Electronic Journal of Probability

Vol. 12 (2007), Paper no. 27, pages 787–807.

Journal URL
http://www.math.washington.edu/~ejpecp/

Large deviations for Dirichlet processes and Poisson-Dirichlet distribution with two parameters*

Shui Feng
Department of Mathematics
and Statistics
McMaster University
Hamilton, Ontario
Canada L8S 4K1
shuifeng@mcmaster.ca

Abstract

Large deviation principles are established for the two-parameter Poisson-Dirichlet distribution and two-parameter Dirichlet process when parameter θ approaches infinity. The motivation for these results is to understand the differences in terms of large deviations between the two-parameter models and their one-parameter counterparts. New insight is obtained about the role of the second parameter α through a comparison with the corresponding results for the one-parameter Poisson-Dirichlet distribution and Dirichlet process.

Key words: GEM representation, Poisson-Dirichlet distribution, two parameter Poisson-Dirichlet distribution, Dirichlet processes, large deviation principles.

AMS 2000 Subject Classification: Primary 60F10; Secondary: 92D10.

Submitted to EJP on November 14, 2006, final version accepted May 22, 2007.

*Research supported by the Natural Science and Engineering Research Council of Canada

1 Introduction

For any α in $[0,1)$ 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. Set

$$X_1^{\theta,\alpha} = U_1, \; X_n^{\theta,\alpha} = (1-U_1)\cdots(1-U_{n-1})U_n, \; n \geq 2. \tag{1.1}$$

Then with probability one

$$\sum_{k=1}^{\infty} X_k^{\theta,\alpha} = 1,$$

and the law of $(X_1^{\theta,\alpha}, X_2^{\theta,\alpha}, ...)$ is called the two-parameter GEM distribution denoted by $GEM(\theta, \alpha)$.

Let $\mathbf{P}(\alpha, \theta) = (P_1(\alpha, \theta), P_2(\alpha, \theta), ...)$ denote $(X_1^{\theta,\alpha}, X_2^{\theta,\alpha}, ...)$ in descending order. The law of $\mathbf{P}(\alpha, \theta)$ is called the two-parameter Poisson-Dirichlet distribution and, following (20), is denoted by $PD(\alpha, \theta)$.

Let $\xi_k, k = 1, ...$ be a sequence of i.i.d. random variables with common diffusive distribution ν on $[0,1]$, i.e., $\nu(x) = 0$ for every x in $[0,1]$. Set

$$\Xi_{\theta,\alpha,\nu} = \sum_{k=1}^{\infty} P_k(\alpha, \theta)\delta_{\xi_k}. \tag{1.2}$$

We call the law of $\Xi_{\theta,\alpha,\nu}$ the two-parameter Dirichlet process, denoted by $Dirichlet(\theta, \alpha, \nu)$.

If $\alpha = 0$ in (1.1), we get the well known GEM distribution, the Poisson-Dirichlet distribution, and Dirichlet process denoted respectively by $GEM(\theta)$, $PD(\theta)$, and $Dirichlet(\theta, \nu)$.

There is a vast literature on GEM distribution, the Poisson-Dirichlet distribution, and Dirichlet process. The areas where they appear include Baysian statistics ((9)), combinatorics ((22)), ecology ((10)), population genetics ((7)), and random number theory ((23)).

The Poisson-Dirichlet distribution was introduced by Kingman (14) to describe the distribution of gene frequencies in a large neutral population at a particular locus. In the population genetics setting it is intimately related to the *Ewens sampling formula* that describes the distribution of the allelic partition of a sample of size n genes selected from the population. The component $P_k(\theta)$ represents the proportion of the kth most frequent alleles. If u is the individual mutation rate and N_e is the effective population size, then the parameter $\theta = 4N_e u$ is the scaled population mutation rate.

The GEM distribution can be obtained from the Poisson-Dirichlet distribution through a procedure called *size-biased sampling*. Here is a brief explanation. Consider a population consisting of individuals of countable number of different types labelled $\{1, 2, ...\}$. Assume that the proportion of type i individual in the population is p_i. A sample is randomly selected from the population and the type of the selected individual is denoted by $\sigma(1)$. Next remove all individuals of type $\sigma(1)$ from the population and then randomly select the second sample. This is repeated to get more samples. Denote the type of the ith selected sample by $\sigma(i)$. Then $(p_{\sigma(1)}, p_{\sigma(2)}, ...)$ is called a *size-biased permutation* of $(p_1, p_2, ...)$. The sequence $X_k^{\theta}, k = 1, 2, ...$ defined in (1.1) with $\alpha = 0$ has the same distribution as the size-biased permutation of $\mathbf{P}(\theta) = \mathbf{P}(0, \theta)$. The name GEM distribution is termed by Ewens after R.C. Grifffiths, S. Engen and J.W. McCloskey for their contributions to the development of the structure. The Dirichlet process first appeared in (9).

The literature on the study of Poisson-Dirichlet distribution and Dirichlet process with two parameters is relatively small but is growing rapidly. Carlton (2) includes detailed calculations of moments and parameter estimations of the two-parameter Poisson-Dirichlet distribution. The most comprehensive study of the two-parameter Poisson-Dirichlet distribution is carried out in Pitman and Yor (20). In (6) and the references therein one can find connections between two-parameter Poisson-Dirichlet distribution and models in physics including mean-field spin glasses, random map models, fragmentation, and returns of a random walk to origin. The two-parameter Poisson-Dirichlet distribution also found its applications in macroeconomics and finance ((1)).

The Poisson-Dirichlet distribution and its two-parameter counterpart have many similar structures including the urn construction in (12) and (8), GEM representation, sampling formula ((18)), etc.. A special feature of the two-parameter Poisson-Dirichlet distribution is included in Pitman (17) where it is shown that the two-parameter Poisson-Dirichlet distribution is the most general distribution whose size-biased permutation has the same distribution as the GEM representation (1.1).

The objective of this paper is to establish large deviation principles (henceforth LDP) for $GEM(\theta, \alpha)$, $PD(\alpha, \theta)$, and $Dirichlet(\theta, \alpha, \nu)$ with positive α when θ approaches infinity. Noting that for the one-parameter model, θ is the scaled population mutation rate. For fixed individual mutation rate u, large θ corresponds to large population size. In the two parameter setting, we no longer have the same explanation. But it can be seen from (1.1) that for nonzero α, large θ plays a very similar role mathematically as in the case $\alpha = 0$.

LDP for $Dirichlet(\theta, \nu)$ has been established in (15) and (3) using different methods. Recently in (4), the LDP is established for $PD(\theta)$. From (1.1), one can see that for every fixed k, the impact of α diminishes as θ becomes large. It is thus reasonable to expect similar LDPs between $GEM(\theta)$ and $GEM(\theta, \alpha)$. But in $PD(\alpha, \theta)$ and $Dirichlet(\theta, \alpha, \nu)$, every term in (1.1) counts. It is thus reasonable to expect that the LDP for $PD(\theta)$ and $Dirichlet(\theta, \nu)$ are different from the corresponding LDPs for $PD(\alpha, \theta)$ and $Dirichlet(\theta, \alpha, \nu)$. But it turns out that the impact of α only appears in the LDP for $Dirichlet(\theta, \alpha, \nu)$.

Result of LDPs turns out to be quite useful in understanding certain critical phenomenon in population genetics. In Gillespie (11) simulations were done for several models in order to understand the roles of mutation and selection forces in the evolution of a population. In the simulations for the infinite-alleles model with selective over-dominance, it was observed that when mutation rate and selection intensity get large together with the population size or θ the selective model behaves like that of a neutral model. In other words, the role of mutation and the role of selection are indistinguishable at certain scale associated with the population size. Through the study of the stationary distribution of the infinitely many alleles diffusion with heterozygote advantage, it was shown in Joyce, Krone and Kurtz (13) that phase transitions occur depending on the relative strength of mutation rate and selection intensity. The result of LDP for $PD(\theta)$ provides a more natural way of studying these phase transitions ((4)).

LDP for $GEM(\theta, \alpha)$ is given in Section 2. Using Perman's formula and an inductive structure, we establish the LDP for $PD(\alpha, \theta)$ in Section 3. The LDP for $Dirichlet(\theta, \alpha, \nu)$ is established in Section 4 using the subordinator representation in (20) and a combination of the methods in (15) and (3). Further comments are included in Section 5.

The reference (5) includes all the terminologies and standard techniques on large deviations used in this article. Since the state spaces encountered here are all compact, there is no need to distinguish between a rate function and a good rate function.

2 LDP for GEM

Let $E = [0,1]$, and E^∞ be the infinite Cartesian product of E. Set

$$\mathcal{E} = \{(x_1, x_2, ...) \in E^\infty : \sum_{k=1}^{\infty} x_k \leq 1\},$$

and consider the map

$$G : E^\infty \to \mathcal{E}, (u_1, u_2, ...) \to (x_1, x_2, ..)$$

with

$$x_1 = u_1, x_n = u_n(1-u_1)\cdots(1-u_{n-1}),\ n \geq 2.$$

By a proof similar to that used in Lemma 3.1 in (4), one obtains the following lemma.

Lemma 2.1. *For each $k \geq 1$, the family of the laws of U_k satisfies a LDP on E with speed θ and rate function*

$$I_1(u) = \begin{cases} \log \frac{1}{1-u}, & u \in [0,1) \\ \infty, & u = 1 \end{cases} \tag{2.3}$$

Theorem 2.2. *The family $\{GEM(\theta,\alpha) : \theta > 0, 0 < \alpha < 1\}$ satisfies a LDP on $\mathcal{E}$ with speed θ and rate function*

$$S(x_1, x_2, ...) = \begin{cases} \log \frac{1}{1-\sum_{i=1}^{\infty} x_i}, & \sum_{i=1}^{\infty} x_i < 1 \\ \infty, & \textit{else.} \end{cases}$$

Proof: Since $U_1, U_2, ..$ are independent, for every fixed n the law of $(U_1, ..., U_n)$ satisfies a LDP with speed θ and rate function $\sum_{i=1}^n I_1(u_i)$. For any $\mathbf{u}, \mathbf{v}$ in E^∞, set

$$|\mathbf{u} - \mathbf{v}| = \sum_{i=1}^{\infty} \frac{|u_i - v_i|}{2^i}.$$

Then for any $\delta'' > 0$ and $\mathbf{u}$ in E^∞, one can choose $n \geq 1$ and small enough $0 < \delta' < \delta < \delta''$ such that

$$\{\mathbf{v} \in E^\infty : \max_{1\leq i\leq n} |v_i - u_i| < \delta'\} \subset \{\mathbf{v} \in E^\infty : |\mathbf{v} - \mathbf{u}| < \delta\},$$
$$\{\mathbf{v} \in E^\infty : |\mathbf{v} - \mathbf{u}| \leq \delta\} \subset \{\mathbf{v} \in E^\infty : \max_{1\leq i\leq n} |v_i - u_i| < \delta''\},$$

which implies

$$\lim_{\delta\to 0} \liminf_{\theta\to\infty} \frac{1}{\theta} \log P\{|(U_1, U_2, ...) - \mathbf{u}| < \delta\} \geq -\sum_{i=1}^{n} I_1(u_i), \tag{2.4}$$

$$\lim_{\delta\to 0} \limsup_{\theta\to\infty} \frac{1}{\theta} \log P\{|(U_1, U_2, ...) - \mathbf{u}| \leq \delta\} \leq -\sum_{i=1}^{n} I_1(u_i). \tag{2.5}$$

Since E^∞ is compact, by letting n approach infinity in (2.4) and (2.5), it follows that the law of $(U_1, U_2, ...)$ satisfies a LDP with speed θ and rate function $\sum_{i=1}^\infty I_1(u_i)$. The effective domain of $\sum_{i=1}^\infty I_1(u_i)$ is

$$\mathcal{C} = \{\mathbf{u} \in E^\infty : u_i < 1, \sum_{i=1}^\infty u_i < \infty\}$$

and

$$\sum_{i=1}^\infty I_1(u_i) = \log \frac{1}{\Pi_{i=1}^\infty (1-u_i)} \text{ on } \mathcal{C}. \tag{2.6}$$

Since the map G is continuous, it follows from contraction principle and Lemma 2.1 that the family $\{GEM(\theta,\alpha) : \theta > 0, 0 < \alpha < 1\}$ satisfies a LDP on $\mathcal{E}$ with speed θ and rate function

$$\inf\{\sum_{i=1}^\infty I_1(u_i) : u_1 = x_1, u_2(1-u_1) = x_2, ...\}. \tag{2.7}$$

For each $1 \le n \le +\infty$,

$$(1-u_1)\cdots(1-u_n) = 1 - \sum_{i=1}^n x_i. \tag{2.8}$$

Hence if $\sum_{i=1}^n x_i = 1$ for some finite n, then one of $u_1, ..., u_n$ is one, and $\mathbf{u}$ is not in $\mathcal{C}$. If $\sum_{i=1}^n x_i < 1$ for all finite n and $\sum_{i=1}^\infty x_i = 1$, then from (2.8)

$$\lim_{n\to\infty} \sum_{i=1}^n \log(1-u_i) = -\infty, \tag{2.9}$$

which implies that $\sum_{i=1}^\infty u_i = \infty$. Thus the inverse of the set $\{(x_1, ...) \in \mathcal{E} : \sum_{i=1}^\infty x_i = 1\}$ under G is disjoint with $\mathcal{C}$. Hence the rate function in (2.7) is the same as $S(x_1, x_2, ...)$.

□

3 LDP for Two-Parameter Poisson-Dirichlet Distribution

In this section, we establish the LDP for $PD(\alpha,\theta)$. Recall that $\mathbf{P}(\alpha,\theta) = (P_1(\alpha,\theta), P_2(\alpha,\theta)...,)$ is the random probability measure with law $PD(\alpha,\theta)$. When $\alpha = 0$, we will write

$$\mathbf{P}(\theta) = \mathbf{P}(0,\theta) = (P_1(\theta), P_2(\theta)...,).$$

3.1 Perman's Formula

For $0 \le \alpha < 1$ and any constant $C > 0, \beta > 0$, let

$$h(x) = \alpha C x^{-(\alpha+1)}, x > 0,$$

and

$$c_{\alpha,\beta} = \frac{\Gamma(\beta+1)(C\Gamma(1-\alpha))^{\beta/\alpha}}{\Gamma(\beta/\alpha+1)}.$$

Let $\psi(t)$ be a density function over $(0,\infty)$ such that for all $\beta > -\alpha$

$$\int_0^\infty t^{-\beta}\psi(t)dt = \frac{1}{c_{\alpha,\beta}}. \tag{3.10}$$

Let $\{\tau_s : s \geq 0\}$ be the stable subordinator with index α. Then $\psi(t)$ is the density function of τ_1. (cf. page 892 in (20).)
Set

$$\psi_1(t,p) = h(tp)t\psi(t\bar{p}), t>0, 0<p<1, \bar{p}=1-p \tag{3.11}$$

$$\psi_{n+1}(t,p) = \begin{cases} h(tp)t\int_{p/\bar{p}}^1 \psi_n(t\bar{p},q)dq, & p \leq 1/(n+1) \\ 0, & \text{else.} \end{cases} \tag{3.12}$$

Then the following result is found in (16)(see also (20)).

Lemma 3.1. (Perman's Formula) *For each $k \geq 1$, let $f(p_1,...,p_k)$ denote the joint density function of $(P_1(\alpha,\theta),...,P_k(\alpha,\theta))$. Then*

$$f(p_1,...,p_k) = c_{\alpha,\theta}\int_0^\infty t^{-\theta}g_k(t,p_1,...,p_k)dt, \tag{3.13}$$

where for $k\geq 2, t>0,\ 0<p_k<\cdots<p_1, \sum_{i=1}^k p_i<1$, and $\hat{p}_k = 1-p_1-\cdots-p_{k-1}$,

$$g_k(t,p_1,...,p_k) = \frac{t^{k-1}h(tp_1)\cdots h(tp_{k-1})}{\hat{p}_k}g_1(t\hat{p}_k,\frac{p_k}{\hat{p}_k}) \tag{3.14}$$

and

$$g_1(t,p) = \sum_{n=1}^\infty (-1)^{n+1}\psi_n(t,p). \tag{3.15}$$

3.2 LDP for $PD(\alpha,\theta)$

We will prove the LDP for $PD(\alpha,\theta)$ by first establishing the LDP for $P_1(\alpha,\theta)$ and $(P_1(\alpha,\theta),...,P_k(\alpha,\theta))$ for any $k\geq 2$.

Lemma 3.2. *The family of the laws of $P_1(\alpha,\theta)$ satisfies a LDP on E with speed θ and rate function $I_1(p)$ given in (2.3).*

Proof: It follows from the GEM representation that

$$E[e^{\lambda\theta U_1}] \leq E[e^{\lambda\theta P_1(\alpha,\theta)}] \text{ for } \lambda\geq 0;$$
$$E[e^{\lambda\theta U_1}] \geq E[e^{\lambda\theta P_1(\alpha,\theta)}] \text{ for } \lambda < 0.$$

On the other hand, from the representation in Proposition 22 of (20) we obtain that

$$E[e^{\lambda\theta P_1(\alpha,\theta)}] \leq E[e^{\lambda\theta P_1(\theta)}] \text{ for } \lambda\geq 0;$$
$$E[e^{\lambda\theta P_1(\alpha,\theta)}] \geq E[e^{\lambda\theta P_1(\theta)}] \text{ for } \lambda < 0.$$

Since both the laws of U_1 and $P_1(\theta)$ satisfy LDPs with speed θ and rate function $I_1(\cdot)$, we conclude from Lemma 2.4 of (4) that the law of $P_1(\alpha,\theta)$ satisfies a LDP with speed θ and rate function $I_1(\cdot)$.

□

Lemma 3.3. *For each fixed $k \geq 2$, let*

$$\nabla_k = \{\mathbf{p} = (p_1, ..., p_k) : 0 \leq p_k \leq \cdots \leq p_1, \sum_{i=1}^{k} p_i \leq 1\},$$

and $\mathcal{P}_{\theta,k}$ be law of $(P_1(\alpha,\theta), ..., P_k(\alpha,\theta))$. Then the family $\{\mathcal{P}_{\theta,k} : \theta > 0\}$ satisfies a LDP on ∇_k with speed θ and rate function

$$I_k(p_1, ..., p_k) = \begin{cases} \log \frac{1}{1-\sum_{i=1}^{k} p_i}, & \sum_{i=1}^{k} p_i < 1 \\ \infty, & \textit{else.} \end{cases}$$

Proof: For $k \geq 2, t > 0$, $0 < p_k < \cdots < p_1, \sum_{i=1}^{k} p_i < 1$, it follows from (3.13), (3.14) and (3.15) that

$$\begin{aligned} f(p_1, ..., p_k) &= \frac{c_{\alpha,\theta}(\alpha C)^{k-1}}{\hat{p}_k (p_1 \cdots p_{k-1})^{\alpha+1}} \int_0^\infty t^{-[\theta+(k-1)\alpha]} g_1(t\hat{p}_k, \frac{p_k}{\hat{p}_k}) d\,t \\ &= \frac{c_{\alpha,\theta}(\alpha C)^{k-1} \hat{p}_k^{\theta+(k-1)\alpha}}{\hat{p}_k^2 (p_1 \cdots p_{k-1})^{\alpha+1}} \int_0^\infty s^{-[\theta+(k-1)\alpha]} g_1(s, \frac{p_k}{\hat{p}_k}) d\,s \qquad (3.16) \\ &= \frac{c_{\alpha,\theta}(\alpha C)^{k-1} \hat{p}_k^{\theta+(k-1)\alpha}}{\hat{p}_k^2 (p_1 \cdots p_{k-1})^{\alpha+1}} \sum_{n=1}^{\infty} (-1)^{n+1} \int_0^\infty s^{-[\theta+(k-1)\alpha]} \psi_n(s, \frac{p_k}{\hat{p}_k}) d\,s \end{aligned}$$

Set

$$\phi_n(u) = \int_0^\infty s^{-[\theta+(k-1)\alpha]} \psi_n(s, u) d\,s.$$

Then

$$\begin{aligned} \phi_1(u) &= \int_0^\infty s^{-[\theta+(k-1)\alpha]} s h(su) \psi(s\bar{u}) d\,s \\ &= (\alpha C) u^{-(\alpha+1)} \int_0^\infty s^{-[\theta+k\alpha]} \psi(s\bar{u}) d\,s \\ &= (\alpha C) u^{-(\alpha+1)} \bar{u}^{\theta+k\alpha-1} \int_0^\infty s^{-[\theta+k\alpha]} \psi(s) d\,s \qquad (3.17) \\ &= (\alpha C) u^{-(\alpha+1)} \bar{u}^{\theta+k\alpha-1} \frac{1}{c_{\alpha,\theta+k\alpha}}, \end{aligned}$$

where $\bar{u} = 1 - u$. For any $n \geq 1$, and $u \leq \frac{1}{n+1}$, it follows from (3.12) that

$$
\begin{aligned}
&\psi_{n+1}(t,u) \qquad\qquad (3.18)\\
&= t\,h(tu)\int_{u/\bar u}^{1}\psi_n(t\bar u,u_1)d\,u_1\\
&= t\,h(tu)\int_{u/\bar u}^{1}[\chi_{\{u_1\le 1/n\}}(t\bar u)h(t\bar u u_1)\int_{u_1/\bar u_1}^{1}\psi_{n-1}(t\bar u\bar u_1,u_2)d\,u_2]d\,u_1\\
&= t\,h(tu)\int_{u/\bar u}^{1/(n-1)}\cdots\int_{u_{n-1}/\bar u_{n-1}}^{1}[(t\bar u h(t\bar u u_1))\cdots(t\bar u\bar u_1\cdots\bar u_{n-2}h(t\bar u\bar u_1\cdots\bar u_{n-2}u_{n-1}))]\\
&\quad\times[\chi_{\{u_1\le 1/n\}}\cdots\chi_{\{u_{n-1}\le 1/2\}}]\psi_1(t\bar u\bar u_1\cdots\bar u_{n-2}\bar u_{n-1},u_n)d\,u_n d\,u_{n-1}\cdots d\,u_1\\
&= t\,h(tu)\int_{u/\bar u}^{1}\cdots\int_{u_{n-1}/\bar u_{n-1}}^{1}[(t\bar u h(t\bar u u_1))\cdots(t\bar u\bar u_1\cdots\bar u_{n-2}h(t\bar u\bar u_1\cdots\bar u_{n-2}u_{n-1}))]\\
&\quad\times[\chi_{\{u_1\le 1/n\}}\cdots\chi_{\{u_{n-1}\le 1/2\}}][(t\bar u\bar u_1\cdots\bar u_{n-2}\bar u_{n-1})h(t\bar u\bar u_1\cdots\bar u_{n-2}\bar u_{n-1}u_n)]\\
&\quad\times\psi(t\bar u\bar u_1\cdots\bar u_{n-2}\bar u_{n-1}\bar u_n)d\,u_n d\,u_{n-1}\cdots d\,u_1\\
&= (\alpha C)^{n+1}t^{-(n+1)\alpha}\int_{u/\bar u}^{1}\int_{u_1/\bar u_1}^{1}\cdots\int_{u_{n-1}/\bar u_{n-1}}^{1}[\chi_{\{u_1\le 1/n\}}\cdots\chi_{\{u_{n-1}\le 1/2\}}]d\,u_n d\,u_{n-1}\cdots d\,u_1\\
&\quad\times\{(uu_1\cdots u_n)^{-(\alpha+1)}\bar u^{-n\alpha}\bar u_1^{-(n-1)\alpha}\cdots\bar u_{n-1}^{-\alpha}\psi(t\bar u\bar u_1\cdots\bar u_{n-2}\bar u_{n-1}\bar u_n)\}.
\end{aligned}
$$

Integrating over t we get

$$
\begin{aligned}
\phi_{n+1}(u) &= (\alpha C)^{n+1}\int_{u/\bar u}^{1}\int_{u_1/\bar u_1}^{1}\cdots\int_{u_{n-1}/\bar u_{n-1}}^{1}[\chi_{\{u_1\le 1/n\}}\cdots\chi_{\{u_{n-1}\le 1/2\}}]d\,u_n d\,u_{n-1}\cdots d\,u_1\\
&\quad\times\{(uu_1\cdots u_n)^{-(\alpha+1)}\bar u^{-n\alpha}\bar u_1^{-(n-1)\alpha}\cdots\bar u_{n-1}^{-\alpha}\\
&\quad\times[\int_0^{\infty}t^{-[\theta+(k+n)\alpha]}\psi(t\bar u\bar u_1\cdots\bar u_{n-2}\bar u_{n-1}\bar u_n)d\,t]\} \qquad (3.19)\\
&= (\alpha C)^{n+1}u^{-(\alpha+1)}\bar u^{(\theta+k\alpha-1)}\frac{1}{c_{\alpha,\theta+(n+k)\alpha}}A_n(\alpha,\theta)(u)\\
&= (\alpha C)^{n}\phi_1(u)\frac{c_{\alpha,\theta+k\alpha}}{c_{\alpha,\theta+(n+k)\alpha}}A_n(\alpha,\theta)(u),
\end{aligned}
$$

where

$$
\begin{aligned}
A_n(\alpha,\theta)(u) &= \int_{u/\bar u}^{1}\int_{u_1/\bar u_1}^{1}\cdots\int_{u_{n-1}/\bar u_{n-1}}^{1}[\chi_{\{u_1\le 1/n\}}\cdots\chi_{\{u_{n-1}\le 1/2\}}]d\,u_n d\,u_{n-1}\cdots d\,u_1\\
&\quad\times\{(u_1\cdots u_n)^{-(\alpha+1)}\bar u_1^{\theta+(k+1)\alpha-1}\cdots\bar u_n^{\theta+(k+n)\alpha-1}\}.
\end{aligned}
$$

Let

$$D=\{(u_1,...,u_n):u_1\in[u/\bar u,1/n],...,u_{n-1}\in[u_{n-2}/\bar u_{n-1},1/2],u_n\in[u_{n-1}/\bar u_n,1]\}.$$

By definition, $A_n(\alpha,\theta)(u)=0$ for $u=\frac{1}{n+1}$.

For $0<u<1/(n+1)$, the Lebesgue measure of D is strictly positive. It follows by direct calculation that

$$lim_{\theta\to\infty}\frac{1}{\theta}\log A_n(\alpha,\theta)(u) \le \operatorname{ess\,sup}\{\sum_{r=1}^{n}\log(1-u_r) : (u_1,...,u_n)\in D\} < 0. \tag{3.20}$$

On the other hand, by Stirling's formula,

$$\lim_{\theta\to\infty}\frac{1}{\theta}\log\frac{c_{\alpha,\theta}}{c_{\alpha,\theta+(n+k)\alpha}}(u) = 0. \tag{3.21}$$

Thus for $n \ge 1$, and $0 < u \le 1/(n+1)$,

$$\lim_{\theta\to\infty}\frac{c_{\alpha,\theta}}{c_{\alpha,\theta+(n+k)\alpha}}A_n(\alpha,\theta) = 0. \tag{3.22}$$

Let

$$\nabla_k^{\circ} = \{(p_1,...,p_k)\in\nabla_k : p_k > 0, \sum_{i=1}^{k} p_i < 1\}.$$

Now for each fixed $(p_1,...,p_k)$ in ∇_k°, set

$$m = \max\{j \ge 1 : p_k/\hat{p}_k \le 1/j\},$$

and $A_0(\alpha,\theta)(u) = 1$ for any $0 < u < 1$.
Then it follows from (3.16), (3.19) and (3.22) that

$$\begin{aligned} f(p_1,...,p_k) &= \frac{c_{\alpha,\theta}(\alpha C)^{k-1}\hat{p}_k^{\theta+(k-1)\alpha}}{\hat{p}_k^2(p_1\cdots p_{k-1})^{\alpha+1}}\phi_1(\frac{p_k}{\hat{p}_k})\sum_{n=1}^{m}(-1)^{n+1}\frac{(\alpha C)^{n-1}c_{\alpha,\theta+k\alpha}}{c_{\alpha,\theta+(n+k-1)\alpha}}A_{n-1}(\alpha,\theta)(\frac{p_k}{\hat{p}_k}) \\ &= \frac{(\alpha C)^k(\hat{p}_{k+1})^{\theta+k\alpha-1}}{(p_1\cdots p_k)^{1+\alpha}}\sum_{n=1}^{m}(-1)^{n+1}\frac{(\alpha C)^{n-1}c_{\alpha,\theta}}{c_{\alpha,\theta+(n+k-1)\alpha}}A_{n-1}(\alpha,\theta)(\frac{p_k}{\hat{p}_k}) \qquad (3.23) \\ &= (\hat{p}_{k+1})^{\theta+k\alpha-1}\frac{(\alpha C)^k}{(p_1\cdots p_k)^{1+\alpha}}[1+o(1)], \end{aligned}$$

which implies that

$$\lim_{\theta\to\infty}\frac{1}{\theta}\log f(p_1,...,p_k) = -\log\frac{1}{1-\sum_{i=1}^{k}p_i} \text{ on } \nabla_k^{\circ}. \tag{3.24}$$

Introduce a metric d_k on ∇_k such that for any $\mathbf{p},\mathbf{q}$ in ∇_k

$$d_k(\mathbf{p},\mathbf{q}) = \sum_{i=1}^{k}|p_i - q_i|.$$

For any $\delta > 0$, set

$$\begin{aligned} V_\delta(\mathbf{p}) &= \{\mathbf{q}\in\nabla_k : d_k(\mathbf{p},\mathbf{q}) < \delta\} \\ \bar{V}_\delta(\mathbf{p}) &= \{\mathbf{q}\in\nabla_k : d_k(\mathbf{p},\mathbf{q}) \le \delta\}. \end{aligned}$$

For every $\mathbf{p} \in \nabla_k^\circ$, one can choose δ small enough such that $V_\delta(\mathbf{p}) \subset \bar{V}_\delta(\mathbf{p}) \subset \nabla_k^\circ$. Let μ denote the Lebesgue measure on ∇_k. Then by Jensen' inequality and (3.24),

$$\begin{aligned}\lim_{\theta\to\infty}\frac{1}{\theta}\log\mathcal{P}_{\theta,k}\{V_\delta(\mathbf{p})\} &= \lim_{\theta\to\infty}\frac{1}{\theta}\log\frac{\mu(V_\delta(\mathbf{p}))}{\mu(V_\delta(\mathbf{p}))}\int_{V_\delta(\mathbf{p})} f(\mathbf{q})\mu(d\,\mathbf{q}) \\ &\geq -\frac{1}{\mu(V_\delta(\mathbf{p}))}\int_{V_\delta(\mathbf{p})} I_k(\mathbf{q})\mu(d\,\mathbf{q}).\end{aligned} \tag{3.25}$$

Letting δ approach zero and using the continuity of $I_k(\cdot)$ at $\mathbf{p}$, one gets

$$\lim_{\delta\to 0}\lim_{\theta\to\infty}\frac{1}{\theta}\log\mathcal{P}_{\theta,k}\{V_\delta(\mathbf{p})\} \geq -I_k(\mathbf{p}). \tag{3.26}$$

Since the family $\{\mathcal{P}_{\theta,k} : \theta > 0\}$ is exponentially tight, a partial LDP holds ((21)). Let J be any rate function associated with certain subsequence of $\{\mathcal{P}_{\theta,k} : \theta > 0\}$. Then it follows from (3.26) that for any $\mathbf{p}$ in ∇_k°

$$J(\mathbf{p}) \leq I_k(\mathbf{p}). \tag{3.27}$$

Because of the continuity of I_k and the lower semi-continuity of J, (3.27) holds on ∇_k.

On the other hand for any $\mathbf{p}$ in ∇_k°,

$$\begin{aligned}\lim_{\theta\to\infty}\frac{1}{\theta}\log\mathcal{P}_{\theta,k}\{\bar{V}_\delta(\mathbf{p})\} &= \lim_{\theta\to\infty}\frac{1}{\theta}\log\int_{\bar{V}_\delta(\mathbf{p})} f(\mathbf{q})\mu(d\,\mathbf{q}) \\ &\leq \lim_{\theta\to\infty}\frac{1}{\theta}\log f(\mathbf{q}_\delta) \\ &= -I_k(\mathbf{q}_\delta),\end{aligned} \tag{3.28}$$

where $\mathbf{q}_\delta$ is in ∇_k° such that

$$f(\mathbf{q}_\delta) = \sup\{f(\mathbf{q}) : \mathbf{q} \in \bar{V}_\delta(\mathbf{p})\}.$$

The existence of such $\mathbf{q}_\delta$ is due to the continuity of f over ∇_k°. Letting δ approach zero, one has

$$\lim_{\delta\to 0}\lim_{\theta\to\infty}\frac{1}{\theta}\log\mathcal{P}_{\theta,k}\{\bar{V}_\delta(\mathbf{p})\} \leq -I_k(\mathbf{p}). \tag{3.29}$$

Next consider the case that $\mathbf{p}$ is such that $p_k > 0, \sum_{i=1}^k p_i = 1$. Then $p_k/\hat{p}_k = 1$. For small enough δ, we have

$$q_k/\hat{q}_k > 1/2 \ \text{ for } \ \mathbf{q} \in \bar{V}_\delta(\mathbf{p}).$$

Thus $A_n(\alpha,\theta)(u) = 0$ for all $n \geq 1$ on $\bar{V}_\delta(\mathbf{p})$ and it follows from (3.23) that

$$\begin{aligned}\lim_{\theta\to\infty}\frac{1}{\theta}\log\mathcal{P}_{\theta,k}\{\bar{V}_\delta(\mathbf{p})\} &= \lim_{\theta\to\infty}\frac{1}{\theta}\log\int_{\bar{V}_\delta(\mathbf{p})} f(\mathbf{q})\mu(d\,\mathbf{q}) \\ &\leq \lim_{\theta\to\infty}\frac{1}{\theta}\log\int_{\bar{V}_\delta(\mathbf{p})} (\hat{q}_{k+1})^{\theta+k\alpha-1}\frac{(\alpha C)^k}{(q_1\cdots q_k)^{1+\alpha}}\mu(d\,\mathbf{q}), \\ &\leq \log(1-a_\delta),\end{aligned} \tag{3.30}$$

where a_δ is such that

$$a_\delta = \inf\{\sum_{i=1}^{k} q_i : \mathbf{q} \in \bar{V}_\delta(\mathbf{p})\} < 1.$$

Letting δ go to zero, one gets

$$\lim_{\delta \to 0} \lim_{\theta \to \infty} \frac{1}{\theta} \log \mathcal{P}_{\theta,k}\{\bar{V}_\delta(\mathbf{p})\} \leq -I_k(\mathbf{p}). \tag{3.31}$$

The only case remains is when there is a $l \leq k$ such that $p_l = 0$. The upper bound in this case is obtained by focusing on a lower dimensional space of the positive coordinates.

Thus we have shown that for every $\mathbf{p}$ in ∇_k

$$\lim_{\delta \to 0} \lim_{\theta \to \infty} \frac{1}{\theta} \log \mathcal{P}_{\theta,k}\{V_\delta(\mathbf{p})\} = \lim_{\delta \to 0} \lim_{\theta \to \infty} \frac{1}{\theta} \log \mathcal{P}_{\theta,k}\{\bar{V}_\delta(\mathbf{p})\} = -I_k(\mathbf{p}), \tag{3.32}$$

which combined with the exponential tightness implies the result.

□

Let

$$\bar{\nabla} = \{(p_1, p_2, ...) : p_1 \geq p_2 \geq \cdots \geq 0, \sum_{i=1}^{\infty} p_i \leq 1\}.$$

and for notational simplicity we use $\mathcal{P}_\theta$ to denote the law of $\mathbf{P}(\alpha, \theta)$ on $\bar{\nabla}$ in the next theorem. Then we have

Theorem 3.4. *The family $\{\mathcal{P}_\theta : \theta > 0\}$ satisfies a LDP with speed θ and rate function*

$$I(\mathbf{p}) = \begin{cases} \log \frac{1}{1 - \sum_{i=1}^{\infty} p_i}, & (p_1, p_2, ...) \in \bar{\nabla},\ \sum_{i=1}^{\infty} p_i < 1 \\ \infty, & \textit{else.} \end{cases} \tag{3.33}$$

Proof: Because $\bar{\nabla}$ is compact, the family $\{\mathcal{P}_\theta : \theta > 0\}$ is exponentially tight. It is thus sufficient to verify the local LDP ((21)). The topology on $\bar{\nabla}$ can be generated by the following metric

$$d(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^{\infty} \frac{|p_i - q_i|}{2^i},$$

where $\mathbf{p} = (p_1, p_2, ...), \mathbf{q} = (q_1, q_2, ...)$. For any fixed $\delta > 0$, let $B(\mathbf{p}, \delta)$ and $\bar{B}(\mathbf{p}, \delta)$ denote the respective open and closed balls centered at $\mathbf{p}$ with radius $\delta > 0$. Set $n_\delta = 2 + [\log_2(1/\delta)]$ where $[x]$ denotes the integer part of x. Set

$$V_{n_\delta}(\mathbf{p}; \delta/2) = \{(q_1, q_2, ...) \in \bar{\nabla} : |q_i - p_i| < \delta/2, i = 1, ..., n_\delta\},$$
$$V((p_1, ..., p_{n_\delta}); \delta/2) = \{(q_1, ..., q_{n_\delta}) \in \nabla_{n_\delta} : |q_i - p_i| < \delta/2, i = 1, ..., n_\delta\}.$$

Then we have

$$V_{n_\delta}(\mathbf{p}; \delta/2) \subset B(\mathbf{p}, \delta).$$

By lemma 3.3 and the fact that

$$\mathcal{P}_\theta\{V_{n_\delta}(\mathbf{p}; \delta/2)\} = \mathcal{P}_{\theta, n_\delta}\{V((p_1, ..., p_{n_\delta}); \delta/2)\},$$

we get that

$$\liminf_{\theta\to\infty}\frac{1}{\theta}\log\mathcal{P}_\theta\{B(\mathbf{p},\delta)\} \ge \liminf_{\theta\to\infty}\frac{1}{\theta}\log\mathcal{P}_{\theta,n_\delta}\{V((p_1,...,p_{n_\delta});\delta/2)\} \tag{3.34}$$
$$\ge -I_{n_\delta}(p_1,...,p_{n_\delta}) \ge -I(\mathbf{p}).$$

On the other hand for any fixed $n\ge 1, \delta_1>0$, let

$$U_n(\mathbf{p};\delta_1) = \{(q_1,q_2,...)\in\bar{\nabla} : |q_i-p_i|\le\delta_1, i=1,...,n\},$$
$$U((p_1,...,p_n);\delta_1) = \{(q_1,...,q_n)\in\nabla_n : |q_i-p_i|\le\delta_1, i=1,...,n\}$$

Then we have

$$\mathcal{P}_\theta\{U_n(\mathbf{p};\delta_1)\} = \mathcal{P}_{\theta,n}\{U((p_1,...,p_n);\delta_1)\},$$

and, for δ small enough,

$$\bar{B}(\mathbf{p},\delta)\subset U_n(\mathbf{p};\delta_1),$$

which implies that

$$\lim_{\delta\to 0}\limsup_{\theta\to\infty}\frac{1}{\theta}\log\mathcal{P}_\theta\{\bar{B}(\mathbf{p},\delta)\} \le \limsup_{\theta\to\infty}\frac{1}{\theta}\log\mathcal{P}_{\theta,n}\{U((p_1,...,p_n),\delta_1)\} \tag{3.35}$$
$$\le -\inf\{I_n(q_1,...,q_n) : (q_1,...,q_n)\in U((p_1,...,p_n),\delta_1)\}.$$

Letting δ_1 go to zero, and then n go to infinity, we get

$$\lim_{\delta\to 0}\limsup_{\theta\to\infty}\frac{1}{\theta}\log\mathcal{P}_\theta\{\bar{B}(\mathbf{p},\delta)\}\le -I(\mathbf{p}), \tag{3.36}$$

which combined with (3.34) implies the result.

□

4 LDP for Two-Parameter Dirichlet Process

Let $M_1(E)$ denote the space of all probability measures on E equipped with the weak topology. For any diffusive ν in $M_1(E)$ with support E, let $\xi_1,\xi_2,..$ be independent and identically distributed with common distribution ν. Let $\Xi_{\theta,\alpha,\nu}$ be the two-parameter Dirichlet process defined in (1.2).

Let $\{\sigma(t) : t\ge 0, \sigma_0 = 0\}$ be a subordinator with Lévy measure $x^{-(1+\alpha)}e^{-x}dx$, $x>0$, and $\{\tau(t) : t\ge 0, \tau_0=0\}$ be a gamma subordinator that is independent of $\{\sigma_t : t\ge 0, \sigma_0 = 0\}$ and has Lévy measure $x^{-1}e^{-x}dx$, $x>0$.

Lemma 4.1. (Pitman and Yor) *Let*

$$\gamma(\alpha,\theta) = \frac{\alpha\tau(\frac{\theta}{\alpha})}{\Gamma(1-\alpha)}. \tag{4.37}$$

For each $n\ge 1$, and each partition $0<t_1<\cdots<t_n=1$ of E, let $A_i=(t_{i-1},t_i]$ for $i=2,...,n$, $A_1=[0,t_1]$, and $a_j=\nu(A_j)$. Set

$$Y_{\alpha,\theta}(t) = \sigma(\gamma(\alpha,\theta)t), t\ge 0.$$

Then the distribution of $(\Xi_{\theta,\alpha,\nu}(A_1),...,\Xi_{\theta,\alpha,\nu}(A_n))$ *is the same as the distribution of*

$$(\frac{Y_{\alpha,\theta}(a_1)}{Y_{\alpha,\theta}(1)},...,\frac{Y_{\alpha,\theta}(\sum_{j=1}^{n}a_j)-Y_{\alpha,\theta}(\sum_{j=1}^{n-1}a_j)}{Y_{\alpha,\theta}(1)}).$$

Proof: Proposition 21 in (20) gives the subordinator representation for $PD(\alpha,\theta)$. The lemma follows from this representation and the construction outlined on page 254 in (19).

□

Let

$$Z_{\alpha,\theta}(t)=\frac{Y_{\alpha,\theta}(t)}{\theta},$$
$$\mathbf{Z}_{\alpha,\theta}(t_1,...,t_n)=(Z_{\alpha,\theta}(a_1),...,Z_{\alpha,\theta}(\sum_{j=1}^{n}a_j)-Z_{\alpha,\theta}(\sum_{j=1}^{n-1}a_j)).$$

By direct calculation, one has

$$\begin{aligned}\varphi(\lambda) &= \log\phi(\lambda)=\log E[e^{\lambda\sigma(1)}]=\int_0^\infty(e^{\lambda x}-1)x^{-(\alpha+1)}e^{-x}\,dx \\ &= \begin{cases}\frac{\Gamma(1-\alpha)}{\alpha}[1-(1-\lambda)^\alpha], & \lambda\le 1\\ \infty, & \text{else.}\end{cases}\end{aligned} \tag{4.38}$$

and

$$\begin{aligned}L(\lambda) &= \lim_{\theta\to\infty}\frac{1}{\theta}\log E[e^{\lambda\tau(\theta)}] \\ &= \begin{cases}\log(\frac{1}{1-\lambda}), & \lambda<1\\ \infty, & \text{else.}\end{cases}\end{aligned} \tag{4.39}$$

For any real numbers $\lambda_1,...\lambda_n$, let $\vec{\lambda}=(\lambda_1,...,\lambda_n)$. Then by direct calculation

$$\begin{aligned}\frac{1}{\theta}\log E[\exp\{\theta\langle\vec{\lambda},\mathbf{Z}_{\alpha,\theta}(t_1,...,t_n)\rangle\}] &= \frac{1}{\theta}\log E[\Pi_{i=1}^n(E_{\tau(\theta/\alpha)}[\exp\{\lambda_i\sigma(1)\}]^{\frac{\alpha a_i}{\Gamma(1-\alpha)}\tau(\theta/\alpha)})] \\ &= \frac{1}{\theta}\log E[\exp\{(\sum_{i=1}^n\frac{\alpha\nu(A_i)}{\Gamma(1-\alpha)}\varphi(\lambda_i))\tau(\frac{\theta}{\alpha})\}] \qquad (4.40)\\ &\to \Lambda(\lambda_1,...,\lambda_n)=\frac{1}{\alpha}L(\frac{\alpha}{\Gamma(1-\alpha)}\sum_{i=1}^n\nu(A_i)\varphi(\lambda_i)).\end{aligned}$$

For $(y_1,...,y_n)$ in R_+^n, set

$$\begin{aligned}J_{t_1,..,t_n}(y_1,...,y_n) &= \sup_{\lambda_1,...,\lambda_n}\{\sum_{i=1}^n\lambda_iy_i-\Lambda(\lambda_1,...,\lambda_n)\} \qquad (4.41)\\ &= \sup_{\lambda_1,...,\lambda_n\in(-\infty,1]^n}\{\sum_{i=1}^n\lambda_iy_i+\frac{1}{\alpha}\log[\sum_{i=1}^n\nu(A_i)(1-\lambda_i)^\alpha]\}\end{aligned}$$

Theorem 4.2. *The family of the laws of $\mathbf{Z}_{\alpha,\theta}(t_1,...,t_n)$ on space R_+^n satisfies a LDP with speed θ and rate function (4.41).*

Proof: First note that both function φ and function L are essentially smooth. Let

$$\mathcal{D}_\Lambda = \{(\lambda_1,...,\lambda_n) : \Lambda(\lambda_1,...,\lambda_n) < \infty\},\ \mathcal{D}_\Lambda^\circ = \text{interior of } \mathcal{D}_\Lambda.$$

It follows from (4.39) and (4.40) that

$$\mathcal{D}_\Lambda = \{(\lambda_1,...,\lambda_n) : \sum_{i=1}^n \nu(A_i)\frac{\alpha}{\Gamma(1-\alpha)}\varphi(\lambda_i) < 1\}.$$

The fact that ν has support E implies that $\nu(A_i) > 0$ for $i = 1,...,n$, and

$$\begin{aligned}
\mathcal{D}_\Lambda &= \{(\lambda_1,...,\lambda_n) : \sum_{i=1}^n \nu(A_i)[1-(1-\lambda_i)^\alpha] < 1\} \\
&= \{(\lambda_1,...,\lambda_n) : \lambda_i \le 1, i = 1,...,n\} \setminus \{(1,...,1)\}, \\
\mathcal{D}_\Lambda^\circ &= \{(\lambda_1,...,\lambda_n) : \lambda_i < 1, i = 1,...,n\}.
\end{aligned}$$

Clearly the function Λ is differentiable on $\mathcal{D}_\Lambda^\circ$ and

$$grad(\Lambda)(\lambda_1,...,\lambda_n) = \frac{1}{\Gamma(1-\alpha)}L'(\frac{\alpha}{\Gamma(1-\alpha)}\sum_{i=1}^n \nu(A_i)\varphi(\lambda_i))(\nu(A_1)\varphi'(\lambda_1),...,\nu(A_n)\varphi'(\lambda_n)).$$

A sequence $\vec{\lambda}_m$ approaches the boundary of $\mathcal{D}_\Lambda^\circ$ from inside implies that at least one coordinate sequence approaches one. Since the interior of $\{\lambda : \varphi(\lambda) < \infty\}$ is $(-\infty,1)$ and φ is essentially smooth, it follows that Λ is steep and thus essentially smooth. The theorem then follows from Gärtner-Ellis theorem ((5)).

□

For $(y_1,...,y_n)$ in R_+^n and $(x_1,...,x_n)$ in E^n, define

$$F(y_1,..,y_n) = \begin{cases} \frac{1}{\sum_{k=1}^n y_k}(y_1,...,y_n), & \sum_{k=1}^n y_k > 0 \\ (0,...,0), & (y_1,...,y_n) = (0,...,0) \end{cases}$$

and

$$I_{t_1,..,t_n}(x_1,...,x_n) = \inf\{J_{t_1,..,t_n}(y_1,...,y_n) : F(y_1,...,y_n) = (x_1,...,x_n)\}. \tag{4.42}$$

Clearly $I_{t_1,...,t_n}(x_1,...,x_n) = +\infty$ if $\sum_{k=1}^n x_k$ is not one. For $(x_1,...,x_n)$ satisfying $\sum_{k=1}^n x_k = 1$, we have

$$\begin{aligned}
I_{t_1,..,t_n}(x_1,...,x_n) &= \inf\{J_{t_1,..,t_n}(ax_1,...,ax_n) : a=\sum_{k=1}^n y_k>0\} \qquad (4.43)\\
&= \inf\{\sup_{(\lambda_1,...,\lambda_n)\in(-\infty,1]^n}\{a\sum_{i=1}^n \lambda_i x_i+\frac{1}{\alpha}\log[\sum_{i=1}^n \nu(A_i)(1-\lambda_i)^\alpha\} : a>0\}\}\\
&= \inf\{\sup_{(\lambda_1,...,\lambda_n)\in(-\infty,1]^n}\{a-\log a\\
&\quad -\sum_{i=1}^n a(1-\lambda_i)x_i+\frac{1}{\alpha}\log[\sum_{i=1}^n \nu(A_i)[a(1-\lambda_i)]^\alpha]\} : a>0\}\\
&= \inf\{a-\log a : a>0\}+\sup_{(\gamma_1,...,\gamma_n)\in R_+^n}\{\frac{1}{\alpha}\log[\sum_{i=1}^n \nu(A_i)\gamma_i^\alpha]-\sum_{i=1}^n \gamma_i x_i\}\\
&= \sup_{(\gamma_1,...,\gamma_n)\in R_+^n}\{\frac{1}{\alpha}\log[\sum_{i=1}^n \nu(A_i)\gamma_i^\alpha]+1-\sum_{i=1}^n \gamma_i x_i\}.
\end{aligned}$$

Theorem 4.3. *The family of the laws of* $(\Xi_{\theta,\alpha,\nu}(A_1),...,\Xi_{\theta,\alpha,\nu}(A_n))$ *on space* E^n *satisfies a LDP with speed* θ *and rate function*

$$I_{t_1,..,t_n}(x_1,...,x_n)=\begin{cases}\sup_{(\gamma_1,...,\gamma_n)\in R_+^n}\{\frac{1}{\alpha}\log[\sum_{i=1}^n \nu(A_i)\gamma_i^\alpha] \\ \qquad +1-\sum_{i=1}^n \gamma_i x_i\}, & \sum_{k=1}^n x_k=1\\ \infty, & \textit{else}\end{cases} \qquad (4.44)$$

Proof: Since $J_{t_1,...,t_n}(0,...,0)=\infty$, the function F is thus continuous on the effective domain of $J_{t_1,..,t_n}$. The theorem then follows from Lemma 4.1 and the contraction principle (remark (c) of Theorem 4.2.1 in (5)).

□

Remark. Let $\Delta_n=\{(x_1,...,x_n)\in E^n : \sum_{i=1}^n x_i=1\}$. Then the result in Theorem 4.3 holds with E^n being replaced by Δ_n.

Let $B_b(E)$ and $C_b(E)$ denote the sets of bounded measurable functions, and bounded continuous functions on E respectively. For each μ in $M_1(E)$, set

$$\begin{aligned}
I^\alpha(\mu) &= \sup_{f\geq 0, f\in C_b(E)}\{\frac{1}{\alpha}\log(\int (f(x))^\alpha \nu(dx))+1-\int f(x)\mu(dx)\}, \qquad (4.45)\\
I^0(\mu) &= \sup_{f\geq 0, f\in B_b(E)}\{\int \log f(x)\nu(dx)+1-\int f(x)\mu(dx)\}, \qquad (4.46)\\
H(\nu|\mu) &= \sup_{g\in C_b(E)}\{\int g(x)\nu(dx)-\log\int e^{g(x)}\mu(dx)\} \qquad (4.47)\\
&= \sup_{g\in B_b(E)}\{\int g(x)\nu(dx)-\log\int e^{g(x)}\mu(dx)\},
\end{aligned}$$

where $H(\nu|\mu)$ is the relative entropy of ν with respect to μ.

Lemma 4.4. *For any* μ *in* $M_1(E)$,

$$I^\alpha(\mu)=\sup\{I_{t_1,..,t_n}(\mu(A_1),...,\mu(A_n)) : 0<t_1<t_2<\cdots<t_n=1; n=1,2,...\}. \qquad (4.48)$$

Proof: It follows from Tietze's continuous extension theorem and Luzin's Theorem that we can replace $C_b(E)$ with $B_b(E)$ in the definition of I^α. This implies that

$$I^\alpha(\mu) \geq \sup\{I_{t_1,..,t_n}(\mu(A_1),...,\mu(A_n)) : 0 < t_1 < t_2 < \cdots < t_n = 1; n = 1,2,...\}.$$

On the other hand, for each nonnegative f in $C_b(E)$, let

$$t_i = \frac{i}{n}, \gamma_i = f(t_i), i = 1,...,n.$$

Then

$$\frac{1}{\alpha}\log(\int (f(x))^\alpha \nu(dx)) - \int f(x)\mu(dx) = \lim_{n\to\infty}\{\frac{1}{\alpha}\log[\sum_{i=1}^n \nu(A_i)\gamma_i^\alpha] - \sum_{i=1}^n \gamma_i\mu(A_i)\}$$
$$\leq \sup\{I_{t_1,..,t_n}(\mu(A_1),...,\mu(A_n)) : 0 < t_1 < t_2 < \cdots < t_n = 1; n = 1,2,...\},$$

which implies

$$I^\alpha(\mu) \leq \sup\{I_{t_1,..,t_n}(\mu(A_1),...,\mu(A_n)) : 0 < t_1 < t_2 < \cdots < t_n = 1; n = 1,2,...\}.$$

□

Remarks. It follows from the proof of Lemma 4.4 that the supremum in (4.48) can be taken over all partitions with $t_1,...,t_{n-1}$ being the continuity points of μ. By monotonically approximating nonnegative $f(x)$ with strictly positive functions from above, it follows from the monotone convergence theorem that the supremum in both (4.45) and (4.46) can be taken over strictly positive bounded functions, i.e.,

$$I^\alpha(\mu) = \sup_{f>0, f\in C_b(E)}\{\frac{1}{\alpha}\log(\int (f(x))^\alpha\nu(dx)) + 1 - \int f(x)\mu(dx)\}, \tag{4.49}$$

$$I^0(\mu) = \sup_{f>0, f\in B_b(E)}\{\int \log f(x)\nu(dx) + 1 - \int f(x)\mu(dx)\}. \tag{4.50}$$

Lemma 4.5.

$$I^0(\mu) = H(\nu|\mu) \tag{4.51}$$

Proof: If ν is not absolutely continuous with respect to μ, then $H(\nu|\mu) = +\infty$. Let A be a set such that $\mu(A) = 0, \nu(A) > 0$ and define

$$f_m(x) = \begin{cases} m, & x \in A \\ 1, & \text{else} \end{cases}$$

Then

$$I^0(\mu) \geq \nu(A)\log m \to \infty \text{ as } m \to \infty.$$

Next we assume $\nu \ll \mu$ and denote $\frac{d\nu}{d\mu}(x)$ by $h(x)$. By definition,

$$H(\nu|\mu) = \int h(x)\log(h(x))\mu(dx).$$

Choosing $f_M(x) = h(x) \wedge M$ in the definition of I^0. Since the function $x \log x$ is bounded below for non-negative x, applying the monotone convergence theorem on $\{x : h(x) \geq e^{-1}\}$, one gets

$$I^0(\mu) \geq \lim_{M\to\infty} \int f_M(x) \log f_M(x)\nu(dx) - \log \int f_M(x)\mu(dx) = H(\nu|\mu). \tag{4.52}$$

On the other hand, it follows by letting $f(x) = e^{g(x)}$ in the definition of $H(\nu|\mu)$ that

$$H(\nu|\mu) = \sup_{f>0, f\in B_b(E)} \{\int \log f(x)\nu(dx) - \log \int f(x)\mu(dx)\}.$$

Since

$$\int f(x)\mu(dx) - 1 \geq \log \int f(x)\mu(dx),$$

we get that

$$H(\nu|\mu) \geq I^0(\mu) \tag{4.53}$$

which combined with (4.52) implies (4.51).

□

Lemma 4.6. *For any μ in $M_1(E)$, $0 \leq \alpha_1 < \alpha_2 < 1$,*

$$I^{\alpha_2}(\mu) \geq I^{\alpha_1}(\mu), \tag{4.54}$$

and for any α in $(0,1)$, one can find μ in $M_1(E)$ satisfying $\nu \ll \mu$ such that

$$I^{\alpha}(\mu) > I^0(\mu). \tag{4.55}$$

Proof: By Hölder's inequality, for any $0 < \alpha_1 < \alpha_2 < 1$,

$$\frac{1}{\alpha_1} \log(\int (f(x))^{\alpha_1}\nu(dx)) \leq \frac{1}{\alpha_2} \log(\int (f(x))^{\alpha_2}\nu(dx)), \tag{4.56}$$

and the inequality becomes strict if $f(x)$ is not constant almost surely under ν. Hence $I^\alpha(\mu)$ is non-decreasing in α over $(0,1)$. It follows from the concavity of $\log x$ that

$$\frac{1}{\alpha} \log(\int (f(x))^{\alpha}\nu(dx)) \geq \int \log f(x)\nu(dx),$$

which implies that $I^\alpha(\mu) \geq I^0(\mu)$ for $\alpha > 0$.
Next choose μ in $M_1(E)$ such that $\nu \ll \mu$ and $\frac{d\nu}{d\mu}(x)$ is not a constant with ν probability one, then $I^\alpha(\mu) > I^{\alpha/2}(\mu) \geq I^0(\mu)$ for $\alpha > 0$.

□

We now ready to prove the main result of this section.

Theorem 4.7. *The family of the laws of $\Xi_{\theta,\alpha,\nu}$ on space $M_1(E)$ satisfies a LDP with speed θ and rate function $I^\alpha(\cdot)$.*

Proof: Let $\{f_j(x) : j = 1, 2, ...\}$ be a countable dense subset of $C_b(E)$ in the supremum norm. The set $\{f_j(x) : j = 1, 2, ...\}$ is clearly convergence determining on $M_1(E)$. Let $|f_j| = \sup_{x \in E} |f_j(x)|$ and

$$\mathcal{C} = \{g_j(x) = \frac{f_j(x)}{|f_j| \vee 1} : j = 1, ...\}.$$

Then $\mathcal{C}$ is also convergence determining.

For any ω, μ in $M_1(E)$, define

$$\rho(\omega, \mu) = \sum_{j=1}^{\infty} \frac{1}{2^j} |\langle \omega, g_j \rangle - \langle \mu, g_j \rangle|. \tag{4.57}$$

Then ρ is a metric on $M_1(E)$ and generates the weak topology.

For any $\delta > 0, \mu \in M_1(E)$, let

$$B(\mu, \delta) = \{\omega \in M_1(E) : \rho(\omega, \mu) < \delta\}, \ \ \overline{B}(\mu, \delta) = \{\omega \in M_1(E) : \rho(\omega, \mu) \leq \delta\}.$$

Since $M_1(E)$ is compact, the family of the laws of $\Xi_{\theta,\alpha}$ is exponentially tight. It thus suffices to show that

$$\lim_{\delta \to 0} \liminf_{\theta \to \infty} \frac{1}{\theta} \log P\{B(\mu, \delta)\} = \lim_{\delta \to 0} \limsup_{\theta \to \infty} \frac{1}{\theta} \log P\{\overline{B}(\mu, \delta)\} = -I^{\alpha}(\mu). \tag{4.58}$$

Choose m large enough, one gets

$$\{\omega \in M_1(E) : |\langle \omega, g_j \rangle - \langle \mu, g_j \rangle| < \delta/2 : j = 1, \cdots, m\} \subset B(\nu, \delta). \tag{4.59}$$

Choose a partition $t_1, \cdots, t_n$ such that

$$\sup\{|g_j(x) - g_j(y)| : x, y \in A_i, i = 1, \cdots, n; j = 1, \cdots, m\} < \delta/8.$$

Choosing $0 < \delta_1 < \frac{\delta}{4n}$, and define

$$V_{t_1,\cdots,t_n}(\mu, \delta_1) = \{(y_1, ..., y_n) \in \Delta_n : |y_i - \mu(A_i)| < \delta_1, i = 1, \cdots, n\}.$$

For any ω in $M_1(E)$, let

$$F(\omega) = (\omega(A_1), ..., \omega(A_n)).$$

If $F(\omega) \in V_{t_1,\cdots,t_n}(\mu, \delta_1)$, then for $j = 1, ..., m$

$$\begin{aligned} |\langle \omega, g_j \rangle - \langle \mu, g_j \rangle| &= |\sum_{i=1}^{n} \int_{A_i} g_j(x)(\omega(dx) - \mu(dx))| \\ &< \frac{\delta}{4} + n\delta_1 < \delta/2, \end{aligned}$$

which implies that

$$\{\omega \in M_1(E) : F(\omega) \in V_{t_1,\cdots,t_n}(\mu, \delta_1)\} \subset \{\omega \in M_1(E) : |\langle \omega, g_j \rangle - \langle \mu, g_j \rangle| < \delta/2 : j = 1, \cdots, m\}.$$

This combined with (4.59) implies that

$$\{\omega \in M_1(E) : F(\omega) \in V_{t_1,\cdots,t_n}(\mu, \delta_1)\} \subset B(\mu, \delta). \tag{4.60}$$

Since $V_{t_1,\cdots,t_n}(\mu, \delta_1)$ is open in Δ_n, it follows from Theorem 4.3 that

$$\begin{aligned}
&\lim_{\delta\to 0}\liminf_{\theta\to\infty}\frac{1}{\theta}\log P\{B(\mu,\delta)\} \qquad (4.61)\\
&\quad\geq \lim_{\delta\to 0}\liminf_{\theta\to\infty}\frac{1}{\theta}\log P\{\omega\in M_1(E): F(\omega)\in V_{t_1,\cdots,t_n}(\mu,\delta_1)\}\\
&\quad= \lim_{\delta\to 0}\liminf_{\theta\to\infty}\frac{1}{\theta}\log P\{(\Xi_{\theta,\alpha,\nu}(A_1),...,\Xi_{\theta,\alpha,\nu}(A_n))\in V_{t_1,\cdots,t_n}(\mu,\delta_1)\}\\
&\quad\geq -I_{t_1,..,t_n}(\mu(A_1),...,\mu(A_n)) \geq -I^{\alpha}(\mu).
\end{aligned}$$

Next we will focus on partitions $t_1, ..., t_n$ such that $t_1, ..., t_{n-1}$ are continuity points of μ. We denote the collection of all such partitions by $\mathcal{P}_\mu$. This implies that $F(\omega)$ is continuous at μ. Hence for any $\delta_2 > 0$, one can choose $\delta > 0$ small enough such that

$$\overline{B}(\mu, \delta) \subset F^{-1}\{V_{t_1,\cdots,t_k}(\mu, \delta_2)\}.$$

Let

$$\overline{V}_{t_1,\cdots,t_k}(\nu, \delta_2) = \{(y_1, ..., y_n) \in \Delta_n : |y_i - \mu(A_i)| \leq \delta_2, i = 1, \cdots, n-1\}.$$

Then we have

$$\begin{aligned}
&\lim_{\delta\to 0}\limsup_{\theta\to\infty}\frac{1}{\theta}\log P\{\overline{B}(\mu,\delta)\} \qquad (4.62)\\
&\quad\leq \limsup_{\theta\to\infty}\frac{1}{\theta}\log P\{(\Xi_{\theta,\alpha,\nu}(A_1),...,\Xi_{\theta,\alpha,\nu}(A_n))\in \overline{V}_{t_1,\cdots,t_n}(\mu,\delta_2)\}.
\end{aligned}$$

By letting δ_2 go to zero and applying Theorem 4.3 again, one gets

$$\lim_{\delta\to 0}\limsup_{\theta\to\infty}\frac{1}{\theta}\log P\{\overline{B}(\mu,\delta)\} \leq -I_{t_1,\cdots,t_n}(\mu(A_1), ..., \mu(A_n)). \tag{4.63}$$

Finally, taking supremum over $\mathcal{P}_\mu$ and taking into account the remark after Lemma 4.4, one gets

$$\lim_{\delta\to 0}\limsup_{\gamma\to 0}\frac{1}{\theta}\log P\{\overline{B}(\mu,\delta)\} \leq -I^{\alpha}(\mu), \tag{4.64}$$

which, combined with (4.61), implies the theorem.

□

5 Further Comments

Our results show that the LDPs for $GEM(\theta, \alpha)$ and $PD(\alpha, \theta)$ have the same rate function. Since $GEM(\theta, \alpha)$ and $PD(\alpha, \theta)$ differs only by the ordering, one would expect to derive the LDP for one from the LDP for the other. Unfortunately the ordering operation is not continuous and it is not easy to establish an exponential approximation. The LDPs for $GEM(\theta, \alpha)$ and $PD(\alpha, \theta)$ also have the same rate function as the LDPs for $GEM(\theta)$ and $PD(\theta)$. Thus α does not play a role in these LDPs. This is mainly due to the topology used. It will be interesting to investigate the possibility of seeing the role of α through establishing the corresponding LDPs on a stronger topology.

The LDPs for $\Xi_{\theta,\alpha,\nu}$ and $\Xi_{\theta,\nu}$ have respective rate functions $I^{\alpha}(\cdot)$ and $I^{0}(\cdot)$. Both $\Xi_{\theta,\alpha,\nu}$ and $\Xi_{\theta,\nu}$ converge to ν for large θ. When θ becomes large, each $P_i(\theta, \alpha)$ is more likely to be small. The introduction of positive α plays a similar role. Thus the mass in $\Xi_{\theta,\alpha,\nu}$ spreads more evenly than the mass in $\Xi_{\theta,\nu}$. Intuitively $\Xi_{\theta,\alpha,\nu}$ is "closer" to ν than $\Xi_{\theta,\nu}$. This observation is made rigorous through the fact that $I^{\alpha}(\cdot)$ can be strictly bigger than $I^{0}(\cdot)$. The monotonicity of $I^{\alpha}(\cdot)$ in α shows that α can be used to measure the relative "closeness" to ν among all $\Xi_{\theta,\alpha,\nu}$ for large θ.

The process $Y_{\alpha,\theta}(t)$ is a process with exchangeable increments. One could try to establish a general LDP result for processes with exchangeable increments and derive the result in Section 4 through contraction principle. The proofs here illustrate most of the procedures needed for pursuing such a general result from which the LDP for $\Xi_{\theta,\alpha,\nu}$ follows.

Acknowledgments The author would like to thank Jim Pitman for clarification of Lemma 4.1, and the referees for their insightful comments and suggestions.

References

[1] M. Aoki (2006). Thermodynamic limit of macroeconomic or financial models: one- and two-parameter Poisson-Dirichlet models. Preprint.

[2] M.A. Carlton (1999). Applications of the two-parameter Poisson-Dirichlet distribution. Unpublished Ph.D. thesis, Dept. of Statistics, University of California, Los Angeles.

[3] D. Dawson and S. Feng (2001). Large deviations for the Fleming-Viot process with neutral mutation and selection, II. *Stoch. Proc. Appl.* **92**, 131–162. MR1815182

[4] D. Dawson and S. Feng (2006). Asymptotic behavior of Poisson-Dirichlet distribution for large mutation rate. *Ann. Appl. Probab.*, Vol.16, No. 2, 562-582. MR2244425

[5] A. Dembo and O. Zeitouni. *Large deviations techniques and applications.* Second edition. Applications of Mathematics, **38**. Springer-Verlag, New York, 1998. MR1619036

[6] B. Derrida (1997). From random walks to spin glasses. *Physica D*, **107**, 186-198. MR1491962

[7] W.J. Ewens. *Mathematical Population Genetics, Vol. I.* Springer-Verlag, New York, 2004. MR0554616

[8] S. Feng and F.M. Hoppe (1998). Large deviation principles for some random combinatorial structures in population genetics and Brownian motion. *Ann. Appl. Probab.*, Vol. 8, No. 4, 975–994. MR1661315

[9] T.S. Ferguson (1973). A Baysian analysis of some nonparametric problems. *Ann. Stat.*, **1**:209–230. MR0350949

[10] R.A. Fisher, A.S. Corbet and C.B. Williams (1943). The relation between the number of species and the number of individuals in a random sample of an animal population. *J. Animal Ecology*, **12**:42–58.

[11] J.H. Gillespie (1999). The role of population size in molecular evolution. *Theoretical Population Biology*, **55** 145–156.

[12] F.M. Hoppe (1984). Pólya-like urns and the Ewens sampling formula. *Journal of Mathematical Biology*, 20, 91–94. MR0758915

[13] P. Joyce, S.M. Krone, and T.G. Kurtz (2003). When can one detect overdominant selection in the infinite-alleles model? *Ann. Appl. Probab.*, **13**, No. 1, 181–212. MR1951997

[14] J.C.F. Kingman (1975). Random discrete distributions. *J. Roy. Statist. Soc. B,* **37**, 1-22. MR0368264

[15] Lynch, J. and Sethuraman, J. (1987). Large deviations for processes with independent increments. *Ann. Probab.*, **15**:610–627. MR0885133

[16] M. Perman (1993). Order statistics for jumps of normalised subordinators. *Stoch. Proc. Appl.,* **46**, 267-281. MR1226412

[17] J. Pitman (1996). Random discrete distributions invariant under size-biased permutation. *Adv. Appl. Probab.,* **28**, 525-539. MR1387889

[18] J. Pitman (1995). Exchangeable and partially exchangeable random partitions. *Prob. Theory Rel. Fields*, **102**, 145–158. MR1337249

[19] J. Pitman (1996). Some developments of the Blackwell-MacQueen urn scheme. In *Statistics, Probability and Game Theory. Papers in Honor of David Blackwell* (T.S. Ferguson, L.S. Shapley, and J.B. MacQueen, eds.), 245–267. IMS Lecture Notes-Monograph series, Vol.**30**. MR1481784

[20] J. Pitman and M. Yor (1997). The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Ann. Probab.*,Vol. 25, No. 2, 855-900. MR1434129

[21] A.A. Puhalskii. On functional principle of large deviations. In V.Sazonov and T. Shervashidze, editors. *New Trends in Probability and Statistics.* pages 198-218. VSP Moks'las, Moskva, 1991. MR1200917

[22] L.A. Shepp and S.P. Lloyd (1966). Ordered cycle length in a random permutation. *Trans. Amer. Math. Soc. Probab.*,Vol. 121, 340-357. MR0195117

[23] A.M. Vershik (1986). The asymptotic distribution of factorizations of natural numbers into prime divisors. *Soviet Math. Dokl.*, 34:57-61.