-V_1)^p \tilde{Z}_{p,\theta})^r \\ &= \sum_{k=0}^r \binom{r}{k} f_p^k (V_1, W_1) [(1-\alpha)_{(p-1)} \tilde{Z}_{p,\theta}]^{r-k} (1-V_1)^{p(r-k)} \\ &= (1-V_1)^{pr} [(1-\alpha)_{(p-1)} \tilde{Z}_{p,\theta}]^r \\ &+ \sum_{k=1}^r \binom{r}{k} f_p^k (V_1, W_1) (1-V_1)^{p(r-k)} [(1-\alpha)_{(p-1)} \tilde{Z}_{p,\theta}]^{r-k}. \end{split}$$

Taking expectation on both sides and using independence, we have

$$\mu_{r,\theta} = \mathcal{E}(1-V_1)^{pr}\mu_{r,\theta} + \sum_{k=1}^r \binom{r}{k} \mathcal{E}(f_p^k(V_1, W_1)(1-V_1)^{p(r-k)})\mu_{r-k,\theta}.$$

Thus,

$$\mu_{r,\theta} = \frac{pr+\theta}{pr} \sum_{k=1}^{r} \binom{r}{k} E(f_p^k(V_1, W_1)(1-V_1)^{p(r-k)}) \mu_{r-k,\theta}.$$

Define two sequences of random variables:

$$M_k := \frac{f_p^k}{k!}, \qquad N_l := \frac{(\tilde{Z}_{\theta}(1-V_1)^p)^l}{l!}.$$

Therefore,

$$\begin{aligned} |\sum_{r=l}^{\infty} \frac{(-t)^{r}}{r!} \mu_{r,\theta}| &= |\sum_{r=l}^{\infty} \frac{pr+\theta}{pr} \frac{(-t)^{r}}{r!} \sum_{k=1}^{r} \binom{r}{k} \mu_{r-k,\theta} \mathbb{E}(f_{p}^{k}(V_{1},W_{1})(1-V_{1})^{p(r-k)})| \\ &= |\mathbb{E}(\sum_{r=l}^{\infty} \sum_{k=1}^{r} (1+\frac{\theta}{pr}) M_{k} N_{r-k}(-t)^{r})| \\ &\leq |\mathbb{E}(\sum_{r=l}^{\infty} \sum_{k=1}^{r} M_{k} N_{r-k}(-t)^{r})| + |\mathbb{E}(\sum_{r=l}^{\infty} \sum_{k=1}^{r} \frac{\theta}{pr} M_{k} N_{r-k}(-t)^{r}))| \\ &=: |\mathbb{E}(A_{1})| + |\mathbb{E}(A_{2})|. \end{aligned}$$

By the inequality in Lemma 6 of [42] and interchanging the order of summation, we obtain ∞

$$|A_1| \le |\sum_{j=1}^{\infty} (-t)^j M_j \sum_{q=0}^{\infty} (-t)^q N_q| + t K_{1l}(t),$$

where

$$K_{1l}(t) := \sum_{j=1}^{l-1} |M_j| t^{j-1} \sum_{i=0}^{l-j-1} |N_i| t^i$$

is an increasing function of t. By Lemma 4 we obtain

$$\left|\sum_{j=1}^{\infty} (-t)^{j} M_{j}\right| = \left|\sum_{j=1}^{\infty} \frac{(-tf_{p})^{j}}{j!}\right| \le tg_{p,1}(t, W_{1})|f_{p}|,$$

and

$$|\sum_{q=0}^{\infty} (-t)^q N_q| = \exp\{-t\tilde{Z}_{\theta}(1-V_1)^p\}$$

$$\leq \max\{1, \exp\{-t\tilde{Z}_{\theta}\}\}$$

$$\leq 1 + \exp\{-t\tilde{Z}_{\theta}\}.$$

Hence

$$|A_1| \le tg_{p,1}(t, W_1)|f_p|(1 + \exp\{-t\tilde{Z}_{\theta}\}) + tK_{1l}(t).$$

It follows that

$$\begin{aligned} |\mathbf{E}(A_1)| &\leq \mathbf{E}|A_1| \\ &\leq t\mathbf{E}(g_{p,1}(t, W_1)|f_p|) + t\mathbf{E}(f_{p,1}(t, W_1)|g_p|)\mathbf{E}\exp\{-t\tilde{Z}_{\theta}\} + t\mathbf{E}(K_{1l}(t)). \end{aligned}$$

Since $E|f_p(V_1, W_1)|^k \sim O(\theta^{-\frac{k}{2}})$ and $Eg_{p,1}(t, W_1)$ only depends on α , there exists $\theta_0 = \theta_0(t) < \infty$ such that $tE(g_{p,1}(t, W_1)|f_p|) < \frac{1}{4}$ for all $\theta > \theta_0(t)$ by the Hölder's inequality. In addition, $K_{1l}(t)$ is a sum of a finite number of random variables, and note that mean of each random variable is bounded for all θ . Thus, $\sup_{\theta > 1} E(K_{1l}(t)) < \theta$

 ∞ . Set $\alpha_1(t) = \sup_{\theta > 1} t \mathbb{E}(K_{1l}(t)) + \frac{1}{4}$. Then, for $\theta > \theta_0(t)$,

$$|\mathbf{E}(A_1)| \le \frac{1}{4} \mathbf{E} \exp\{-t\tilde{Z}_{\theta}\}) + \alpha_1(t) = \frac{1}{4} \sum_{r=0}^{\infty} \frac{(-t)^r}{r!} \mu_{r,\theta} + \alpha_1(t).$$

We also have the same inequality for A_2 :

$$|\mathbf{E}(A_2)| \le \frac{1}{4} \sum_{r=0}^{\infty} \frac{(-t)^r}{r!} \mu_{r,\theta} + \alpha_2(t).$$

Since A_2 involves the factor $\frac{\theta}{pr}$, it needs a more delicate estimation but the argument is similar to that of A_1 . Readers can find a similar detailed proof in Lemma 9 of [42], which we will not repeat here.

By the inequalities for A_1 and A_2 as above, we have

$$\left|\sum_{r=l}^{\infty} \frac{(-t)^{r}}{r!} \mu_{r,\theta}\right| \leq \frac{1}{2} \sum_{r=0}^{\infty} \frac{(-t)^{r}}{r!} \mu_{r,\theta} + \alpha_{1}(t) + \alpha_{2}(t).$$

Thus,

$$\left|\sum_{r=l}^{\infty} \frac{(-t)^{r}}{r!} \mu_{r,\theta}\right| - \frac{1}{2} \sum_{r=l}^{\infty} \frac{(-t)^{r}}{r!} \mu_{r,\theta} \le \frac{1}{2} \sum_{r=0}^{l-1} \frac{t^{r}}{r!} |\mu_{r,\theta}| + \alpha_{1}(t) + \alpha_{2}(t).$$

It follows that

$$\frac{1}{2} \left| \sum_{r=l}^{\infty} \frac{(-t)^r}{r!} \mu_{r,\theta} \right| \le \frac{1}{2} \sum_{r=0}^{l-1} \frac{t^r}{r!} |\mu_{r,\theta}| + \alpha_1(t) + \alpha_2(t).$$

Therefore, $\left|\sum_{r=l}^{\infty} \frac{(-t)^r}{r!} \mu_{r,\theta}\right|$ is bounded by a sum of a finite number of terms, which are uniformly bounded for θ . It can be concluded that

$$\sup_{\theta > \theta_0} |\sum_{r=l}^{\infty} \frac{(-t)^r}{r!} \mu_{r,\theta}| < \infty.$$

4.4 The Proof of Theorem

Proof. For $\sigma = c\theta^{m-\frac{1}{2}-\beta}$, $\beta \ge 0$ and $m = 2, 3, \ldots$, we know that the characteristic function of $Z_{p,\theta}^{\beta}$ can be calculated as follows:

$$\psi_p(x) = \operatorname{E}[\exp(ixZ_{p,\theta}^{\beta})] = \frac{\operatorname{E}[e^{ixZ_{p,\theta}}e^{-\sigma H_m}]}{\operatorname{E}[e^{-\sigma H_m}]} = \frac{\operatorname{E}[e^{ixZ_{p,\theta}-c(1-\alpha)_{(m-1)}\theta^{-\beta}Z_{m,\theta}}]}{\operatorname{E}[e^{-c(1-\alpha)_{(m-1)}\theta^{-\beta}Z_{m,\theta}}]}$$

Based on Lemma 1.1,

$$(Z_{2,\theta}, Z_{3,\theta}, \ldots) \Rightarrow (Z_{2,\alpha}, Z_{3,\alpha}, \ldots)$$
as θ goes to infinity, where $(Z_{2,\alpha}, Z_{3,\alpha}, \ldots)$ is multivariate normal with mean 0 and the covariance matrix Σ given by (4.4).

For $\beta = 0$, combining the uniform integrability of $e^{-tZ_{p,\theta}}$ with $Z_{p,\theta} \Rightarrow Z_{p,\alpha}$ for any $p = 2, 3, \ldots$ we have

$$\mathbf{E}[\mathrm{e}^{-c(1-\alpha)_{(m-1)}Z_{m,\theta}}] \to \mathbf{E}[\mathrm{e}^{-c(1-\alpha)_{(m-1)}Z_{m,\alpha}}],$$

as $\theta \to \infty$. Therefore,

$$\mathbf{E}[\mathbf{e}^{ixZ_{p,\theta}-c(1-\alpha)_{(m-1)}Z_{m,\theta}}] \to \mathbf{E}[\mathbf{e}^{ixZ_{p,\alpha}-c(1-\alpha)_{(m-1)}Z_{m,\alpha}}]$$

It follows that when θ goes to infinity,

$$\psi_p(x) \to \exp\{-\frac{x^2}{2}\operatorname{var} Z_{p,\alpha} - icx(1-\alpha)_{(m-1)}\operatorname{cov}(Z_{p,\alpha}, Z_{m,\alpha})\}.$$

Define $b_p = (1 - \alpha)_{(m-1)} \operatorname{cov}(Z_{p,\alpha}, Z_{m,\alpha})$ and let $Z_{p,\alpha}^* = Z_{p,\alpha} - cb_p$, it follows that

$$Z_{p,\theta}^{\beta} \Rightarrow Z_{p,\alpha}^{*} \text{ for } p = 2, 3, \dots, \text{ as } \theta \to \infty.$$

Using the same method of calculating the characteristic function of finite linear combination of $Z_{p,\theta}$, we obtain

$$(Z_{2,\theta}^{\beta}, Z_{3,\theta}^{\beta}, \ldots) \Rightarrow (Z_{2,\alpha}^{*}, Z_{3,\alpha}^{*}, \ldots), \text{ as } \theta \to \infty.$$

For $\beta > 0, \theta > 1$,

$$\exp\{-c(1-\alpha)_{(m-1)}\theta^{-\beta}Z_{m,\theta}\} \le \max\{\exp\{-c(1-\alpha)_{(m-1)}Z_{m,\theta}\}, 1\}.$$

From the dominated convergence theorem, we have

$$\mathbf{E}[\mathrm{e}^{-c(1-\alpha)_{(p-1)}\theta^{-\beta}Z_{q,\theta}}] \to 1,$$

as $\theta \to \infty$. Therefore, $(Z_{2,\theta}^{\beta}, Z_{3,\theta}^{\beta}, \ldots) \Rightarrow (Z_{2,\alpha}, Z_{3,\alpha}, \ldots)$, as $\theta \to \infty$. \Box

Chapter 5

Transition Function of the Two-Parameter Diffusion Process

5.1 Abstract

To characterize a Markov process, we rely on three basic tools: transition function, semigroup and generator. Ethier [13] derived the transition function of the IMNA model through an expansion of eigenfunctions. In this chapter, we use this method to derive the transition function for the two-parameter extension of the IMNA model. An explicit transition density is given in Section 5.3 with respect to the two-parameter Poisson-Dirichlet distribution being the stationary distribution of the process. We give .

Combining this transition density with the two-parameter correlation measure derived in Chapter 3, we obtain the probability generating function of random sampling from the two-parameter model at a fixed time point in Section 5.4. This result can be further used to compare with the equilibrium result which corresponds to the Pitman sampling formula.

It should be noted that this chapter is part of a published paper coauthored with Professor Feng, Professor Sun and Professor Wang [27], to which the author of the thesis is an equal contributor. The references of the chapter are indexed to adapt to the thesis, and therefore they differ from those in the original paper.

5.2 Introduction

Consider a continuous time Markov process $\{X_t, t \geq 0\}$ with continuous sample paths and state space S. Such process is also called the diffusion process. Given the past history up to time s, the conditional probability

$$P(X_t \in dy | X_s = x, X_{u_i} = x_{u_i}, 0 \le u_i < s, x_{u_i} \in S, \forall i = 1, 2, \ldots)$$

of the process at a future time t is given by the transition function

$$P(X_t \in dy | X_s = x) = P(s, x, t, dy).$$

The process considered here is time homogeneous in a sense that

$$P(s, x, t, dy) = P(0, x, t - s, dy).$$

For convenience, denote the transition probability P(0, x, t, dy) by P(t, x, dy). The corresponding Markov semigroup operator is a family of linear operators $T_t, t \ge 0$ on C(S) the space of continuous functions on S such that

$$T_t f(x) = E_x[f(X_t)] = \int_S f(y) P(t, x, dy).$$

The infinitesimal generator L of the semigroup T_t is defined as

$$Lf = \lim_{t \to 0} \frac{T_t f - f}{t}$$

with properly defined domain. Assume that the law of X_0 is given by a probability measure μ on S satisfying

$$\int_{S} P(t, x, dy) \mu(dx) = \mu(dy)$$

for any t > 0. Then, the process X_t starting with μ has the same distribution at any time as the initial state. Such process is called a stationary distribution and μ is the stationary distribution. Equivalently, under the semigroup T_t of the process the stationary measure μ satisfies

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

for $f \in C(S)$. Finally, the process with Markov semigroup T_t is reversible with a probability measure μ such that

$$\int_{S} f(x)T_{t}g(x)\,\mu(dx) = \int_{S} g(x)T_{t}f(x)\,\mu(dx)$$

for all $f, g \in C(S)$.

For any $0 \le \alpha < 1$ and $\theta > -\alpha$, the two-parameter extension model is a symmetric diffusion process taking values in infinite dimensional ordered simplex

$$\nabla_{\infty} = \{(x_1, x_2, \ldots) : x_1 \ge x_2 \ge \cdots \ge 0, \sum_{i=1}^{\infty} x_i = 1\}$$

with generator

$$L_{\theta,\alpha} = \frac{1}{2} \left\{ \sum_{i,j=1}^{\infty} x_i (\delta_{ij} - x_j) \frac{\partial^2}{\partial x_i \partial x_j} - \sum_{i=1}^{\infty} (\theta x_i + \alpha) \frac{\partial}{\partial x_i} \right\}.$$
 (5.1)

The domain of $L_{\theta,\alpha}$ is $\mathcal{D}(L_{\theta,\alpha}) = span\{1, \varphi_2, \varphi_3, \ldots\} \subset C(\nabla_{\infty})$, where $\varphi_n(x) = \sum_{i=1}^{\infty} x_i^n$ is defined on ∇_{∞} and extends continuously to $\overline{\nabla}_{\infty}$ which is the closure of ∇_{∞} in $[0, 1]^{\infty}$. The process was first established by Petrov [52] and lately has been further studied by Feng and Sun [26].

The process is reversible with respect to the two-parameter Poisson-Dirichlet distribution $PD(\alpha, \theta)$ which is defined as follows. Let U_k , k = 1, 2, ..., be a sequence of independent random variables such that U_k has $Beta(1 - \alpha, \theta + k\alpha)$ distribution with density function given by

$$f(x) = \frac{\Gamma(\theta + 1 + (k - 1)\alpha)}{\Gamma(1 - \alpha)\Gamma(\theta + k\alpha)} x^{-\alpha} (1 - x)^{\theta + k\alpha - 1}, \ 0 < x < 1.$$

Set

$$V_1 = U_1, V_n = (1 - U_1) \cdots (1 - U_{n-1})U_n, \ n \ge 2.$$

Arrange $(V_1, V_2, ...)$ in descending order and denote it as $P = (P_1, P_2, ...)$. Then the law of P is called the two-parameter Poisson-Dirichlet distribution.

The case $\alpha = 0$ corresponds to the IMNA diffusion model constructed in [15], where the generator of the process is given by

$$L = \frac{1}{2} \sum_{i,j=1}^{\infty} x_i (\delta_{ij} - x_j) \frac{\partial^2}{\partial x_i \partial x_j} - \frac{\theta}{2} \sum_{i=1}^{\infty} x_i \frac{\partial}{\partial x_i}$$
(5.2)

with the same domain of $L_{\theta,\alpha}$. The process has the unique reversible measure given by the Poisson-Dirichlet distribution with parameter θ i.e. $PD(0,\theta)$. From [13] we know that the complete set of eigenvalues of L consists of $0, -\lambda_2, \lambda_3, \ldots$, where $\lambda_m = m(m-1+\theta)/2, m \ge 2$. 0 is a simple eigenvalue and for $m \ge 2$, the multiplicity of λ_m is $\pi(m) - \pi(m-1)$, where $\pi(m)$ is the number of partitions of the integer m. The transition probability is obtained in [13] and the transition density also exists with respect to the invariant measure.

We could compare this model with its labeled version which is the Fleming-Viot process with parent independent mutation. Let S denote the type space which is a Polish space and $\mathcal{P}(S)$ be the set of probability measure on S equipped with the usual weak topology. $C_b(S)$ denotes the set of bounded, continuous functions on S. Fleming-Viot process is a $\mathcal{P}(S)$ -valued process with the generator

$$(\mathcal{L}\phi)(\mu) = \frac{1}{2} \int_{S} \int_{S} \mu(dx) (\delta_{x}(dy) - \mu(dy)) \frac{\delta^{2}\phi(\mu)}{\delta\mu(x)\delta\mu(y)} + \int_{S} \mu(dx) A\left(\frac{\delta\phi(\mu)}{\delta\mu(\cdot)}\right)(x),$$
(5.3)

where $\delta\phi(\mu)/\delta\mu(x) = \lim_{\epsilon\to 0+} \epsilon^{-1}(\phi(\mu + \epsilon\delta_x) - \phi(\mu))$ and A called the mutation operator is the generator of a Feller semigroup on C(S). The domain of \mathcal{L} is defined as follows $\mathcal{D}(\mathcal{L}) = \{\phi : \phi(\mu) = f(\langle \varphi_1, \mu \rangle, \dots, \langle \varphi_k, \mu \rangle), f \in C_b^{\infty}(\mathbf{R}^k), \varphi_1, \dots, \varphi_k \in$ $C_b(S) \cap D(A), \mu \in \mathcal{P}(S), k \geq 1\}$, where $\langle \varphi, \mu \rangle = \int_S \varphi \, d\mu$ and $C_b^{\infty}(\mathbf{R})$ denotes the set of all bounded, infinitely differentiable functions on **R**. Particularly we consider the parent independent mutation operator which is given by

$$(Af)(x) = \frac{1}{2}\theta \int_{S} (f(\xi) - f(x))\nu_0(d\xi), \quad f \in C_b(S)$$
(5.4)

with $\theta > 0$ and $\nu_0 \in \mathcal{P}(S)$. It is shown that the stationary distribution for the labeled model is the distribution of $\sum_{i=1}^{\infty} P_i \delta_{\xi_i}$ where $(P_1, P_2, ...)$ has the Poisson-Dirichlet distribution $PD(\theta)$, and $\xi_1, \xi_2, ...$ are i.i.d. random variables on S distributed as ν_0 independent of $(P_1, P_2, ...)$. The distribution of $\sum_{i=1}^{\infty} P_i \delta_{\xi_i}$ is also known as Dirichlet process. An explicit transition function of the process is obtained in [14] and transition density does not exist. The transition function implies that the process concentrates on pure atomic probability measure on S. Stannat [59] shows that a log-Sobolev inequality holds for the Fleming-Viot process with parent independent mutation if and only if the number of alleles or type space S is finite.

To associate these two models, we consider the mapping $\Phi(\mu)$ defined as the sequence of descending order statistics of the masses of the atoms of μ . Clearly the

image of the Fleming-Viot process defined as above under the projection of Φ is exactly the neutral infinite alleles model with generator L characterized by [15]. But in the case of $S = [0, 1], \nu_0$ being the Lebesgue measure, the set of eigenvalues of the generator \mathcal{L} has one more element $-\lambda_1$ than that of L. Comparing these two models, it implies that the ordering and un-labeling increase the spectral gap by an amount of $1 + \theta/2$.

In the two-parameter case $L_{\theta,\alpha}$, the structure of eigenvalues and eigenfunctions are similar to the IMNA diffusion model. The description of this fact will be given in the following theorem. We aim to obtain an explicit expression of the transition function. The Pitman sampling formula is of vital importance in our derivation, and will be presented in the next section with various known properties of this formula.

5.3 Preliminaries

In this section, we summarize several important results that are known or are easy to derive. First, we need to introduce some notation. Let \mathbb{N} denote the set of non-negative integers. Set $a_{(0)} = a_{[0]} = 1$ and for $n \in \mathbb{N}$,

$$a_{(n)} = a(a+1)\cdots(a+n-1), \ a_{[n]} = a(a-1)\cdots(a-n+1), \ n \ge 1.$$

For $k \in \mathbb{N}$ set

$$J_k = \{ \vec{n} = (n_1, \cdots, n_k) \in \mathbb{N}^k : n_1 \ge \cdots n_k \ge 2 \}$$

and $J = \bigcup_{k=1}^{\infty} J_k$. Let $|\vec{n}| = n_1 + \cdots + n_k$ and $\varphi_{\vec{n}} = \varphi_{n_1} \cdots \varphi_{n_k}$, $\vec{n} \in J_k$. For each $\ell \in \mathbb{N}, \ \vec{m} = (m_1, \ldots, m_\ell) \in \mathbb{N}_\ell$, set

$$p^{\circ}_{\vec{m}}(x) = \sum_{i_1,\dots,i_\ell \text{distinct}} x^{m_1}_{i_1} \cdots x^{m_\ell}_{i_\ell}$$

and

$$p_{\vec{m}}(x) = \binom{m}{m_1 \cdots m_\ell} \frac{1}{\alpha_1! \cdots \alpha_m!} p_{\vec{m}}^{\circ}(x), \quad x \in \nabla_{\infty}.$$

where $m = m_1 + \cdots + m_\ell$ and $\alpha_i = \#\{j : m_j = i\}$. For each $n \in \mathbb{N}$ define $P_n \in C(\nabla_\infty \times \nabla_\infty)$ by

$$P_n(x,y) = \sum_{\vec{n} \in I: |\vec{n}|=n} \frac{p_{\vec{n}}(x)p_{\vec{n}}(y)}{\int_{\nabla_{\infty}} p_{\vec{n}}(x) PD(\alpha,\theta)(dx)},$$

where $I = \bigcup_{i=1}^{\infty} I_l$ and $I_l = \{ \vec{n} = (n_1, \cdots, n_l) \in \mathbb{N}^l : n_1 \ge \cdots \ge n_l \ge 1 \}.$

Theorem 5.1. (1) The generator $L_{\theta,\alpha}$ defined on $\mathcal{D}(L_{\theta,\alpha})$ is closable in $C(\overline{\nabla}_{\infty})$. The closure, denoted by $L_{\theta,\alpha}$ for notational simplicity, generates a $\overline{\nabla}_{\infty}$ -valued diffusion process $X_{\theta,\alpha}(t)$, the two-parameter unlabeled infinite-alleles diffusion process;

- (2) The process $X_{\theta,\alpha}(t)$ is reversible with respect to $PD(\alpha, \theta)$;
- (3) For each $k \geq 1$, $\vec{n} = (n_1, \ldots, n_l)$ in \mathbb{N}^l ,

$$\int_{\overline{\nabla}_{\infty}} p_{\vec{n}}(x) PD(\alpha, \theta) dx = \frac{n!}{\theta_{(n)}} \prod_{r=0}^{l-1} (\theta + r\alpha) \prod_{j=1}^{n} \frac{((1-\alpha)_{(j-1)})^{\alpha_j}}{(j!)^{\alpha_j} \alpha_j!},$$

where $l = \sum_{j=1}^{n} a_j$.

Proof. (1) and (2) are obtained in [52]. Alternate proof of (1) using Dirichlet forms can be found in [26]. The formula (3) is the well-known Pitman sampling formula (cf. [53]). \Box

Let $L^2(\theta, \alpha)$ denote the space of square integrable measurable functions on the probability space $(\nabla_{\infty}, PD(\alpha, \theta))$. Replace 2 with any $p \ge 1$, we can define $L^p(\theta, \alpha)$ and $\|\cdot\|_{L^p(\theta,\alpha)}$ accordingly.

Theorem 5.2. (1) The spectrum of $L_{\theta,\alpha}$ is $\{0, -\lambda_2, -\lambda_3, \ldots\}$. 0 is a simple eigenvalue and for $m \ge 2$, the multiplicity of $-\lambda_m$ is $\pi(m) - \pi(m-1)$, where $\pi(m)$ is the total number of partitions of integer m;

(2) For m = 0, 2, 3..., let W_m be the eigenspace corresponding to eigenvalue $-\lambda_m$ and \oplus denote the direct sums. Then

$$L^2(\theta, \alpha) = \oplus_m W_m;$$

(3) For $k \geq 1$, let $J_k = \{\vec{m} = (m_1, \ldots, m_k) \in \mathbb{N}^k : m_1 \geq m_2 \cdots \geq m_k \geq 2\}$, and $J = \bigcup_{i=1}^{\infty} J_i$. The space $L^2(\alpha, \theta)$ has an orthonormal basis $\{1\} \cup \{\chi_{\mathbf{m}} : \vec{m} \in J\}$ such that $\{\chi_{\mathbf{m}} : \vec{m} \in J, |\vec{m}| = m\}$ is an orthonormal basis for W_m .

Proof. The spectrum in (1) is identified in [52] and the multiplicity is verified in [26]. (2) and (3) can be derived following an argument similar to that used in the proof of Theorem 2.3 in [13]. \Box

Now we are ready to give the main result of this chapter.

5.4 Transition Density Function

We first give the following essential lemma to derive the density function.

Lemma 5.3. For any $\vec{m} = (m_1, \ldots, m_k)$ in J satisfying $|\vec{m}| = m \ge 2$, and any $n \ge m$,

$$\int_{\nabla_{\infty}} P_n(x,\cdot) p_{\vec{m}}(x) PD(\alpha,\theta)(dx) - \frac{n_{[m]}}{(n+\theta)_{(m)}} p_{\vec{m}}(\cdot)$$

is a polynomial of order m-1 or less.

Proof. Fix $\vec{m} = (m_1, \ldots, m_k), |\vec{m}| = m \ge 2$. For each $\ell \in \mathbb{N}, \vec{n} \in I_\ell, \vec{n} = n \ge m$,

$$\frac{\int p_{\vec{m}}^{\circ} p_{\vec{n}}^{\circ} PD(\alpha, \theta)(dx)}{\int p_{\vec{n}}^{\circ} PD(\alpha, \theta)(dx)} = \sum_{\substack{\ell \wedge k \\ j = 0}}^{\ell \wedge k} \prod_{\substack{t=\ell \\ k=\ell}}^{\ell+k-j-1} (\theta + t\alpha) \sum_{\substack{\Lambda \subset \{1,\dots,k\} \\ \sigma:\Lambda \to \{1,\dots,\ell\} i \in \Lambda}} \sum_{\substack{\alpha:\Lambda \to \{1,\dots,\ell\} i \in \Lambda \\ \text{one to one}}} \prod_{i \notin \Lambda} (n_{\sigma(i)} - \alpha)_{(m_i)} \prod_{i \notin \Lambda} (1 - \alpha)_{(m_i-1)} \frac{1}{(n+\theta)_{(m)}}.$$

Since $(n_{\sigma(i)} - \alpha)_{(m_i)} = (-\alpha)(1 - \alpha)_{(m_i-1)} + \sum_{r_i=1}^{m_i} c_{m_i r_i}(\alpha)(n_{\sigma(i)})_{[r_i]}$, we have

$$\prod_{i\in\Lambda} (n_{\sigma(i)} - \alpha)_{(m_i)} = \sum_{\gamma=0}^j \sum_{\Gamma\subset\Lambda\atop |\Gamma|=\gamma} (-\alpha)^{j-\gamma} \prod_{i\in\Lambda\backslash\Gamma} (1-\alpha)_{(m_i-1)} \sum_{\vec{r}\in\Gamma(\vec{m},\Gamma)} \prod_{i\in\Gamma} c_{m_ir_i}(\alpha)(n_{\sigma(i)})_{[r_i]},$$

where

$$\Gamma(\vec{m},\Gamma) = \{\vec{r} = (r_1,\ldots,r_k) : 1 \le r_i \le m_i, i \in \Gamma; r_i = 0, i \notin \Gamma\}.$$

Thus,

$$\frac{\int p_{\vec{m}}^{\circ} PD(\alpha,\theta)(dx)}{\int p_{\vec{n}}^{\circ} PD(\alpha,\theta)(dx)} = \sum_{\substack{\ell \wedge k \\ j=0}}^{\ell \wedge k} \prod_{\substack{t=\ell \\ l = j}}^{\ell \wedge k} (\theta + t\alpha) \sum_{\substack{\Lambda \subset \{1,\dots,k\} \\ |\Lambda| = j}} \sum_{\substack{\sigma:\Lambda \to \{1,\dots,\ell\} \\ \text{one to one}}} \sum_{\substack{\gamma=0 \\ |\Gamma| = \gamma}} \sum_{\substack{\Gamma \subset \Lambda \\ |\Gamma| = \gamma}}^{\Gamma(\Lambda)} (-\alpha)^{j-\gamma} \sum_{i \in \Gamma} \prod_{i \in \Gamma} c_{m_i r_i}(\alpha) (n_{\sigma(i)})_{[r_i]} \prod_{i \notin \Gamma} (1-\alpha)_{(m_i-1)} \frac{1}{(n+\theta)_{(m)}}.$$

By interchanging the order of summation, we have the above

$$= \sum_{j=0}^{\ell \wedge k} \prod_{t=\ell}^{\ell+k-j-1} (\theta+t\alpha) \sum_{\substack{\Lambda \subset \{1,\dots,k\} \\ |\Lambda|=j}} \sum_{\gamma=0}^{j} \sum_{\substack{\Gamma \subset \Lambda \\ |\Gamma|=\gamma}} (-\alpha)^{j-\gamma} \prod_{i \notin \Gamma} (1-\alpha)_{(m_i-1)} \sum_{\substack{\sigma:\Gamma \to \{1,\dots,\ell\} \\ \text{one to one}}} \frac{(\ell-\gamma)!}{(j-\gamma)!} \\ \sum_{\substack{r \in \Gamma(\vec{m},\Gamma) \\ i \in \Gamma}} \prod_{i \in \Gamma} c_{m_i r_i}(\alpha) (n_{\sigma(i)})_{[r_i]} \cdot \frac{1}{(n+\theta)_{(m)}} \\ = \sum_{j=0}^{\ell \wedge k} \prod_{t=\ell}^{\ell+k-j-1} (\theta+t\alpha) \sum_{\substack{\gamma=0 \\ \gamma=0}}^{j} \sum_{\substack{\Gamma \subset \{1,\dots,k\} \\ |\Gamma|=\gamma}} \binom{k-\gamma}{j-\gamma} (-\alpha)^{j-\gamma} \frac{(\ell-\gamma)!}{(j-\gamma)!} \prod_{i \notin \Gamma} (1-\alpha)_{(m_i-1)} \\ \sum_{\substack{\sigma:\Gamma \to \{1,\dots,\ell\} \\ i \in \Gamma}} \sum_{\substack{\tau \in \Gamma(\vec{m},\Gamma) \\ i \in \Gamma}} \prod_{i \in \Gamma} c_{m_i r_i}(\alpha) (n_{\sigma(i)})_{[r_i]} \cdot \frac{1}{(n+\theta)_{(m)}} \\ = \sum_{\gamma=0}^{\ell \wedge k} \left(\sum_{j=\gamma}^{\ell \wedge k} \prod_{t=\ell}^{\ell+k-j-1} (\theta+t\alpha) \binom{k-\gamma}{j-\gamma} (-\alpha)^{j-\gamma} \frac{(\ell-\gamma)!}{(j-\gamma)!} \right) \sum_{\substack{\Gamma \subset \{1,\dots,k\} \\ |\Gamma|=\gamma}} \prod_{i \notin \Gamma} (1-\alpha)_{(m_i-1)} \\ \sum_{\substack{\sigma:\Gamma \to \{1,\dots,\ell\} \\ one to one}} \sum_{\tau \in \Gamma(\vec{m},\Gamma)} \prod_{i \in \Gamma} c_{m_i r_i}(\alpha) (n_{\sigma(i)})_{[r_i]} \cdot \frac{1}{(n+\theta)_{(m)}}. \end{cases}$$

Also, we have

$$\begin{split} &\sum_{j=\gamma}^{\ell \wedge k} \prod_{t=\ell}^{\ell+k-j-1} (\theta + t\alpha) \binom{k-\gamma}{j-\gamma} (-\alpha)^{j-\gamma} \frac{(\ell-\gamma)!}{(j-\gamma)!} \\ &= \sum_{j=\gamma}^{\ell \wedge k} \frac{(\theta + \ell\alpha) \cdots (\theta + (\ell+k-j-1)\alpha)}{(k-j)!} \binom{\ell-\gamma}{j-\gamma} (-\alpha)^{j-\gamma} (k-\gamma)! \\ &= \sum_{j=\gamma}^{\ell \wedge k} \frac{(\frac{\theta}{\alpha} + \ell) \cdots (\frac{\theta}{\alpha} + \ell + k - j - 1)}{(k-j)!} \binom{\ell-\gamma}{j-\gamma} (-1)^{j-\gamma} \alpha^{k-\gamma} (k-\gamma)! \\ &= \alpha^{k-\gamma} (k-\gamma)! \sum_{j=0}^{(\ell-\gamma) \wedge (k-\gamma)} \binom{\frac{\theta}{\alpha} + \ell + k - j - \gamma - 1}{k-j-\gamma} \binom{\ell-\gamma}{j} (-1)^{j} \\ &= \alpha^{k-\gamma} (k-\gamma)! \binom{\frac{\theta}{\alpha} + k - 1}{k-\gamma} \\ &= (\theta + \gamma\alpha) \cdots (\theta + (k-1)\alpha), \end{split}$$

where the last but one equality will be verified at the end of proof. Therefore,

$$\frac{\int p_{\vec{m}}^{\circ} p_{\vec{n}}^{\circ} dPD(\alpha, \theta)}{\int p_{\vec{n}}^{\circ} dPD(\alpha, \theta)} = \sum_{\substack{\ell \wedge k \\ \gamma = 0}}^{\ell \wedge k} (\theta + \gamma \alpha) \cdots (\theta + (k - 1)\alpha) \sum_{\substack{\Gamma \subset \{1, \dots, k\} \\ |\Gamma| = \gamma}} \prod_{i \notin \Gamma} (1 - \alpha)_{(m_i - 1)} \sum_{\substack{\sigma: \Gamma \to \{1, \dots, \ell\} \\ \text{one to one}}} \sum_{\vec{r} \in \Gamma(\vec{m}, \Gamma)} \prod_{\substack{i \in \Gamma}} c_{m_i r_i}(\alpha) (n_{\sigma(i)})_{[r_i]} \frac{1}{(n + \theta)_{(m)}}.$$

Thus, we have

$$\begin{split} &\int P_n(x,y)p_{\vec{m}}(x)\,dPD(\alpha,\theta) \\ &= \sum_{\ell=1}\sum_{\substack{\vec{n}\in I_\ell\\|\vec{n}|=n}} \binom{m}{m_1\cdots m_k} \frac{1}{\beta_1!\cdots \beta_m!} \frac{\int p_{\vec{m}}^{\circ} p_{\vec{n}}^{\circ} dPD(\alpha,\theta)}{\int p_{\vec{n}}^{\circ} dPD(\alpha,\theta)} p_{\vec{n}}(y) \\ &= \sum_{\ell=1}\sum_{\substack{\vec{n}\in I_\ell\\|\vec{n}|=n}} \binom{m}{m_1\cdots m_k} \frac{1}{\beta_1!\cdots \beta_m!} \sum_{\gamma=0}^{\ell\wedge k} (\theta+\gamma\alpha)\cdots(\theta+(k-1)\alpha) \sum_{\substack{\Gamma\subset\{1,\dots,k\}\\|\Gamma|=\gamma}} \prod_{\substack{\vec{n}\in I_\ell\\|\vec{n}|=n}} (1-\alpha)_{(m_i-1)} \sum_{\substack{\sigma:\Gamma\to\{1,\dots,\ell\}\\\sigma\in\sigma\text{ to one}}} \sum_{\vec{n}\in\Gamma} \prod_{\vec{n}\in\Gamma} c_{m_ir_i}(\alpha)(n_{\sigma(i)})_{[r_i]} \frac{p_{\vec{n}}(y)}{(n+\theta)_{(m)}} \\ &= \frac{n_{[m]}}{(n+\theta)_{(m)}} \binom{m}{m_1\cdots m_k} \frac{1}{\beta_1!\cdots \beta_m!} \sum_{\gamma=0}^k (\theta+\gamma\alpha)\cdots(\theta+(k-1)\alpha) \\ &\sum_{\substack{\Gamma\subset\{1,\dots,k\}\\|\vec{\Gamma}|=\gamma}} \prod_{i\notin\Gamma} (1-\alpha)_{(m_i-1)} \sum_{\vec{r}\in\Gamma(\vec{m},\Gamma)} \frac{n_{[r]}}{n_{[m]}} \prod_{i\in\Gamma} c_{m_ir_i}(\alpha) \\ &\cdot \{\sum_{\ell=\gamma\vee 1}^n \sum_{\substack{\vec{n}\in I_\ell\\|\vec{\sigma}|=n}} \sum_{\substack{\sigma:\Gamma\to\{1,\dots,\ell\}\\\sigma\in\tau\text{ to one}}} \frac{\prod_{i\in\Gamma} (n_{\sigma(i)})_{[r_i]}}{n_{[r]}} p_{\vec{n}}(y)\}, \end{split}$$

where the expression in the braces is exactly $p_{\vec{r}}^{\circ}(y)$. To see this, we first notice that for $\vec{r} = (r_1, \ldots, r_{\gamma})$, $p_{\vec{r}}^{\circ}(y)$ is the probability of a random sampling of size $r = |\vec{r}|$ from a population with allele frequencies y_1, y_2, \ldots with γ types and each type contains r_i individuals for $1 \leq i \leq \gamma$. The probability of this event can be calculated in an alternative way. Suppose we select a random sample of size n with ℓ types ($\ell \geq \gamma$) and each type has n_i alleles for $1 \leq i \leq \ell$. The probability of getting such a sample is $p_{\vec{n}}(y)$. Then we choose γ types denoted by σ out of ℓ types from this sample. Since for each type we only need to select r_i individuals. the corresponding probability would be $\frac{\prod_{i \in \Gamma} (n_{\sigma(i)})_{[r_i]}}{n_{[r]}}$. Equating the probability of there two procedures we get

$$p_{\vec{r}}^{\circ}(y) = \sum_{\ell=\gamma \vee 1}^{n} \sum_{\substack{\vec{n} \in I_{\ell} \\ |\vec{n}|=n}} \sum_{\substack{\sigma: \Gamma \to \{1, \dots, \ell\} \\ \text{ one to one}}} \frac{\prod_{i \in \Gamma} (n_{\sigma(i)})_{[r_i]}}{n_{[r]}} p_{\vec{n}}(y).$$

In the following we show the proof of the combinatorics identity. To prove

$$\sum_{j=0}^{(\ell-\gamma)\wedge(k-\gamma)} \binom{\frac{\theta}{\alpha}+\ell+k-j-\gamma-1}{k-j-\gamma} \binom{\ell-\gamma}{j} (-1)^j = \binom{\frac{\theta}{\alpha}+k-1}{k-\gamma},$$

we only need to show that

$$\sum_{j=0}^{\ell \wedge (k-\gamma)} \binom{\frac{\theta}{\alpha} + k + \ell - j - 1}{k - \gamma - j} \binom{\ell}{j} (-1)^j = \binom{\frac{\theta}{\alpha} + k - 1}{k - \gamma}, \tag{5.5}$$

where we put ℓ instead of $\ell - \gamma$. Consider the Taylor expansion of $(1 - x)^{-(\frac{\theta}{\alpha} + \gamma)}$, we have

$$(1-x)^{-(\frac{\theta}{\alpha}+\gamma)} = \sum_{i=0}^{\infty} {\binom{\frac{\theta}{\alpha}+\gamma+i-1}{i}} x^i,$$

since $(1-x)^{-n} = \sum_{i=0}^{\infty} {\binom{-n}{i}} (-x)^i = \sum_{i=0}^{\infty} {\binom{n+i-1}{i}} x^i$. The coefficient of $x^{k-\gamma}$ in the expansion is $\left(\frac{\frac{\theta}{\alpha}+\gamma+k-\gamma-1}{k-\gamma}\right) = {\binom{\theta}{\alpha}+k-1}{k-\gamma}$ which in the right side of (1). In addition, since $(1-x)^{-(\frac{\theta}{\alpha}+\gamma)} = (1-x)^{-(\frac{\theta}{\alpha}+\ell+\gamma)} \cdot (1-x)^{\ell}$, and

$$(1-x)^{-\left(\frac{\theta}{\alpha}+\ell+\gamma\right)} = \sum_{i=0}^{\infty} \left(\frac{\frac{\theta}{\alpha}+\ell+\gamma+i-1}{i}\right) x^{i},$$
$$(1-x)^{\ell} = \sum_{i=0}^{\ell} \binom{\ell}{i} (-1)^{i} x^{i},$$

we have

$$(1-x)^{-(\frac{\theta}{\alpha}+\ell+\gamma)} \cdot (1-x)^{\ell} = \sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\ell} \binom{\frac{\theta}{\alpha}+\ell+\gamma+i_1-1}{i_1} \binom{\ell}{i_2} (-1)_2^i x^{i_1+i_2}.$$

Thus, the coefficient of $x^{k-\gamma}$ is

$$\sum_{i_1+i_2=k-\gamma} \binom{\frac{\theta}{\alpha}+\ell+\gamma+i_1-1}{i_1} \binom{\ell}{i_2} (-1)^{i_2} = \sum_{i_2=0}^{\ell \wedge (k-\gamma)} \binom{\frac{\theta}{\alpha}+k+\ell-i_2-1}{k-\gamma-i_2} \binom{\ell}{i_2} (-1)^{i_2},$$

which is the left side of 5.5. Equating the coefficients in these two ways, we got the identity. $\hfill \Box$

Let T(t) denote the semigroup defined on $C(\overline{\nabla}_{\infty})$ generated by $L_{\theta,\alpha}$, and $\{1, \chi_{\vec{m}} : \vec{m} \in J\}$ be the orthonormal basis of $L^2(\theta, \alpha)$ in Theorem 5.2. Then for f in $C(\overline{\nabla}_{\infty})$, the following equality holds

$$T(t)f(x) = (f,1) + \sum_{m=2}^{\infty} e^{-\lambda_m t} \sum_{\vec{m} \in J: |\vec{m}|=m} (f,\chi_{\vec{m}})\chi_{\vec{m}}(x),$$

where (\cdot, \cdot) denotes the scalar product in $L^2(\theta, \alpha)$. Thus, if

$$\sum_{m=2}^{\infty} e^{-\lambda_m t} \sum_{\vec{m} \in J: |\vec{m}|=m} \chi_{\vec{m}}^2 \in L^1(\theta, \alpha),$$
(5.6)

then T(t) has a transition density

$$p(t, x, y) = 1 + \sum_{m=2}^{\infty} e^{-\lambda_m t} q_m(x, y)$$

with

$$q_m(x,y) = \sum_{\vec{m} \in J: |\vec{m}|=m} \chi_{\vec{m}}(x)\chi_{\vec{m}}(y), \quad x,y \in \nabla.$$

As in [13], we can show that for any $m \ge 1$,

$$\oplus_{n=0}^{m} W_n = span\{p_{\vec{n}} : \vec{n} \in \{0\} \cup J, |\vec{n}| \le m\},\$$

and Lemma 5.3 can be applied to the orthonormal basis $\{\chi_{\vec{m}} : \vec{m} \in J\}$, and for any $\vec{m}, \vec{m}' \in 0 \cup J$ with $m = |\vec{m}| \leq |\vec{m}'| \leq n$,

$$\int_{\overline{\nabla}_{\infty}} \int_{\overline{\nabla}_{\infty}} P_n(x, y) \chi_{\vec{m}}(x) \chi_{\vec{m}'}(y) PD(\alpha, \theta)(dx) PD(\alpha, \theta)(dy) = \frac{n_{[m]}}{(n+\theta)_m} \delta_{\vec{m}\vec{m}'}.$$
 (5.7)

Write $p_n(x, y)$ as a linear combination of $\{\chi_{\vec{m}}(x)\chi_{\vec{m}'}(y) : |\vec{m}| \lor |\vec{m}'| \le n\}$. It follows from 5.7 that the coefficient of $\chi_{\vec{m}}(x)\chi_{\vec{m}'}(y)$ is zero if $\vec{m} \ne \vec{m}'$. Therefore, we get a linear expression of $p_n(x, y)$ in terms of $\{q_m(x, y) : m \le n\}$.

$$P_n(x,y) = 1 + \sum_{m=2}^n \frac{n_{[m]}}{(n+\theta)_{(m)}} \sum_{\vec{m} \in J : |\vec{m}|=m} \chi_{\vec{m}}(x) \chi_{\vec{m}}(y)$$
$$= 1 + \sum_{m=2}^n \frac{n_{[m]}}{(n+\theta)_{(m)}} q_m(x,y).$$

Solving $q_m(x, y)$ in terms of $P_n(x, y)$, yields that for any $m \ge 2$,

$$q_m(x,y) = \frac{2m-1+\theta}{m!} \sum_{n=0}^m (-1)^{m-n} \binom{m}{n} (n+\theta)_{(m-1)} P_n(x,y)$$

where $P_0(x,y) = 1$. Since $\int_{\nabla_{\infty}} p_{\vec{n}} dPD(\alpha,\theta) = \frac{n!}{n_1! \cdots n_\ell! \alpha_1! \cdots \alpha_n!} \frac{\prod_{i=0}^{\ell-1} (\theta+i\alpha)}{\theta_{(n)}} \prod_{i=1}^{\ell} (1-\alpha)_{n_i-1}$, it follows that

$$||P_n||_{\infty} \le \max_{\vec{n} \in I: |\vec{m}|=n} (\int_{\nabla_{\infty}} p_{\vec{n}} dPD(\alpha, \theta))^{-1} \le \frac{\theta_{(n)}}{(\theta(1-\alpha))^n \wedge 1}$$

Then, for $m \ge 2$ there exist C, d > 1 such that

$$||q_m||_{\infty} \le (2m - 1 + \theta) \sum_{n=0}^{m} (n + \theta)_{m-1} \frac{\theta_{(n)}}{(\theta(1 - \alpha))^n \wedge 1} \le (cm^d)^m.$$
(5.8)

which leads to (5.6) and the following upper bound for the transition density function.

Theorem 5.4. The transition function of the process $X_{\theta,\alpha}(t)$ has a density function p(t, x, y) with respect to $PD(\alpha, \theta)$ given by

$$p(t, x, y) = 1 + \sum_{m=2}^{\infty} e^{-\lambda_m t} q_m(x, y),$$
 (5.9)

and there exists a constant c > 1 such that

$$p(t, x, y) \le ct^{c(\log t)/t}, \quad t > 0, x, y \in \overline{\nabla}_{\infty}.$$
 (5.10)

Proof. Obviously, (5.6) follows from (5.8) and that $\lambda_m = m(m-1+\theta)/2$. So, it suffices to prove (5.10). Since $\sup_{x,y} p(t,x,y)$ is decreasing in t, we only have to consider the case that $t \in (0, 1/2]$. By (5.8), (5.9) and the fact that $\lambda_m = \frac{1}{2}m(m-1+\theta)$, we have

$$p(t, x, y) \le 1 + \sum_{m=2}^{\infty} 2^{-m} \exp[\Psi_t(m)],$$
 (5.11)

where

$$\Psi_t(m) = -\frac{1}{2}m(m-1+\theta)t + m\log(2C) + dm\log m, \quad m \ge 2.$$

Noting that

$$d\log m \le \frac{mt}{4} + d\log \frac{4d}{t} - d,$$

we conclude that

$$\Psi_t(m) \le -\frac{m^2 t}{4} + c_1 m \log t^{-1} \le \frac{c_2 (\log t^{-1})^2}{t}, \ t \in (0, 1/2]$$

holds for some constants $c_1, c_2 > 0$ and all m > 0. The proof is then completed by combining this with (5.11).

5.5 Random Sampling at Time t

Using the above transition density, we find the transient sampling distribution for the two-parameter diffusion model. Our result can be seen as a two-parameter generalization of Theorem 1 in [31]. There the author's approach is based on the fact that the IMNA diffusion process can be approximated by the finite alleles Wright-Fisher model. However, a finite-dimensional approximation of the two-parameter model is not found yet. Therefore, we turn to use the correlation measure of the two-parameter Poisson-Dirichlet point process obtained in Chapter 3 in our derivation.

Theorem 5.5. Consider the two-parameter diffusion process X(t) with generator $L_{\alpha,\theta}$ (5.1) with initial value $x = (x_1, x_2, ...)$. The probability generating function

$$\mathbf{E}[\prod_{i=1}^{n} u_i^{a_i} | X(0) = x]$$

of a sample $\mathbf{a} = (a_1, a_2, \dots, a_r)$ of size r at time t is the coefficient of φ^r in

$$G_0 + \sum_{m=2}^{\infty} \frac{2m - 1 + \theta}{m!} e^{-\lambda_m t} \sum_{n=0}^m (-1)^{m-n} \binom{m}{n} (n+\theta)_{(m-1)} G_n,$$

where

$$G_0 = G_1 = \frac{r!}{\theta_{(r)}} \left(1 - \alpha \sum_{l=1}^r (1 - \alpha)_{(l-1)} u_l \varphi^l / l! \right)^{-\theta/\alpha},$$

and for $m \geq 2$

$$G_{m} = r! \frac{\theta_{n}}{\theta_{(n+r)}} \sum_{\boldsymbol{n}=(n_{1},\dots,n_{k})} \psi_{\boldsymbol{n}}(x) \left(1 - \alpha \sum_{l=1}^{r} (1-\alpha)_{(l-1)} u_{l} \varphi^{l} / l!\right)^{(-k+\frac{\theta}{\alpha})}$$
$$\prod_{i=1}^{k} \left(1 + \sum_{l=1}^{r} (n_{i} - \alpha)_{(l)} \varphi^{l} u_{l} / l!\right)$$

Proof. Since the probability generating function is $r! \times$ the coefficient of φ^r in

$$\mathbf{E}[\prod_{i=1}^{\infty} (1 + \sum_{l=1}^{r} (\varphi X_{i}(t))^{l} u_{l}/l!) | X(0) = x]$$

Due to the expression of transition density q(t, x, y) we only need to compute

$$E[\prod_{i=1}^{\infty} (1 + \sum_{l=1}^{r} (\varphi X_i)^l u_l / l!)]$$
(5.12)

and

$$PSF(\mathbf{n})^{-1} \mathbf{E}[\psi_{\mathbf{n}}(X) \prod_{i=1}^{\infty} (1 + \sum_{l=1}^{r} (\varphi X_{i})^{l} u_{l}/l!)], \qquad (5.13)$$

where **n** = $(n_1, ..., n_k)$,

$$PSF(\mathbf{n}) = \frac{\theta(\theta + \alpha) \cdots (\theta + (k-1)\alpha)}{\theta_{(n)}} (1-\alpha)_{(n_1-1)} \cdots (1-\alpha)_{(n_k-1)}$$

and the expectation is taken with respect to the $PD(\alpha, \theta)$ distribution. For simplicity we only give the proof of (5.12). (5.13) can be computed similarly using the argument in Theorem 3.1. Setting $g(z) = \sum_{l=1}^{n} \varphi^l z^l u_l / l!$ and using the correlation function obtained in Theorem 3.1, we have

$$\begin{split} \mathbf{E}[\prod_{i=1}^{\infty} (1 + \sum_{l=1}^{r} (\varphi X_{i}(t))^{l} u_{l}/l!)] \\ &= 1 + \sum_{n=1}^{\infty} \frac{c_{n,\alpha,\theta}}{n!} \int_{\Delta_{n}} \prod_{i=1}^{n} \frac{g(v_{i})}{v_{i}^{1+\alpha}} (1 - \sum_{j=1}^{n} v_{j})^{\theta+\alpha n-1} dv_{1} \cdots dv_{n} \\ &= 1 + \sum_{n=1}^{\infty} \frac{c_{n,\alpha,\theta}}{n!} \sum_{l_{1},\dots,l_{n}=1}^{r} \frac{\varphi^{l_{1}+\dots+l_{n}} u_{l_{1}} \cdots u_{l_{n}}}{l_{1}! \cdots l_{n}!} \frac{\Gamma(l_{1}-\alpha) \cdots \Gamma(l_{n}-\alpha) \Gamma(\theta+\alpha n)}{\Gamma(l_{1}+\dots+l_{n}+\theta)} \\ &= 1 + \sum_{n=1}^{\infty} \frac{\Gamma(\theta)\alpha^{n}}{n!} \left(\frac{\theta}{\alpha}\right) \sum_{(n)} \sum_{l_{1},\dots,l_{n}=1}^{r} \frac{\varphi^{l_{1}+\dots+l_{n}} u_{l_{1}} \cdots u_{l_{n}} (1-\alpha)_{(l_{1}-1)} \cdots (1-\alpha)_{(l_{n}-1)}}{l_{1}! \cdots l_{n}! \Gamma(l_{1}+\dots+l_{n}+\theta)}. \end{split}$$

Therefore, the coefficient of φ^r in the above corresponds to the coefficient of φ^r in the Taylor expansion of $G_0/\theta_{(r)}$. It is also worth noticing that by letting $\alpha \to 0$, the limit of G_0 is $r!\theta_{(r)}^{-1}\exp\{\theta\sum_{l=1}^r u_l\varphi^l/l\}$, which coincides with the result obtained by Griffiths in [31]. This is also true for $G_m, m \ge 2$.

Chapter 6

The Quasi-invariant Property and Its Application

6.1 Abstract

Suppose that S is a Borel space with probability measure μ and G is a group consisting of measurable transformations on S. The quasi-invariance of μ with respect to the transformation group G means that for any element $T \in G$, the image measure $T\mu$ and the original measure μ are mutually absolutely continuous. We obtain two results based on the quasi-invariance property of the gamma process a discrete random measure on S under the multiplication transformation group. Under the induced transformation on the normalized atom masses, we establish the quasi-invariance property of the two-parameter Poisson-Dirichlet distribution. On the other hand, we show the equivalence between the reversibility of a class of measure-valued branching process and the quasi-invariance property of their stationary distributions with respect to the multiplication group. In the reversible case, the cocycle of the quasiinvariance is given by the integral of the drift term along a specific transformation flow.

6.2 Introduction

A subordinator $(x_t, t \ge 0)$ is an increasing process with stationary independent increments. Assume that (x_t) has no drift component, the corresponding Laplace transform is

$$E[\exp(-\lambda x_t)] = \exp(-t \int_0^\infty (1 - \exp(-\lambda z))\Lambda(dz)),$$

where the Lévy measure Λ on $(0, \infty)$ is the intensity measure for the Poisson point process of jumps $(x_s - x_{s-}, s \ge 0)$. The gamma process denoted by $(\gamma_s, s \ge 0)$ is a subordinator with Lévy measure given by

$$\Lambda(dx) = x^{-1} e^{-x} dx, x > 0.$$
(6.1)

Therefore, the Laplace transform of the gamma process is

$$E[\exp\{-\lambda\gamma_s\}] = \exp\{-s\int_0^\infty (1 - \exp(-\lambda x))x^{-1}e^{-x}dx\} = \frac{1}{(1+\lambda)^s},$$

indicating that the marginal distribution of γ_s is the gamma(s) distribution, i.e. for $s \ge 0$

$$\mathsf{P}(\gamma_s \in dx) = \Gamma(s)^{-1} x^{s-1} e^{-x} dx, x > 0,$$

where $\Gamma(s) = \int_0^\infty \lambda^{s-1} e^{-\lambda} d\lambda$.

Another typical subordinator is the α -stable process τ_s with Lévy measure

$$\Lambda_{\alpha}(x,\infty) = Cx^{-\alpha}, x > 0$$

for some constant C > 0, and thus the Laplace transform is

$$\mathbf{E}[\exp\{-\lambda\tau_s\}] = \exp\{-sK\lambda^{\alpha}\}, \quad K = C\Gamma(1-\alpha).$$

In fact, the above pure jump subordinators can be realized as a random discrete measure

$$\eta = \sum Z_i \delta_{S_i},$$

where the sizes of jumps $(Z_1, Z_2, ...)$ are distinct points of a Poisson point process on $[0, \infty)$ with mean measure given by the Lévy measure of subordinator. Independently, the jumping locations $(S_1, S_2, ...)$ forms a sequence of i.i.d. random variables on a standard Borel space S with a non-atomic probability measure ν . As a random discrete measure on the space (S, ν) the Laplace transform of η is

$$\mathbf{E}[\exp\{-\langle a,\eta\rangle\}] = \exp\{\langle \log\psi_{\Lambda}(a),\nu\rangle\},\$$

where a is an arbitrary non-negative bounded Borel function on S and

$$\psi_{\Lambda}(a) = \exp\{-\int_0^\infty (1 - e^{-xa})d\Lambda(x)\}.$$

Here and after $\langle a, \eta \rangle$ always represents $\int_{S} a(s)\eta(ds)$.

Consider the gamma process $\gamma(\theta)$, which is equivalent to $\eta = \sum Z_i \delta_{S_i}$ with the Lévy measure $\Lambda(dx) = \theta x^{-1} e^{-x} dx, x > 0$. Its Laplace transform is given by

$$\mathbf{E}[\exp\{-\langle a,\eta\rangle\}] = \exp\{-\theta\langle \log(1+a),\nu\rangle\}.$$

Denote the law of the gamma process η by P_{ν} .

It follows that the law P_{ν} is concentrated on the cone

$$D = \{\sum z_i \delta_{s_i} : z_i > 0, s_i \in S, \sum z_i < \infty\}$$

consisting of finite positive discrete measures on S with infinite number of atoms and equipped with the usual weak convergence topology. We decompose $\eta = \sum Z_i \delta_{S_i}$ into two components as

$$\eta = (\bar{\eta}, \eta(S)),$$

where the first coordinate $\bar{\eta} = \sum \frac{Z_i}{\sum Z_j} \delta_{S_i} = \sum Y_i \delta_{S_i}$ is the probability measure after

normalization, and the second denotes the total mass $\sum Z_i$. Furthermore, the normalized random measure $\bar{\eta} = \sum \frac{Z_i}{\sum Z_j} \delta_{S_i} = \sum Y_i \delta_{S_i}$ called the Dirichlet process is independent of the total mass of $\eta(S) = \sum Z_j = \gamma(\theta)$. Denote the distribution of $\bar{\eta}$ as $\Pi_{\theta,\nu}$. The Dirichlet process has extensive applications in the Bayesian statistics (see e.g. [29]).

Consider the following transformation on $\mathcal{M}(S)$ i.e. the space of finite positive Borel measures on S with the topology of weak convergence:

$$M_a\eta(ds) = a(s)\eta(ds),$$

where a belongs to the space of nonnegative bounded measurable functions on S with $\inf_{s} a(s) > 0$. The induced transformation on $\mathcal{P}(S)$ the set of probability measure on S with weak convergence topology is given by

$$S_a\bar{\eta}(ds) = a(s)\eta(ds)/\langle a,\eta\rangle.$$

Note that both M_a and S_a form a transformation group on $\mathcal{M}(S)$ and $\mathcal{P}(S)$:

$$M_{a+b}\eta = M_a(M_b\eta),$$

and

 $S_{a+b}\bar{\eta} = S_a(S_b\bar{\eta}).$

In particular, $(M_a)^{-1} = M_{a^{-1}}$ and $(S_a)^{-1} = S_{a^{-1}}$. [62] and [63] show that for the gamma process the image measure $M_a P_{\nu}$ and P_{ν} are mutually absolutely continuous and the density is given by

$$\frac{M_a P_\nu}{P_\nu}(d\eta) = \exp\{-\langle \log a, \nu \rangle - \langle a^{-1} - 1, \eta \rangle\}, \quad P_\nu - a.s..$$

For the Dirichlet process $\Pi_{\theta,\nu}$, the quasi-invariance also holds with the density given by [35] as

$$\frac{S_a \Pi_{\theta,\nu}}{\Pi_{\theta,\nu}} (d\bar{\eta}) = \exp\{-\theta \langle \log a, \bar{\nu} \rangle + \log \langle a^{-1}, \bar{\eta} \rangle\}, \quad \Pi_{\theta,\nu} - a.s..$$
(6.2)

Denote the cone

$$C = \{x = (y_1, y_2, \ldots) : y_1 \ge y_2 \ge \cdots \ge 0, \sum y_i < \infty\} \subseteq \ell^1.$$

We define a map $T: D \to C$ by arranging the atom masses in decreasing order

$$T\eta = (z_1, z_2, \ldots), \quad \text{if} \quad \eta = \sum z_i \delta_{s_i},$$

and

$$T\bar{\eta} = (z_1 / \sum z_i, z_2 / \sum z_i, \ldots) = (Y_1, Y_2, \ldots).$$

When η is the gamma process, the corresponding $T\bar{\eta}$ has the law of the Poisson-Dirichlet distribution with parameter θ . In the same manner, the $PD(\alpha, 0)$ distribution can be derived as the law of $T\bar{\eta}$ when η is the α -stable subordinator.

In [62] the authors use this property to show the quasi-invariance of the Poisson-Dirichlet distributions denoted by $PD(\theta)$. The induced mapping on the ordered atom masses $(Y_1, Y_2, ...)$ of the Dirichlet process is defined as

$$S_a Y = \Phi(a(S_1)Y_1/\rho, a(S_2)Y_2/\rho, \ldots),$$

where $\rho = \sum a(S_i)Y_i$ and Φ is a map that arranges the coordinates in non-increasing order. It is easy to see that S_a also forms a transformation group on the ordered infinite-dimensional simplex space

$$\nabla_{\infty} = \{(y_1, y_2, \ldots) : y_1 \ge y_2 \ge \cdots \ge 0, \sum y_i = 1\}$$

with topology inherited from $[0, 1]^{\infty}$.

Under the mapping S_a , it is shown that the image distribution $S_a PD(\theta)$ is equivalent to the $PD(\theta)$ distribution. Furthermore, Arguin [2] utilizes this transformation to characterize the $PD(\alpha, 0)$ distribution. He shows that the $PD(\alpha, 0)$ distribution is invariant under S_a with proper multiplicators a. Here, we further show that the quasi-invariance of the two-parameter Poisson-Dirichlet distribution $PD(\alpha, \theta)$ still holds with respect to the transformation group S_a parametrized by bounded Borel measurable functions a on S with $\inf_s a(s) > 0$ and a sequence of i.i.d. ν_0 distributed random variables S_1, S_2, \ldots on S.

6.3 Quasi-invariance of the $PD(\alpha, \theta)$ distribution

First recall the definition of the two-parameter Poisson-Dirichlet distribution. For $0 \leq \alpha < 1$ and $\theta > -\alpha$, let $U_k, k = 1, 2, ...$ be a sequence of independent random variables such that U_k has $Beta(1 - \alpha, \theta + k\alpha)$ distribution with density function given by

$$f(x) = \frac{\Gamma(\theta + 1 + (k - 1)\alpha)}{\Gamma(1 - \alpha)\Gamma(\theta + k\alpha)} x^{-\alpha} (1 - x)^{\theta + k\alpha - 1}, \ 0 < x < 1.$$

Set $(V_1, V_2, ...)$ as

$$V_1 = U_1, V_n = (1 - U_1) \cdots (1 - U_{n-1})U_n, \ n \ge 2.$$
(6.3)

Arrange $(V_1, V_2, ...)$ in descending order and denote it as $P = (P_1, P_2, ...)$. Then the law of P is called the two-parameter Poisson-Dirichlet distribution denoted by $PD(\alpha, \theta)$. The law of $(V_1, V_2, ...)$ is called the two-parameter GEM distribution denoted by $GEM(\alpha, \theta)$. $PD(\alpha, \theta)$ is a probability concentrating on ∇_{∞} . $PD(0, \theta)$ corresponds to the Poisson-Dirichlet distribution with parameter $PD(\theta)$.

The α -stable process on the space (S, ν_0) with $\nu_0(S) = 1$ is a subordinator with Lévy measure

$$d\Lambda_{\alpha} = \frac{c\alpha}{\Gamma(1-\alpha)} x^{-\alpha-1} dx,$$

where $x > 0, 0 < \alpha < 1$ and c > 0. Denote the law of the α -stable process by $P_{\alpha,c}^{\nu_0}$. The Laplace transform of $P_{\alpha,c}^{\nu_0}$ is

$$\mathbf{E}_{P^{\nu_0}_{\alpha,c}}[\exp\{-\langle b,\tau\rangle\}] = \exp\{-c\langle b^{\alpha},\nu_0\rangle\}$$

for an arbitrary nonnegative measurable function b on S with $\langle b^{\alpha}, \nu_0 \rangle < \infty$. Consider $(M_a \tau)(ds) = a(x)\tau(ds)$ such that a is a bounded nonnegative Borel measurable functions on S with $\inf_s a(s) > 0$. Then the Laplace transform of $M_a P_{\alpha,c}^{\nu_0}$ is

$$E_{M_{a}P_{\alpha,c}^{\nu_{0}}}[\exp\{-\langle b,\tau\rangle\}] = E_{P_{\alpha,c}^{\nu_{0}}}[\exp\{-\langle ab,\tau\rangle\}]$$

$$= \exp\{-c\langle a^{\alpha}b^{\alpha},\nu_{0}\rangle\}$$

$$= \exp\{-cA_{\alpha}\langle b^{\alpha},\frac{a^{\alpha}\nu_{0}}{A_{\alpha}}\rangle\}$$

$$= E_{P_{\alpha,c}^{\nu_{0}^{\alpha}}}[\exp\{-\langle b,\tau\rangle\}],$$

where $A_{\alpha} = \langle a^{\alpha}, \nu_0 \rangle$ and $\nu_0^{\alpha}(ds) = a^{\alpha}(s)\nu_0(ds)/A_{\alpha}$. Therefore, we have the following equivalence of the distributions

$$M_a P^{\nu_0}_{\alpha,c} = P^{\nu_0^{\alpha}}_{\alpha,cA_{\alpha}}.$$
(6.4)

Pitman and Yor [57] show that the two-parameter Poisson-Dirichlet distribution $PD(\alpha, \theta)$ with $\theta \neq 0$ is absolutely continuous with respect to the $PD(\alpha, 0)$ distribution. Namely, consider the law $P_{\alpha,\theta}^{c,\nu_0}$ which has the density

$$\frac{P^{c,\nu_0}_{\alpha,\theta}}{P^{\nu_0}_{\alpha,c}}(d\eta) = \frac{c_{\alpha,\theta}}{\eta(S)^{\theta}}$$

with respect to the α -stable law $P_{\alpha,c}^{\nu_0}$. Here the normalizing constant $c_{\alpha,\theta} = c^{\theta/\alpha} \frac{\Gamma(\theta+1)}{\Gamma(\theta/\alpha+1)}$. The simplicial part $T\bar{\eta}$ where η distributed as $(P_{\alpha,\theta}^{c,\nu_0})$ has the two-parameter Poisson-Dirichlet distribution $PD(\alpha,\theta)$. This is essential obtaining the main result.

Theorem 6.1. Suppose that $\{S_i, i = 1, 2, ...\}$ is a sequence of i.i.d. ν_0 distributed random variables on a Borel space S and a is is a bounded nonnegative Borel measurable functions on S with $\inf_s a(s) > 0$. Independently, set $P = (P_i)_{i=1}^{\infty} \sim PD(\alpha, \theta)$ and $S_a P = \Phi(a(S_1)P_1/\rho, a(S_2)P_2/\rho, ...)$, where $\rho = \sum a(S_i)P_i$, Φ is a map that arranges the coordinates in non-increasing order. Then the two-parameter Poisson-Dirichlet distribution $PD(\alpha, \theta)$ with $\theta \neq 0$ is quasi-invariant under the operator S_a and

$$\frac{S_a P D(\alpha, \theta)}{P D(\alpha, \theta)} (dp) = \langle a^{\alpha}, \nu_0 \rangle^{-\theta/\alpha} \frac{1}{\Gamma(\theta)} \int_0^\infty \sigma^{\theta-1} \prod_{i=1}^\infty E_{\nu_0^{\alpha}} [e^{-\sigma a^{-1}(X_i)p_i}] d\sigma.$$

Proof. Since the simplicial part $T(P_{\alpha,\theta}^{c,\nu_0})$ and $T(P_{\alpha,c}^{\nu_0})$ correspond to $PD(\alpha,\theta)$ and $PD(\alpha,0)$ distributions for $\theta \neq 0$. Therefore, we have

$$\begin{split} E[\Phi((\frac{a(S_i)P_i}{\sum a(S_i)P_i})_{i=1}^{\infty})] &= E_{M_a P_{\alpha,\theta}^{c,\nu_0}}[\Phi(T(\eta))] \\ &= E_{P_{\alpha,\theta}^{c,\nu_0}}[\Phi(T(M_a\eta))] \\ &= E_{P_{\alpha,c}^{\nu_0}}[\Phi(T(M_a\eta))\frac{c_{\alpha,\theta}}{\eta(X)^{\theta}}] \\ &= E_{P_{\alpha,c}^{\nu_0}}[\Phi(T(M_a\eta))\frac{c_{\alpha,\theta}}{\langle a^{-1}, M_a\eta \rangle^{\theta}}] \\ &= E_{M_a P_{\alpha,c}^{\nu_0}}[\Phi(T(\eta))\frac{c_{\alpha,\theta}}{\langle a^{-1}, \eta \rangle^{\theta}}] \\ &= E_{P_{\alpha,\theta}^{\nu_0}}[\Phi(T(\eta))\frac{c_{\alpha,\theta}}{\langle a^{-1}, \eta \rangle^{\theta}}] \\ &= E_{P_{\alpha,\theta}^{cA_\alpha,\nu_0}}[\Phi(T(\eta))\frac{c_{\alpha,\theta}}{\langle a^{-1}, \eta \rangle^{\theta}}] \end{split}$$

We further have the above equals to

$$= E_{P_{\alpha,\theta}^{cA_{\alpha},\nu_{0}^{\alpha}}} [\Phi(T(\eta)) \frac{1}{A_{\alpha}^{\theta/\alpha} \langle a^{-1}, \bar{\eta} \rangle^{\theta}}]$$

$$= E_{PD(\alpha,\theta)} [\Phi((P_{i})_{i=1}^{\infty}) E_{(\nu_{0}^{\alpha})^{\infty}} [\frac{1}{A_{\alpha}^{\theta/\alpha} (\sum a^{-1}(S_{i})P_{i})^{\theta}}]].$$

This means that

$$\frac{S_a PD(\alpha, \theta)}{PD(\alpha, \theta)}(dp) = E_{(\nu_0^{\alpha})^{\infty}} [\frac{1}{A_{\alpha}^{\theta/\alpha} (\sum a^{-1}(S_i)p_i)^{\theta}}],$$

where $1/A_{\alpha}^{\theta/\alpha} = \langle a^{\alpha}, \nu_0 \rangle^{-\theta/\alpha}$. Using $\lambda^{-\theta} = \int_0^\infty \frac{\sigma^{\theta-1}}{\Gamma(\theta)} e^{-\lambda\sigma} d\sigma$ we can rewrite

$$E_{(\nu_0^{\alpha})^{\infty}}\left[\frac{1}{(\sum a^{-1}(S_i)p_i)^{\theta}}\right] = E_{(\nu_0^{\alpha})^{\infty}}\left[\int_0^{\infty} \frac{\sigma^{\theta-1}}{\Gamma(\theta)} e^{-\sigma \sum a^{-1}(S_i)p_i} d\sigma\right]$$
$$= \frac{1}{\Gamma(\theta)} \int_0^{\infty} \sigma^{\theta-1} \prod_{i=1}^{\infty} E_{\nu_0^{\alpha}}\left[e^{-\sigma a^{-1}(S_i)p_i}\right] d\sigma$$

Therefore,

$$\frac{S_a PD(\alpha, \theta)}{PD(\alpha, \theta)} (dp) = \langle a^{\alpha}, \nu_0 \rangle^{-\theta/\alpha} \frac{1}{\Gamma(\theta)} \int_0^\infty \sigma^{\theta-1} \prod_{i=1}^\infty E_{\nu_0^{\alpha}} [e^{-\sigma a^{-1}(S_i)p_i}] d\sigma.$$

Remark 6.2. It is easily seen that when $\theta = 0$, the density equals one, which means that $S_a P_{\alpha,\theta}$ and $P_{\alpha,\theta}$ are identically distributed which coincides with Arguin's result [2]. On the other hand, by letting $\alpha \to 0$, we have

$$\langle a^{\alpha}, \nu_0 \rangle^{-\theta/\alpha} \to \exp\{-\theta \langle \log a, \nu_0 \rangle\}$$

and $\nu_0^{\alpha} = \nu_0$ when $\alpha = 0$. This means the density we obtained is consistent with that for the Poisson-Dirichlet distribution $PD(\theta)$ [62].

6.4 Reversibility of Branching Processes

Recall that the Dirichlet process $\Pi_{\theta,\nu}$ is the stationary distribution of Fleming-Viot process with parent independent mutation defined as follows. Let S denote the type

space which is a compact Polish space, and $\mathcal{P}(S)$ be the set of probability measure on S equipped with the usual weak convergence topology. Fleming-Viot process (FV) with parent independent mutation is a $\mathcal{P}(S)$ -valued process with the generator given by

$$(\mathcal{L}\phi)(\mu) = \frac{1}{2} \int_{S} \int_{S} \mu(dx) (\delta_x(dy) - \mu(dy)) \frac{\delta^2 \phi(\mu)}{\delta \mu(x) \delta \mu(y)} + \int_{S} \mu(dx) A\left(\frac{\delta \phi(\mu)}{\delta \mu(\cdot)}\right)(x),$$

where $\delta\phi(\mu)/\delta\mu(x) = \lim_{\epsilon \to 0+} \epsilon^{-1}(\phi(\mu + \epsilon \delta_x) - \phi(\mu))$, and mutation operator

$$(Af)(x) = \frac{1}{2}\theta \int_{S} (f(\xi) - f(x))\nu(d\xi).$$

Here and after, we always assume ν is a probability measure. The domain of \mathcal{L} is given by the space of function Φ of the form $\Phi = P(\langle f_1, \mu \rangle, \dots, \langle f_n, \mu \rangle)$ where P is a polynomial function with n variables and $f_1, \dots, f_n \in D(A)$. Handa [36] used the quasi-invariance of $\Pi_{\theta,\nu}$ to characterize the reversibility of FV process with parent independent mutation. Reversibility means that the FV process is reversible with respect to $\Pi_{\theta,\nu}$ i.e.

$$\int_{\mathcal{P}(S)} \psi(\mu) \mathcal{L}\phi(\mu) \Pi_{\theta,\nu}(d\mu) = \int_{\mathcal{P}(S)} \mathcal{L}\psi(\mu)\phi(\mu) \Pi_{\theta,\nu}(d\mu), \qquad \psi, \phi \in Dom(\mathcal{L})$$

The condition for FV process incorporating mutation and selection to be reversible was dealt with by Li et.al. [48]. Handa [36] solved a more general case, in which he considered the recombination factor.

The gamma process is the stationary distribution of a typical measure-valued branching process with immigration. Together with FV process, they are two typical superprocesses. Readers can refer to [6] for a general account of superprocess. A measure-valued branching diffusion with immigration (MBI) takes values in $\mathcal{M}(S)$, the set of finite positive Borel measures on S with the topology of weak convergence. The generator of MBI is

$$\tilde{\mathcal{L}}(\phi)(\mu) = \frac{1}{2} \int_{S} \mu(dx) \frac{\delta^2 \phi(\mu)}{\delta \mu(x)^2} + \frac{1}{2} \theta \int_{S} \nu(dx) \frac{\delta \phi(\mu)}{\delta \mu(x)} - \frac{1}{2} \lambda \int_{S} \mu(dx) \frac{\delta \phi(\mu)}{\delta \mu(x)}, \quad (6.5)$$

where $\lambda \in \mathbf{R}$ and θ, ν is defined as above. As pointed by Shiga [58], the above process is reversible with respect to the gamma process P_{ν} when $\lambda = 1$.

Here we look into a general branching process with immigration μ_t with the following generator:

$$LF(\mu) = \int_{S} \mu(dx)a(x)\frac{\delta^{2}F(\mu)}{\delta\mu(x)^{2}} + \int_{S} \mu(dx)\int_{0}^{\infty} n(x,ds)(F(\mu+s\delta_{x})-F(\mu))$$

$$-s\frac{\delta F(\mu)}{\delta\mu(x)} + \int_{S} (\nu(dx)-\mu(dx)b(x))\frac{\delta F(\mu)}{\delta\mu(x)}$$

$$+ \int_{S} \mu(dx)(A\frac{\delta F(\mu)}{\delta\mu(\cdot)})(x),$$

(6.6)

where the domain of L is

$$\mathcal{D}(L) = \{ F(\mu) = \Phi(\langle f_1, \mu \rangle, \dots, \langle f_n, \mu \rangle); n \ge 1, f_i \in D(\mathcal{A}) \bigcap C(S)_+, \Phi \in C_0^2(\mathbb{R}^n) \}$$

In this case, the branching mechanism can be characterized as follows. Define

$$\Psi(x,\lambda) = -a(x)\lambda^2 + \int_0^\infty (1 - e^{-\lambda s} - \lambda s)n(x,ds) - b(x)\lambda,$$

where $a \ge 0, a, b \in C(S)$,

$$\sup_{x\in S}\int_0^\infty s\wedge s^2n(x,ds)<\infty$$

and

$$(x,\lambda) \mapsto \int_0^\infty (1 - e^{-\lambda s} - \lambda s) n(x,ds) \in C(S \times R_+).$$

The solution of the following equation

$$\frac{\partial \psi_t(f)}{\partial t} = A\psi_t(f) + \Psi(x, \psi_t(f)),$$

with $\psi_0(f) = f$ gives the Laplace transform of μ_t in the form of

$$E_{\mu}[e^{-\langle f,\mu_t\rangle}] = \exp\{-\int_0^t \langle \psi_s(f),\nu\rangle ds - \langle \psi_t(f),\mu\rangle\}.$$

The corresponding Carré du champ is

$$\Gamma(F,G) = \frac{1}{2} \{ L(FG) - LF \cdot G - F \cdot LG \}$$

= $\langle a(x) \frac{\delta F(\mu)}{\delta \mu(x)} \frac{\delta G(\mu)}{\delta \mu(x)}, \mu \rangle$
+ $\langle \int_0^\infty n(x, ds) (F(\mu + s\delta_x) - F(\mu)) (G(\mu + s\delta_x) - G(\mu)), \mu \rangle,$

where $F, G \in \mathcal{D}(L)$. In order to satisfy the following equation

$$\Gamma(FH,G) + \Gamma(HG,F) - \Gamma(H,FG) = 2H\Gamma(F,G), \tag{6.7}$$

which is essential in our proof, we have to set n(x, ds) = 0. Suppose that the process L is reversible with respect to some measure Π on $\mathcal{M}(S)$, which is equivalent to say

$$\int \Phi L \Psi d\Pi = \int \Psi L \Phi d\Pi, \qquad \Phi, \Psi \in \mathcal{D}(L)$$

Define

$$\langle f, c(\mu) \rangle = \langle f, \nu \rangle - \langle bf, \mu \rangle + \langle Af, \mu \rangle$$

and

$$\Lambda(f,\mu) = \int_0^1 \langle f/a, c(S_{uf}(\mu)) \rangle du, \qquad (6.8)$$

where $S_f \mu(dx) = e^f \mu(dx)$. S_f forms a transformation group on $\mathcal{M}(S)$, where f is a bounded measurable function on S such that

$$S_{f+g}\mu = S_f(S_g\mu),$$

and $(S_f)^{-1} = S_{-f}$. Let K be a linear subspace of bounded measurable functions of S and $\Lambda : K \times \mathcal{P}(S) \to R$ be such that for each $f \in K$ the function $\mu : \mu \to \Lambda(f,\mu)$ is Borel measurable. Next we will try to show that $\Pi \circ S_f$ and Π are equivalent and the density is given by $e^{\Lambda(f,\mu)}$ for $f \in K$. It means that Π is K-quasi-invariant with cocycle $\Lambda(f,\mu)$. As a result, $\Lambda(f,\mu)$ has to satisfy the cocycle identity:

$$\Lambda(f+g,\mu) = \Lambda(f, S_g\mu) + \Lambda(g,\mu).$$

So as to let Λ defined by (6.8) satisfy the cocycle identity it can be seen that A = 0. In this case, $\Lambda(f, \mu) = \langle f/a, \nu \rangle - \langle (e^f - 1)b/a, \mu \rangle$.

Theorem 6.3. Suppose the branching process with immigration $\{\nu_t, t \ge 0\}$ is defined by generator L (6.6) with A = n = 0 and $a, b \in C(S), a \ge 0$. The domain of L is given by

$$\mathcal{D}(L) = \{F(\mu) = \Phi(\langle f_1, \mu \rangle, \dots, \langle f_n, \mu \rangle); n \ge 1, f_i \in C(S)_+, \Phi \in C_0^2(\mathbb{R}^n)\}$$

Let Π be a Borel probability measure on $\mathcal{M}(S)$. Then L is reversible with respect to Π if and only if Π is quasi-invariant under the transformation S_f with f a bounded measurable function on S and density given by $e^{\Lambda(f,\mu)}$ defined by (6.8).

Proof. We first show that the reversibility implies the quasi-invariance. Define

 $\mu_t = S_{-atf}\mu$

and

$$\Phi_t(\mu) = \Phi(\mu_t) = \Phi(\langle f_1, \mu_t \rangle, \dots, \langle f_n, \mu_t \rangle),$$

where $\Phi \in \mathcal{D}(L), f_i \in \mathcal{B}(S), i = 1, 2, \dots, n$. Set

$$Z(t) = \int \Phi(\mu_t) e^{-\Lambda(atf,\mu_t)} \Pi(d\mu).$$

We will obtain the quasi-invariance by showing Z'(t) = 0. First we note that

$$\frac{d}{dt}\Phi(\mu_t) = \sum_{i=1}^n \frac{\partial\Phi}{\partial x_i} \langle f_i(-af), \mu_t \rangle$$

and

$$\frac{\delta\Phi_t(\mu)}{\delta\mu(x)} = \sum_{i=1}^n \frac{\partial\Phi}{\partial x_i} f_i e^{-atf}.$$

As a result, the above two identities lead to

$$\langle af \frac{\delta \Phi_t(\mu)}{\delta \mu(x)}, \mu \rangle = \sum_{i=1}^n \frac{\partial \Phi}{\partial x_i} \langle af f_i e^{-atf}, \mu \rangle = \sum_{i=1}^n \frac{\partial \Phi}{\partial x_i} \langle af f_i, \mu_t \rangle = -\frac{d}{dt} \Phi(\mu_t).$$

Therefore,

$$\begin{split} \Lambda(atf,\mu_t) &= \int_0^1 \langle tf, c(S_{uatf}(\mu_t)) \rangle du \\ &= t \int_0^1 \langle f, c(S_{-(1-u)atf}(\mu)) \rangle du \\ &= t \int_0^1 \langle f, c(S_{-uatf}(\mu)) \rangle du \\ &= \int_0^t \langle f, c(S_{-uaf}(\mu)) \rangle du \\ &= \int_0^t \langle f, c(\mu_u) \rangle du. \end{split}$$

From this, we have

$$\frac{d}{du}\langle f, c(\mu_u)\rangle = -\langle \mu, af \frac{\delta\langle f, c(\mu_u)\rangle}{\delta\mu(x)}\rangle.$$

Define $\tilde{\Phi}_t(\mu) = \Phi(\mu_t)e^{-\Lambda(atf,\mu_t)}$. Then we obtain the following

$$\frac{\delta \tilde{\Phi}_t(\mu)}{\delta \mu(x)} = \frac{\delta \Phi_t(\mu)}{\delta \mu(x)} e^{-\Lambda(atf,\mu_t)} - \tilde{\Phi}_t(\mu) \int_0^t \frac{\delta \langle f, c(\mu_u)}{\delta \mu(x)} du,$$

and

$$\begin{aligned} \langle af \frac{\delta \tilde{\Phi}_t(\mu)}{\delta \mu(x)} . \mu \rangle &= \langle af \frac{\delta \Phi_t(\mu)}{\delta \mu(x)}, \mu \rangle e^{-\Lambda(atf,\mu_t)} - \tilde{\Phi}_t(\mu) \int_0^t \langle af \frac{\delta \langle f, c(\mu_u) \rangle}{\delta \mu(x)}, \mu \rangle du \\ &= -\frac{d}{dt} \Phi(\mu_t) e^{-\Lambda(atf,\mu_t)} + \tilde{\Phi}_t(\mu) (\langle f, c(\mu_t) \rangle - \langle f, c(\mu) \rangle). \end{aligned}$$

Therefore,

$$Z'(t) = \int \frac{d}{dt} \Phi(\mu_t) e^{-\Lambda(atf,\mu_t)} - \tilde{\Phi}_t(\mu) \langle f, c(\mu_t) \rangle \Pi(d\mu)$$

$$= \int -\langle af \frac{\delta \tilde{\Phi}_t(\mu)}{\delta \mu(x)} \cdot \mu \rangle + \tilde{\Phi}_t(\mu) (\langle f, c(\mu_t) \rangle - \langle f, c(\mu) \rangle) - \tilde{\Phi}_t(\mu) \langle f, c(\mu_t) \rangle \Pi(d\mu)$$

$$= \int -\langle af \frac{\delta \tilde{\Phi}_t(\mu)}{\delta \mu(x)} \cdot \mu \rangle - \tilde{\Phi}_t(\mu) \langle f, c(\mu) \rangle \Pi(d\mu).$$

Let $\Psi(\mu) = \langle f, \mu \rangle$. Then $L\Psi(\mu) = \langle f, c(\mu) \rangle$ and $\frac{\delta\Psi(\mu)}{\delta\mu(x)} = f(x)$. The above is equivalent to $Z'(t) = \int -\Gamma(\tilde{\Phi}, \Psi) - \tilde{\Phi}_t L\Psi \Pi(d\mu)$

$$Z'(t) = \int -\Gamma(\tilde{\Phi}_t, \Psi) - \tilde{\Phi}_t L \Psi \Pi(d\mu).$$

Clearly, $\tilde{\Phi}_t, \Psi \in \mathcal{D}(L)$. If L is reversible w.r.t. the measure $\Pi(d\mu)$, then Z'(t) = 0. Therefore, Z(1) = Z(0) which is equivalent to

$$\int \Phi(S_{-af}\mu)e^{-\Lambda(af,S_{-af}\mu)}\Pi(d\mu) = \int \Phi(\mu)\Pi(d\mu).$$

Thus, we have

$$\int \Phi(S_{-f}\mu)e^{-\Lambda(f,S_{-f}\mu)}\Pi(d\mu) = \int \Phi(\mu)\Pi(d\mu),$$

for any $f \in \mathcal{B}(S)$. Replacing $\Phi(\mu)$ by $\Phi(\mu)e^{\Lambda(f,\mu)}$, we have

$$\int \Phi(S_{-f}\mu)\Pi(d\mu) = \int \Phi(\mu)e^{\Lambda(f,\mu)}\Pi(d\mu),$$

i.e. $\int \Phi(\mu) \Pi \circ S_f(d\mu) = \int \Phi(\mu) e^{\Lambda(f,\mu)} \Pi(d\mu).$

Secondly, we try to show the above quasi-invariance implies reversibility. From the above argument, we can see that if $\Pi \circ S_f \ll \Pi$ with the density $e^{\Lambda(f,\mu)}$, then Z(t) must be a constant. This induces that Z'(t) = 0. Hence

$$0 = Z'(0) = \int -\langle af \frac{\delta \Phi(\mu)}{\delta \mu(x)} . \mu \rangle - \Phi(\mu) \langle f, c(\mu) \rangle \Pi(d\mu).$$

Therefore,

$$0 = \int -\Gamma(\Phi, \Psi) - \Phi L \Psi \Pi(d\mu), \qquad (6.9)$$

for $\Phi \in \mathcal{D}(L)$ and $\Psi = \langle f, \mu \rangle$. Suppose $\Psi(\mu) = \Psi_1(\mu)\Psi_2(\mu) = \langle f_1, \mu \rangle \langle f_2, \mu \rangle$. Using (6.7), we have

$$-\int \Phi L(\Psi_1 \Psi_2) \Pi(d\mu) = \int \Phi \{-2\Gamma(\Psi_1, \Psi_2) - L\Psi_1 \cdot \Psi_2 - \Psi_1 L\Psi_2\} \Pi(d\mu)$$

$$= -\int (\Gamma(\Phi \Psi_1, \Psi_2) + \Gamma(\Phi \Psi_2, \Psi_1) - \Gamma(\Phi, \Psi_1 \Psi_2)) \Pi(d\mu)$$

$$+ \int \Gamma(\Phi \Psi_1, \Psi_2) + \Gamma(\Phi \Psi_2, \Psi_1) \Pi(d\mu)$$

$$= \int \Gamma(\Phi, \Psi_1 \Psi_2) \Pi(d\mu).$$

This indicates that (6.9) holds for $\Psi(\mu) = \prod_{i=1}^{n} \langle f_i, \mu \rangle$, for $f_i \in C(S)$. Therefore, (6.9) holds for $\Psi \in \mathcal{D}(L)$.

Example 6.4. Consider the following generator:

$$L_1 F(\mu) = \frac{1}{2} \int_S \mu(dx) \frac{\delta^2 F(\mu)}{\delta \mu(x)^2} + \frac{1}{2} \int_S (\theta \nu_0(dx) - \lambda \mu(dx)) \frac{\delta F(\mu)}{\delta \mu(x)},$$

where $a(x) = \frac{1}{2}$, $\nu(dx) = \frac{\theta}{2}\nu_0(dx)$, $\nu_0 \in \mathcal{P}(S)$ and $b(x) = \frac{\lambda}{2}$, A = 0. The cocycle

$$\Lambda(f,\mu) = \langle f/\frac{1}{2}, \theta\nu_0/2 \rangle - \langle (e^f - 1)\frac{\lambda}{2}/\frac{1}{2}, \mu \rangle = \theta \langle f, \nu_0 \rangle - \lambda \langle e^f - 1, \mu \rangle.$$

On the other hand, it is known that L_1 is reversible w.r.t. the random measure $\Gamma_{\theta,\nu_0}^{\lambda^{-1}}$ whose Laplace transform is given by

$$\int e^{-\langle f,\mu\rangle} \Gamma_{\theta,\nu_0}^{\lambda^{-1}}(d\mu) = \exp\{-\theta \langle \log(1+\lambda^{-1}f),\nu_0\rangle\}.$$

Then, we have

$$\int e^{-\langle g, S_{-f}\mu \rangle} \Gamma_{\theta,\nu_0}^{\lambda^{-1}}(d\mu) = \int e^{-\langle e^{-f}g,\mu \rangle} \Gamma_{\theta,\nu_0}^{\lambda^{-1}}(d\mu)$$

$$= \exp\{-\theta \langle \log(1+\lambda^{-1}e^{-f}g),\nu_0 \rangle\}$$

$$= \exp\{-\theta \langle \log e^{-f} + \log(e^f + \lambda^{-1}g).\nu_0 \rangle\}$$

$$= \exp\{\theta \langle f,\nu_0 \rangle\} \exp\{-\theta \langle \log(1+\lambda^{-1}(\lambda e^f + g - \lambda).\nu_0 \rangle\}$$

$$= \exp\{\theta \langle f,\nu_0 \rangle\} \int e^{-\langle \lambda e^f + g - \lambda,\mu \rangle} \Gamma_{\theta,\nu_0}^{\lambda^{-1}}(d\mu)$$

which also gives the density of $\Gamma_{\theta,\nu_0}^{\lambda^{-1}} \circ S_f$ w.r.t $\Gamma_{\theta,\nu_0}^{\lambda^{-1}}$ is $\exp\{\theta\langle f,\nu_0\rangle\}e^{-\langle\lambda e^f-\lambda,\mu\rangle}$

Remark 6.5. The quasi-invariance property of the gamma process was also considered by Stannat [60]. He associated this property with an analogue of the Cameron-Martin-Girsanov theorem for the classical Wiener measure. The transformed gamma process is the symmetrizing measure for the process whose generator is a first-order perturbabtion of the original one.

Chapter 7 Summary

In this thesis we study the impact of various influential factors on the population evolutionary process.

In Chapter 2, we study the limiting distribution of age-ordered family frequencies of a linear birth process with immigration as time goes to infinity. The result shows that the cumulative effect of time let the distribution of age-ordered family proportions converge to the two-parameter GEM distribution.

In Chapter 3 we first use the sampling formula to derive the Laplace transform of the two-parameter Poisson-Dirichlet distribution. Also, the sampling formula is derived from the Laplace transform. This indicates the robustness of sampling formula which is used a basic tool in statistical inference of population genetics. Furthermore, letting mutation rate go to infinity we obtain a Gaussian limit distribution of the scaled homozygosity functions for the IMNA model at a fixed time. Since there exists a similar behavior at the equilibrium state, we conclude that with any starting point the model will look like its equilibrium in a finite time under the influence of large mutation. In the end of Chapter 3, we use the sampling formula of the selection model to derive the Laplace transform.

The selection impact is further studied for the two-parameter model in Chapter 4. We consider the selection intensity given by different powers and study the limiting behavior as mutation goes to infinity. The asymptotic behavior for the two-parameter Poisson-Dirichlet distribution with selection looks like the neutral model. Therefore, the impact of selection is hardly detected when mutation rate and selection intensity both approach infinity.

Chapter 5 considers the generalized two-parameter infinite alleles model. The behavior of the model at finite time is described by the transition function, which is absolutely continuous w.r.t. the stationary distribution $PD(\alpha, \theta)$. The transition

density is derived by the eigenfunction expansion. With the transition density we find the probability generating function of random sampling from the two-parameter extension model at a fixed time, as a comparison of the sampling formula at equilibrium.

In the last chapter, we establish the quasi-invariance property of the $PD(\alpha, \theta)$ distribution with respect to the Markovian transformation group. In the other aspect, we show the equivalence between the reversibility of a class measure-valued branching process with immigration and the quasi-invariance of its invariant measure with respect to the multiplication transformation group. This conclusion implicitly establishes the connection of the stationary distribution under two different kinds of mappings.

References

- [1] C. Antoniak (1974). Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. *Annals of Statistics* **2**, 1152–1174.
- [2] L.-P. Arguin (2007). A dynamical characterization of Poisson-Dirichlet distributions. *Electronic Communication in Probability* 12, 283–290.
- [3] J. Bertoin (2008). Two-parameter Poisson-Dirichlet measures and reversible exchangeable fragmentation-coalescence processes. *Combinatorics, Probability* and Computing 17, no. 3, 329–337.
- [4] M. A. Carlton (1999). Applications of the Two-Parameter Poisson-Dirichlet Distribution. PhD thesis, U.C.L.A.
- [5] D.J. Daley and D. Vere-Jones (2003). An introduction to the theory of point processes. Vol. I, Elementary theory and methods, Second edition, Springer-Verlag, New York.
- [6] D. Dawson (1993). Measure-valued Markov processes, Ecole d'Eté de Probabilités de Saint Flour, Lecture Notes in Math, 1541, pp.1-260, Springer-Verlag, Berlin.
- [7] D.A. Dawson and S. Feng (2006). Asymptotic behavior of the Poisson-Dirichlet distribution for large mutation rate. *Annals of Applied Probability* **16**, 562–582.
- [8] P. Donnelly (1986). Partition structures, Pólya urns, the Ewens sampling formula, and the age of allels. *Theoretical Population Biology* **30**, 271–288.
- [9] P. Donnelly and P. Joyce (1989). Continuity and weak convergence of ranked and size-biased permutations on the infinite simplex. *Stochastic Processes and their Applications* **31**, 89–104.

- [10] P. Donnelly and S. Tavaré (1986). The ages of alleles and a coalescent. Advances in Applied Probability 18, 1–19.
- [11] P. Donnelly and S. Tavaré (1987). The population genealogy of the infinitelymany neutral alleles model. *Journal of Mathematical Biology* 25, 381–391.
- [12] S.N. Ethier (1990). The distribution of the frequencies of age-ordered alleles in a diffusion model. *Advances in Applied Probability* **22**, 519–532.
- [13] S.N. Ethier (1992). Eigenstructure of the infinitely-many-neutral-alleles diffusion model. Journal of Applied Probability 29, 487–498.
- [14] S.N. Ethier and R.C. Griffiths (1993). The transition function of a Fleming-Viot process. Annals of Probability 21, 1571–1590.
- [15] S.N. Ethier and T.G. Kurtz (1981). The infinitely-many-neutral-alleles diffusion model. Advances in Applied Probability 13, 429–452.
- [16] S.N. Ethier and T.G. Kurtz (1986). Markov processes: characterization and convergence. John Wiley & Sons, Inc., New York.
- [17] S.N. Ethier and T.G. Kurtz (1994). Convergence to Fleming-Viot processes in the week atomic topology. *Stochastic Processes and their Applications* 54, 1–27.
- [18] W.J. Ewens (1972). The sampling theory of selectively neutral alleles. Theoretical Population Biology 3, 87–112.
- [19] W.J. Ewens (1988). Population genetics theory the past and the future. In S. Lessard, editor, Mathematical and Statistical Problems in Evolution. University of Montreal Press, Montreal.
- [20] W.J. Ewens (2004). Mathematical Population Genetics, I. Theoretical Introduction. Springer-Verlag, New York.
- [21] S. Feng (2007). Large deviations for Dirichlet processes and Poisson-Dirichlet distribution with two parameters. *Electronic Journal of Probability* 12, 787– 807.
- [22] S. Feng (2010). The Poisson-Dirichlet distribution and related topics: models and asymptotic behaviors. Springer, Heidelberg.

- [23] S. Feng and F.Q. Gao (2008). Moderate deviations principles for Poisson-Dirichlet distribution. Annals of Applied Probability 18, no. 5, 1794–1824.
- [24] S. Feng and F.Q. Gao (2010). Asymptotic results for the two-parameter Poisson-Dirichlet distribution. *Stochastic Processes and their Applications* 120, 1159–1177.
- [25] S. Feng and F.M. Hoppe (1998). Large deviation principles for some random combinatorial structures in population genetics and Brownian motion. Annals of Applied Probability 8, no. 4, 975–994.
- [26] S. Feng and W. Sun (2010). Some diffusion processes associated with two parameter Poisson-Dirichlet distribution and Dirichlet process. *Probability The*ory and Related Fields 148, 501–525.
- [27] S. Feng, W. Sun, F.Y. Wang, and F. Xu (2011). Functional inequalities for the two parameter extension of the infinitely-many-neutral-alleles diffusion. *Journal of Functional Analysis* 260, 399–413.
- [28] S. Feng and F.Y. Wang (2007). A class of infinite-dimensional diffusion processes with connection to population genetics. *Journal of Applied Probability* 44, 938–949.
- [29] T.S. Ferguson (1973). A Bayesian analysis of some nonparametric problems. Annals of Statistics 1, 209–230.
- [30] J.H. Gillespie (1999). The role of population size in molecular evolution. *Theoretical Population Biology* 55, 145–156.
- [31] R.C. Griffiths (1979). Exact sampling distribution from the infinite neutral alleles model. Advances in Applied Probability 11, 326–354.
- [32] R.C. Griffiths (1979a). A transition density expansion for a multi-allele diffusion model. Advances in Applied Probability 11, 310–325.
- [33] R.C. Griffiths (1979b). On the distribution of allele frequencies in a diffusion model. *Theoretical Population Biology* 15, 140–158.
- [34] R.C. Griffiths (1988). On the distribution of points in a Poisson Dirichlet process. *Journal of Applied Probability* 25, 336–345.
- [35] K. Handa (2001). Quasi-invariant measures and their characterization by conditional probabilities. *Bulletin of Mathematical Sciences* **125**, 583–604.

- [36] K. Handa (2002). Quasi-invariance and reversibility in the Fleming-Viot process. Probability Theory and Related Fields 122, 545–566.
- [37] K. Handa (2005). Sampling formulae for symmetric selection. *Electronic Com*munications in Probability 10, 223–234.
- [38] K. Handa (2009). The two-parameter Poisson-Dirichlet point process. Bernoulli 15, 1082–1116.
- [39] F.M. Hoppe (1984). Plya-like urns and the Ewens' sampling formula. Journal of Mathematical Biology 20, no. 1, 91–94.
- [40] F.M. Hoppe (1987). The sampling theory of neutral alleles and an urn model in population genetics. *Journal of Mathematical Biology* **25**, no. 2, 123–159.
- [41] P. Joyce, S.M. Krone, and T.G. Kurtz (2002). Gaussian limits associated with the Poisson-Dirichlet distribution and Ewens sampling formula. *Annals* of Applied Probability **12**, 101–124.
- [42] P. Joyce, S.M. Krone, and T.G. Kurtz (2003). When can one detect overdominant selection in the infinite-alleles model? Annals of Applied Probability 13, 181–212.
- [43] J.F.C. Kingman (1975). Random discrete distributions. Journal of the Royal Statistical Society, Series B 37, 1–22.
- [44] J.F.C. Kingman (1977). The population structure associated with the Ewens sampling formula. *Theoretical Population Biology* 11, 274–283.
- [45] J.F.C. Kingman (1978). The representation of partition structures. Journal of the London Mathematical Society 18, 374–380.
- [46] J.F.C. Kingman (1993). *Poisson processes*. Oxford University Press.
- [47] A.N. Kolmogoroff (1937). Zur Umkehrbarkeit der statistischen Naturgesetze. Mathematische Annalen 113, 766–772.
- [48] Z. Li, T. Shiga, and L. Yao (1999). A reversibility problem for Fleming-Viot processes. *Electronic Communication in Probability* 4, 71–82.
- [49] E. Nelson (1958). The adjoint Markoff process. Duke Mathematical Journal 25, 671–690.
- [50] C.P. Patil and C. Taillie (1977). Diversity as a concept and its implications for random communities. *Bull. Int. Statist. Inst.* XLVII, 497–515.
- [51] M. Perman, J. Pitman, and M. Yor (1992). Size-biased sampling of Poisson point processes and excursions. *Probability Theory and Related Fields* 92, 21– 39.
- [52] L. Petrov (2009). Two-parameter family of infinite-dimensional diffusions on the Kingman simplex. *Functional Analysis and Its Applications* **43**, 279–296.
- [53] J. Pitman (1992). The two-parameter generalization of Ewens' random partition structure. Technical Report 345, Department of Statistics, University of California, Berkeley.
- [54] J. Pitman (1995). Exchangeable and partially exchangeable random partitions. Probability Theory and Related Fields 102, 145–158.
- [55] J. Pitman (1996). Random discrete distributions invariant under size-biased permutation. Advances in Applied Probability 28, 525–539.
- [56] J. Pitman (1996a). Some developments of the Blackwell-MacQueen urn scheme. Statistics, probability and game theory, 245–267, IMS Lecture Notes Monogr. Ser. 30, Inst. Math. Statist., Hayward, CA.
- [57] J. Pitman and M. Yor (1997). The two-parameter Poisson-Dirichlet distribution derived from a stable subordiantor. Annals of Probility 25, no. 2, 855–900.
- [58] T. Shiga (1991). Infinite-dimensional diffusion processes arising in population genetics. Lecture notes based on lectures given at Taiwan University.
- [59] W. Stannat (2000). On the validity of the log-Sobolev inequality for symmetric Fleming-Viot operators. Annals of Probability 28, No. 2, 667–684.
- [60] W. Stannat (2003). Spectral properties for a class of continuous state branching processes with immigration. *Journal of Functional Analysis* **201**, 185–227.
- [61] S. Tavaré (1987). The birth process with immigration, and the genealogical structure of large populations. *Journal of Mathematical Biology* 25, no. 2, 161–168.

- [62] N. Tsilevich, A. Vershik, and M. Yor (2000). Distinguished properties of the gamma process and related topics. Prpublication du Laboratoire de Probabilits et Modeles Alatoires. No. 575.
- [63] N. Tsilevich, A. Vershik, and M. Yor (2001). An infinite-dimensional analogue of the Lebesgue measure and distinguished properties of the Gamma process. *Journal of Functional Analysis* 185, no. 1, 274–296.
- [64] M.-K. von Renesse, M. Yor, and L. Zambotti (2008). Quasi-invariance properties of a class of subordinators. *Stochastic Processes and their Applications* 118, 2038–2057.
- [65] G.A. Watterson (1976). The stationary distribution of the infinitely-many neutral alleles diffusion model. *Journal Applied Probability* 13, 639-651; correction. ibid. 14 (1977) 897.
- [66] S. Wright (1949). Adaption and selection. In Genetics, Paleontology, and Evolution, ed., G.L. Jepson, E. Mayr, G.G. Simpson, 365–389. Princeton University Press.
- [67] F. Xu (2009). A central limit theorem associated with the transformed twoparameter Poisson–Dirichlet distribution. *Journal of Applied Probability* 46, no. 2, 392–401.
- [68] F. Xu (2011). The sampling formula and Laplace transform associated with the two-parameter Poisson-Dirichlet distribution. To appear in Advances in Applied Probability (December 2011).
- [69] M. Yor (2007). Some remarkable properties of gamma processes. Advances in Mathematical Finance I, 37–47.