The Annals of Applied Probability
2001, Vol. 00, No. 00, 000–000

GAUSSIAN LIMITS ASSOCIATED WITH THE POISSON–DIRICHLET DISTRIBUTION AND THE EWENS SAMPLING FORMULA

BY PAUL JOYCE,[1] STEPHEN M. KRONE[1] AND THOMAS G. KURTZ[2]

University of Idaho, University of Idaho and University of Wisconsin

In this paper we consider large θ approximations for the stationary distribution of the neutral infinite alleles model as described by the the Poisson–Dirichlet distribution with parameter θ. We prove a variety of Gaussian limit theorems for functions of the population frequencies as the mutation rate θ goes to infinity. In particular, we show that if a sample of size n is drawn from a population described by the Poisson–Dirichlet distribution, then the conditional probability of a particular sample configuration is asymptotically normal with mean and variance determined by the Ewens sampling formula. The asymptotic normality of the conditional sampling distribution is somewhat surprising since it is a fairly complicated function of the population frequencies. Along the way, we also prove an invariance principle giving weak convergence at the process level for powers of the size-biased allele frequencies.

1. Preliminaries and main results. Gillespie (1999) showed that population size plays a significant role in molecular evolutionary dynamics. He considered various models in which the per individual mutation rate u is held fixed and the population size N is increased. In the case of the neutral model of evolution this is equivalent to considering the limiting distributions when the scaled mutation rate $\theta = 4Nu$ goes to infinity. From the point of view of genetic diffusions, Gillespie's simulations suggest that large θ limits are both natural and interesting. In particular, they shed light on the difficulty of detecting certain types of selection.

For many population genetics models, statistical inference is complicated by the fact that both the population and sample are random. Given a particular evolutionary scenario, the composition of the current population is just one of many possible that could have arisen under the same evolutionary forces. When one considers a sample drawn from such a population, there are two probability distributions of interest. Firstly, one can consider the conditional probability of observing a particular sample given the current population, and secondly, one can ask for the unconditional probability of observing a particular sample. The unconditional distribution of the sample is calculated by averaging over all possible population frequencies. It is interesting to note, for the neutral infinite alleles model with parameter θ, that the conditional

Received August 2000; revised March 2001.

[1]Supported in part by NSF Grants DMS-96-26764 and DMS-00-72198.
[2]Supported in part by NSF Grant DMS-99-71571.

AMS 1991 *subject classifications*. Primary 62J70, 92D10; secondary 60F05.
Key words and phrases. Poisson–Dirichlet, GEM, neutral infinite alleles model, sampling distribution, Ewens sampling formula, Gaussian process, mutation rate.

probability distribution can be approximated by the unconditional probability distribution if θ is large. In fact we establish a normal limit theorem that can be used to assess the error in this approximation.

In this paper, we deal exclusively with the neutral infinite alleles model. However, our original motivation came from trying to understand some surprising non-neutral simulations in Gillespie (1999) in which the selection intensity and the mutation rate get large together. The results of the present paper form the theoretical foundation for comparing the neutral model to models with selection when θ is assumed to be large. This will be treated in a forthcoming paper [cf. Joyce, Krone, and Kurtz (2001)].

We begin by establishing some notation and describing the model in more detail. The *neutral infinite alleles model* is a diffusion process which arises from a Wright–Fisher model in which each mutation gives rise to a completely new allele. The reader is referred to Ethier and Kurtz (1986) or Ewens (1979) for an account of this model. To get a nontrivial stationary measure, it turns out that one must order the allele frequencies in some way. Thus, we consider the ordered infinite simplex

$$\nabla \equiv \left\{(x_1, x_2, \dots) : x_1 \geq x_2 \geq \cdots \geq 0,\ \sum_{i=1}^{\infty} x_i = 1\right\}.$$

For the neutral infinite alleles model with *mutation rate* θ, the stationary distribution, μ, of the allele frequencies $\mathbf{X} = (X_1, X_2, \dots) \in \nabla$ in descending order is given by the Poisson–Dirichlet distribution with parameter θ [cf. Kingman (1977)]. We will abbreviate this with the notation $\mu \sim \mathscr{PD}(\theta)$ or $\mathbf{X} \sim \mathscr{PD}(\theta)$.

Our main results deal with the partition structure of a sample drawn from a random population with relative frequencies $\mathbf{X} \sim \mathscr{PD}(\theta)$. Let $\mathbf{a} = (a_1, a_2, \dots, a_n)$ denote an *allelic partition* of a sample of n genes; that is, a_i $(i = 1, \dots, n)$ gives the number of distinct alleles each of which is represented exactly i times in the sample. Clearly, $a_i \geq 0$, $\sum_{i=1}^n i a_i = n$, and $\sum_{i=1}^n a_i \equiv k$ gives the number of distinct alleles in the sample. We write $\mathbf{A}_n = (A_1, A_2, \dots, A_n)$ for the random partition obtained by independently sampling n genes according to the relative frequencies of the different alleles in the population. If the allele proportions are given by $\mathbf{x} = (x_1, x_2, \dots) \in \nabla$, then the conditional sampling probability $P(\mathbf{A}_n = \mathbf{a}|\mathbf{X} = \mathbf{x})$ is given by the *multinomial sampling function* (cf. Kingman (1977)):

$$P(\mathbf{A}_n = \mathbf{a}|\mathbf{X} = \mathbf{x}) = \phi_{\mathbf{a}}(\mathbf{x}) \equiv \frac{n!}{\prod_{i=1}^n (i!)^{a_i}} \sum_{\nu \in \mathscr{A}_{\mathbf{a}}} x_1^{\nu_1} x_2^{\nu_2} \cdots, \tag{1}$$

where $\mathscr{A}_{\mathbf{a}} \equiv \{\nu = (\nu_1, \nu_2, \dots) \in \mathbb{Z}_+ \times \mathbb{Z}_+ \times \cdots :\ \#(i : \nu_i = j) = a_j\}$ represents the set of allele counts which are consistent with the partition $\mathbf{a}$. In the neutral infinite alleles model, the stationary probability of observing the allelic partition $\mathbf{a}$ in a sample of size n,

$$P(\mathbf{A}_n = \mathbf{a}) = \int_{\nabla} \phi_{\mathbf{a}}(\mathbf{x})\, \mu(d\mathbf{x}) = E\left(\phi_{\mathbf{a}}(\mathbf{X})\right), \tag{2}$$

can be expressed by the *Ewens sampling formula*

$$P(\mathbf{A}_n = \mathbf{a}) = \mathrm{ESF}(\theta, \mathbf{a}) \equiv \frac{n!}{\theta_{(n)}} \prod_{j=1}^{n} \left(\frac{\theta}{j}\right)^{a_j} \frac{1}{a_j!}, \tag{3}$$

where $\theta_{(n)} = \theta(\theta+1)\cdots(\theta+n-1)$.

Note that $E\phi_{\mathbf{a}}(\mathbf{X}) = P(\mathbf{A}_n = \mathbf{a}) \to 0$ as $\theta \to \infty$, if $\mathbf{a} \neq (n, 0, \dots, 0)$. We will show that $\phi_{\mathbf{a}}(\mathbf{X}) \Rightarrow 0$ and discuss the rate of convergence by deriving a normal limit theorem. The main goal of this paper is to show that, when $\mathbf{X} \sim \mathscr{P}D(\theta)$, the *conditional sampling function* $\phi_{\mathbf{a}}(\mathbf{X})$ is asymptotically normal for large θ (and fixed n), with asymptotic mean given by the $\mathrm{ESF}(\theta, \mathbf{a})$ and asymptotic variance given by $(\mathrm{ESF}(\theta, \mathbf{a}))^2/\theta$. A precise statement appears in Theorem 2. To establish the asymptotic distribution of $\phi_{\mathbf{a}}(\mathbf{X})$ we will first consider some simple partitions. For a subsample of size $m \leq n$, the probability that all m individuals in the sample are the same type is given by $\sum_{i=1}^{\infty} X_i^m$. We refer to this sum as the *mth population moment*. It can be thought of as a measure of the mth-order population homozygosity. In the case of $m = 2$ the above formula is referred to as the population homozygosity and plays a significant role in population genetic analysis. Note that the Ewens sampling formula is consistent in the sense that the sampling distribution for a sample of m individuals is the same as that of a subsample of m from a larger sample of size n [cf. Joyce (1998)].

We will show that the asymptotics of $\phi_{\mathbf{a}}(\mathbf{X})$ are determined by the asymptotics of the population moments. To this end, define $\mathbf{Z}_\theta = (Z_{2,\theta}, Z_{3,\theta}, \dots) \in \mathbb{R}^\infty$ where

$$Z_{m,\theta} = \sqrt{\theta}\left(\frac{\theta^{m-1}\sum_{i=1}^{\infty} X_i^m - (m-1)!}{(m-1)!}\right), \qquad m \geq 2, \tag{4}$$

is the scaled mth population moment.

To see that $Z_{m,\theta}$ is an appropriate transformation of $\sum_{i=1}^{\infty} X_i^m$, note that, according to the Ewens sampling formula (3),

$$E\left(\sum_{i=1}^{\infty} X_i^m\right) = P\left(\mathbf{A}_m = (0, 0, \dots, 1)\right) = \frac{m!\,\theta}{\theta_{(m)} m},$$

when $\mathbf{X} \sim \mathscr{P}D(\theta)$. Therefore, as $\theta \to \infty$, we get $E(\theta^{m-1}\sum_{i=1}^{\infty} X_i^m) \to (m-1)!$ which suggests

$$\frac{\theta^{m-1}\sum_{i=1}^{\infty} X_i^m}{(m-1)!} - 1 \Rightarrow 0.$$

To obtain a nontrivial limit for the above equation we scale by $\sqrt{\theta}$.

THEOREM 1. *Suppose* $\mathbf{X} \sim \mathscr{P}D(\theta)$ *and let* $\mathbf{Z}_\theta = (Z_{2,\theta}, Z_{3,\theta}, \dots) \in \mathbb{R}^\infty$ *where* $Z_{m,\theta}$ *is defined by* (4). *Then*

$$\mathbf{Z}_\theta \Rightarrow \mathbf{Z} \qquad \text{as} \quad \theta \to \infty,$$

where $\mathbf{Z}$ *is an* $\mathbb{R}^\infty$*-valued random vector. The joint distribution of any finite number of components of* $\mathbf{Z}$ *has a multivariate normal distribution with mean* $\mathbf{0}$ *and*

$$\text{Cov}(Z_i, Z_j) = \frac{(i+j-1)! - i!j!}{(i-1)!(j-1)!}, \qquad i, j = 2, 3, \ldots. \tag{5}$$

REMARK. Asymptotic normality of the homozygosity ($\sum X_i^2$) was observed Griffiths (1979) for a similar K-allele model, in the limit as K and θ go to infinity together. He, in effect, proved that $Z_{2,\theta}$ converges in distribution to Z_2, which is a special case of Theorem 1. For other asymptotic results involving the neutral model see Watterson and Guess (1977).

The proof of Theorem 1 will require several preliminary results and appears in the next section. In our quest to prove that $\phi_{\mathbf{a}}(\mathbf{X})$, when properly scaled, is asymptotically normal, we will be aided by the following lemma which demonstrates that the large θ asymptotics of $\phi_{\mathbf{a}}(\mathbf{X})$ are determined by those of $\mathbf{Z}_\theta$. This will allow us to focus on simple allelic sub-partitions of the form $\mathbf{a} = (0, \ldots, 0, m)$; that is, all the information we need is contained in the mth population moments.

LEMMA 1. *Suppose* $\mathbf{X} \sim \mathscr{P}D(\theta)$ *and let* $\mathbf{a}$ *be a partition corresponding to a sample of size* n *drawn from a population with ordered frequencies* $\mathbf{X}$*. If* $a_1 \neq n$*, then*

$$\sum_{i=2}^{n} a_i Z_{i,\theta} = \sqrt{\theta}\left(\frac{\phi_{\mathbf{a}}(\mathbf{X}) - ESF(\theta, \mathbf{a})}{ESF(\theta, \mathbf{a})}\right) + R_{\mathbf{a}}(\theta)$$

where $\theta^{1/2-\varepsilon} R_{\mathbf{a}}(\theta) \to 0$ *as* $\theta \to \infty$ *for all* $\varepsilon > 0$.

This lemma will be proved in Section 2.

In the next theorem, we combine these results to show how the scaled conditional sampling function converges to the projection of a multivariate normal onto the allelic partition. First note that, as the mutation rate $\theta \to \infty$, with sample size n held fixed, all the probability moves to the allelic partition $\mathbf{a}_1 \equiv (n, 0, 0 \ldots, 0)$ with $a_1 = n$, giving all different alleles in the sample. This is intuitively clear from the infinite alleles assumption. It also follows quickly from (3):

$$P(\mathbf{A}_n = \mathbf{a}_1) = E\phi_{\mathbf{a}_1}(\mathbf{X}) = \frac{n!}{\theta_{(n)}} \frac{\theta^n}{n!} \to 1, \tag{6}$$

and for $\mathbf{a} \neq \mathbf{a}_1$,

$$E\phi_{\mathbf{a}}(\mathbf{X}) \to 0, \tag{7}$$

as $\theta \to \infty$.

Therefore, to establish the appropriate scaling factor required for asymptotic normality of $\phi_{\mathbf{a}}(\mathbf{X})$, we will need to consider the cases $\mathbf{a} \neq \mathbf{a}_1$ and $\mathbf{a} = \mathbf{a}_1$ separately. The following theorem is our main result.

THEOREM 2. *Suppose* $\mathbf{X} \sim \mathscr{P}D(\theta)$ *and let* $\phi_{\mathbf{a}}$ *be the multinomial sampling function corresponding to a sample of size* n, *given by* (1). *Let* $(Z_2, Z_3, \dots, Z_n)$ *have a multivariate normal distribution with mean* $\mathbf{0}$ *and covariance given by* (5). *Then the following limits hold as* $\theta \to \infty$:

(i) *If* $\mathbf{a} = (a_1, \dots, a_n)$ *is an allelic partition such that* $a_1 \neq n$, *then*

$$\sqrt{\theta}\left(\frac{\phi_{\mathbf{a}}(\mathbf{X}) - ESF(\theta, \mathbf{a})}{ESF(\theta, \mathbf{a})}\right) \Rightarrow \sum_{i=2}^{n} a_i Z_i. \tag{8}$$

(ii) *If* $\mathbf{a}_1 = (n, 0, \dots, 0)$ *and* $\mathbf{a}_2 = (n-2, 1, 0, \dots, 0)$ *are the allelic partitions with* n *singletons and* $n-2$ *singletons, respectively, then*

$$\sqrt{\theta}\left(\frac{\phi_{\mathbf{a}_1}(\mathbf{X}) - ESF(\theta, \mathbf{a}_1)}{ESF(\theta, \mathbf{a}_2)}\right) \Rightarrow Z_2 \sim \mathrm{N}(0, 2). \tag{9}$$

Theorem 2 follows immediately from Lemma 1 and Theorem 1 provided at least two individuals in the sample have the same type; that is $a_1 \neq n$. The case where $a_1 = n$ requires a separate argument and appears in the next section after the proof of Theorem 1.

2. Proofs.

The Poisson–Dirichlet distribution is difficult to deal with directly, so the first step in proving the above results is to express this stationary distribution in terms of the *GEM distribution* [cf. Donnelly and Joyce (1989)]. To define the GEM distribution, suppose that

$$U_1, U_2, \dots \quad \text{are i.i.d. Beta}\,(1, \theta) \text{ random variables} \tag{10}$$

with common density $f(x;\theta) = \theta(1-x)^{\theta-1}$, $x \in [0,1]$. Then set

$$\begin{aligned} V_{1,\theta} &= U_1, \\ V_{k,\theta} &= (1-U_1)(1-U_2)\cdots(1-U_{k-1})U_k \qquad \text{for} \quad k \geq 2. \end{aligned} \tag{11}$$

The random point $(V_{1,\theta}, V_{2,\theta}, \dots)$ in the unordered infinite simplex

$$\Delta \equiv \left\{(x_1, x_2, \dots) : x_i \geq 0 \qquad \forall i,\ \sum_{i=1}^{\infty} x_i = 1\right\}$$

is said to have the GEM(θ) distribution [cf. Donnelly and Joyce (1989)]. This gives the neutral stationary allele proportions in "size-biased order" rather than descending order. The construction in (11) is sometimes referred to as "stick breaking." Note that $\nabla \subseteq \Delta$ and so $\mu \sim \mathscr{P}D(\theta)$ can be thought of as a probability measure on Δ which puts all its mass on ∇. If $\rho : \Delta \to \nabla$ is the "descending order map," then we have the relationship $\rho(\mathbf{V}_\theta) \sim \mathscr{P}D(\theta)$ if $\mathbf{V}_\theta = (V_{1,\theta}, V_{2,\theta}, \dots) \sim \mathrm{GEM}(\theta)$ [cf. Donnelly and Joyce (1989)]. In particular, since the multinomial sampling function is invariant under reordering of the variables, we can write

$$\phi_{\mathbf{a}}(\mathbf{X}) = \phi_{\mathbf{a}}(\mathbf{V}_\theta). \tag{12}$$

PROOF OF THEOREM 1. To guide the reader through the technical material that follows, a few words on our strategy are in order. Our ultimate goal is to establish asymptotic normality for $\sum_{i=1}^{\infty} V_{i,\theta}^m$ [which is equal to $\sum_{i=1}^{\infty} X_i^m$ by (12)] as $\theta \to \infty$. This is a four step process.

1. We truncate the series and consider $\sum_{i=1}^{[\theta t]} V_{i,\theta}^m$.
2. Since $V_{1,\theta}^m = U_1^m$, and $V_{i,\theta}^m = U_i^m \exp\{m \sum_{j=1}^{i-1} \log(1-U_j)\}$ when $i \geq 2$, we next establish limit theorems for $\sum_{i=1}^{[\theta t]} U_i^m$ and $\sum_{i=1}^{[\theta t]} \log(1-U_i)$, properly scaled. This is done in Lemmas 2 and 3.
3. Writing $\sum_{i=1}^{[\theta t]} V_{i,\theta}^m$ as a stochastic integral involving U_i^m and $\log(1-U_i)$, we then use a theorem of Kurtz and Protter (1991) to establish a limit result for $\sum_{i=1}^{[\theta t]} V_{i,\theta}^m$, again with suitable scaling. This is done in Theorem 3.
4. We extend the results of Theorem 3, by taking the limit as $t \to \infty$. This extension requires that the tails of the series go to zero uniformly in θ, and this is established in Lemma 4 and then used to prove Theorem 1.

Recall that $U_1, U_2, \ldots$ are i.i.d. Beta$(1, \theta)$ random variables. In light of statement 2 above, we define a sequence of $\mathbb{R}^n$-valued random vectors $\mathbf{J}_\theta(1)$, $\mathbf{J}_\theta(2), \ldots$ by

$$\begin{aligned} \mathbf{J}_\theta(i) &\equiv (J_{1,\theta}(i), J_{2,\theta}(i), \ldots, J_{n,\theta}(i)) \\ &= \left(-\theta \log(1-U_i), (\theta U_i)^2, \ldots, (\theta U_i)^n\right). \end{aligned} \tag{13}$$

The factor of θ which has been introduced is natural, as will be seen below; for example, all the moments of θU_i are easy to compute. All of the limit theorems that follow will involve functions of the components of the $\mathbf{J}_\theta(i)$'s. The next two lemmas establish limit theorems for the $\mathbb{R}^n$-valued processes

$$\mathbf{Y}_\theta(t) = (Y_{1,\theta}(t), \ldots, Y_{n,\theta}(t)) = \frac{1}{\theta} \sum_{i=1}^{[\theta t]} \mathbf{J}_\theta(i), \tag{14}$$

$$\mathbf{W}_\theta(t) = (W_{1,\theta}(t), \ldots, W_{n,\theta}(t)) = \frac{1}{\sqrt{\theta}} \sum_{i=1}^{[\theta t]} (\mathbf{J}_\theta(i) - E(\mathbf{J}_\theta(i))), \tag{15}$$

with sample paths in $D_{\mathbb{R}^n}[0, \infty)$. Here and below, $D_{\mathbb{R}^n}[0, \infty)$ denotes the space of right-continuous functions from $[0, \infty)$ to $\mathbb{R}^n$ which possess left limits under the Skorohod topology. Let

$$\mathbf{C} = (C_{ij}) \qquad \text{where } C_{ij} = (i+j)! - i!\,j!. \tag{16}$$ □

LEMMA 2. *There exists a process* $\mathbf{W}(\cdot) \equiv (W_1(\cdot), W_2(\cdot), \ldots, W_n(\cdot))$ *with continuous sample paths and independent Gaussian increments satisfying* $\mathbf{W}(0) = \mathbf{0}$, $E(\mathbf{W}(t)) = \mathbf{0}$ *and* $\mathrm{Cov}(W_i(t), W_j(t)) = tC_{ij}$ *such that, as* $\theta \to \infty$, $\mathbf{W}_\theta(\cdot) \Rightarrow \mathbf{W}(\cdot)$ *in* $D_{\mathbb{R}^n}[0, \infty)$. *Also,* $W_i(t)W_j(t) - tC_{ij}$ *is an* $(\mathcal{F}_t^{\mathbf{W}})$*–martingale.*

In particular, the $W_i(t)$'s are correlated Brownian motions with $\mathrm{Var}\, W_i(t) = C_{ii} t$.

PROOF. To prove the lemma, we will apply the martingale central limit theorem [cf. Ethier and Kurtz (1984), page 339]. To this end, we use equation (13) to define

$$\boldsymbol{\xi}_\theta(i) \equiv \frac{1}{\sqrt{\theta}}\left(\mathbf{J}_\theta(i) - E(\mathbf{J}_\theta(i))\right), \quad i \geq 1.$$

It follows from (15) that

$$\mathbf{W}_\theta(t) = \sum_{k=1}^{[\theta t]} \boldsymbol{\xi}_\theta(k).$$

Letting $\mathscr{F}_k^\theta = \sigma\{\boldsymbol{\xi}_\theta(i) : 1 \leq i \leq k\}$ and denoting by $\boldsymbol{\xi}_\theta(k)'$ the column vector giving the transpose of $\boldsymbol{\xi}_\theta(k)$, we define the symmetric $n \times n$ matrix-valued process

$$\mathbf{C}_\theta(t) \equiv \sum_{k=1}^{[\theta t]} E\left[\boldsymbol{\xi}_\theta(k)'\boldsymbol{\xi}_\theta(k) | \mathscr{F}_{k-1}^\theta\right] = \sum_{k=1}^{[\theta t]} E\left[\boldsymbol{\xi}_\theta(k)'\boldsymbol{\xi}_\theta(k)\right],$$

the last equality following by independence of the $\mathbf{J}_\theta(i)$'s. The $\boldsymbol{\xi}_\theta(k)$ form a martingale difference array; that is, $\mathbf{W}_\theta(t)$ is a martingale with respect to the filtration $(\mathscr{F}_{[\theta t]}^\theta)$.

Verification of the hypotheses in the martingale central limit theorem rest on some standard results for moments of Beta $(1, \theta)$ random variables. If $U_k \sim$ Beta $(1, \theta)$, then

$$E(\theta U_k)^m = m! \frac{\theta^{m+1}}{\theta_{(m+1)}}, \tag{17}$$

$$\operatorname{Cov}\left((\theta U_k)^i, (\theta U_k)^j\right) = (i+j)! \frac{\theta^{i+j+1}}{\theta_{(i+j+1)}} - i! \frac{\theta^{i+1}}{\theta_{(i+1)}} \cdot j! \frac{\theta^{j+1}}{\theta_{(j+1)}}, \tag{18}$$

where $\theta_{(j)} = \theta(\theta+1)(\theta+2)\cdots(\theta+j-1)$. In addition, note that $(1-U_k)^\theta \sim$ Unif $(0, 1)$ and therefore

$$J_{1,\theta}(k) \equiv -\theta \log(1 - U_k) \sim \operatorname{Exp}(1). \tag{19}$$

It follows from (13) and (18) that, for $i, j = 2, \dots, n$,

$$\operatorname{Cov}\left(J_{i,\theta}(k), J_{j,\theta}(k)\right) = (i+j)! \frac{\theta^{i+j+1}}{\theta_{(i+j+1)}} - i! \frac{\theta^{i+1}}{\theta_{(i+1)}} \cdot j! \frac{\theta^{j+1}}{\theta_{(j+1)}} \tag{20}$$

and this $\to C_{ij}$, as $\theta \to \infty$. To calculate the covariance of $J_{1,\theta}(k)$ and $J_{j,\theta}(k)$, we begin with

$$\begin{aligned} E\left[J_{1,\theta}(k) J_{j,\theta}(k)\right] &= E\left[J_{1,\theta}(k)(\theta U_k)^j\right] \\ &= \theta^j E\left[J_{1,\theta}(k)\left(1 - \exp\{-J_{1,\theta}(k)/\theta\}\right)^j\right]. \end{aligned}$$

By the usual Taylor expansion we get

$$1 - \exp\{-J_{1,\theta}(k)/\theta\} = J_{1,\theta}/\theta + R_\theta$$

where R_θ is the remainder and $|R_\theta| \leq \dfrac{J_{1,\theta}(k)^2}{\theta^2}$. Therefore,

$$\begin{aligned}\lim_{\theta\to\infty} & E\left[J_{1,\theta}(k)J_{j,\theta}(k)\right] \\ &= \lim_{\theta\to\infty} \theta^j E\left[\left(J_{1,\theta}(k)/\theta\right)^j J_{1,\theta}(k)\right] = E\left[J_{1,\theta}(k)^{j+1}\right] = (j+1)!\end{aligned}$$

and hence

$$\lim_{\theta\to\infty} \operatorname{Cov}\left(J_{1,\theta}(k), J_{j,\theta}(k)\right) = (j+1)! - j! = C_{1j}. \tag{21}$$

Also, $\operatorname{Cov}(J_{1,\theta}(k), J_{1,\theta}(k)) = 1 = C_{11}$.

Now note that the entries of the matrix $\mathbf{C}_\theta(t)$ are given by

$$(\mathbf{C}_\theta(t))_{i,j} = \frac{[\theta t]}{\theta} \operatorname{Cov}\left(J_{i,\theta}(k), J_{j,\theta}(k)\right) \qquad (i, j = 1, 2, \cdots, n). \tag{22}$$

It follows from the above limits that

$$\lim_{\theta\to\infty} \mathbf{C}_\theta(t) = t\mathbf{C}, \tag{23}$$

where $\mathbf{C}$ is the matrix defined in (16).

Similarly, straightforward calculations show for each finite T,

$$\lim_{\theta\to\infty} E[\sup_{t\leq T} |\mathbf{C}_\theta(t) - \mathbf{C}_\theta(t-)|] \leq \lim_{\theta\to\infty} \frac{1}{\theta} \sum_{j=1}^{n}\sum_{i=1}^{n} C_{ij} = 0, \tag{24}$$

$$\lim_{\theta\to\infty} E[\sup_{t\leq T} |\mathbf{W}_\theta(t) - \mathbf{W}_\theta(t-)|^2] = 0, \tag{25}$$

and $\mathbf{W}_\theta(t)'\mathbf{W}_\theta(t) - \mathbf{C}_\theta(t)$ defines a matrix-valued $(\mathscr{F}^\theta_{[\theta t]})$–martingale.

Thus the conditions of the martingale central limit theorem are satisfied and we have $\mathbf{W}_\theta \Rightarrow \mathbf{W}$ where $\mathbf{W}$ is a process with independent Gaussian increments with $\mathbf{W}(0) = \mathbf{0}$, $E(\mathbf{W}(t)) = \mathbf{0}$ and covariance matrix $E(\mathbf{W}(t)'\mathbf{W}(t)) = t\mathbf{C}$. □

LEMMA 3. *Let $\mathbf{Y}_\theta(t)$ be defined by equation* (14) *and set* $\mathbf{y}(t) = t(1!, 2!, \ldots, n!)$. *Then, as $\theta \to \infty$, $\mathbf{Y}_\theta(\cdot) \xrightarrow{P} \mathbf{y}(\cdot)$ in $D_{\mathbb{R}^n}[0, \infty)$.*

PROOF. The result follows immediately from Lemma 2 and the fact that, as $\theta \to \infty$, $E\mathbf{J}_\theta(k) \to (1!, 2!, \ldots, n!)$. □

We are now ready to derive a large θ limit for the truncated sum $\sum_{k=1}^{[\theta t]} V^m_{k,\theta}$. To motivate the scaling in the following theorem, first note that it is natural to multiply by θ^m so that the easily-handled terms $(\theta U_k)^m$ will appear. The resulting sum has mean

$$
\begin{aligned}
E\left[\theta^m \sum_{k=1}^{[\theta t]} V_{k,\theta}^m\right] &= \sum_{k=1}^{[\theta t]} E\,[(\theta U_k)^m]\,E\left[\prod_{j=1}^{k-1}(1-U_j)^m\right] \\
&= m!\frac{\theta^{m+1}}{\theta_{(m+1)}}\sum_{k=1}^{[\theta t]} E\left[\exp\left\{-\frac{m}{\theta}\sum_{j=1}^{k-1} J_{1,\theta}(j)\right\}\right] \\
&= m!\frac{\theta^{m+1}}{\theta_{(m+1)}}\sum_{k=1}^{[\theta t]}\left(\frac{1}{1+m/\theta}\right)^{k-1} \\
&= m!\frac{\theta^{m+1}}{\theta_{(m+1)}}\frac{1-(1+m/\theta)^{-[\theta t]}}{1-(1+m/\theta)^{-1}} \\
&\sim m!\frac{\theta^{m+1}}{\theta_{(m+1)}}\frac{\theta}{m}\left(1-e^{-mt}\right) \quad \text{(for large } \theta) \\
&= m!\frac{\theta^{m+1}}{\theta_{(m+1)}}\theta\int_0^t e^{-ms}\,ds.
\end{aligned}
\tag{26}
$$

This suggests that centering $\theta^m \sum_{k=1}^{[\theta t]} V_{k,\theta}^m$ by this last expression (which will be easier to work with than the exact mean) and dividing by $\sqrt{\theta}$ should produce a nontrivial limit as $\theta \to \infty$. The following theorem gives an invariance principle for partial sums of the powers of the size-biased allele frequencies.

THEOREM 3. *Let $\mathbf{H}_\theta = (H_{2,\theta}, H_{3,\theta}, \dots, H_{n,\theta})$ and $\mathbf{H} = (H_2, H_3, \dots, H_n)$ be processes with sample paths in $D_{\mathbb{R}^{n-1}}[0,\infty)$, defined for $2 \le m \le n$ by*

$$
H_{m,\theta}(t) \equiv \theta^{m-1/2}\sum_{k=1}^{[\theta t]} V_{k,\theta}^m - m!\frac{\theta^{m+1}}{\theta_{(m+1)}}\sqrt{\theta}\int_0^t e^{-ms}\,ds \tag{27}
$$

and

$$
H_m(t) \equiv \int_0^t e^{-ms} dW_m(s) - m\cdot m!\int_0^t e^{-ms} W_1(s)\,ds, \tag{28}
$$

where $\mathbf{W} = (W_1, W_2, \dots, W_n)$ is the $\mathbb{R}^n$–valued process with independent Gaussian increments defined in Lemma 2. Then $\mathbf{H}_\theta \Rightarrow \mathbf{H}$ in $D_{\mathbb{R}^{n-1}}[0,\infty)$, as $\theta \to \infty$.

PROOF. Recall that $\mathbf{Y}_\theta(t) = (Y_{1,\theta}(t), Y_{2,\theta}(t), \dots, Y_{n,\theta}(t))$ is given by

$$
Y_{1,\theta}(t) = -\sum_{i=1}^{[\theta t]}\log(1-U_i) \quad \text{and} \quad Y_{m,\theta}(t) = \frac{1}{\theta}\sum_{i=1}^{[\theta t]}(\theta U_i)^m, \qquad 2 \le m \le n,
$$

$$
\exp\{-mY_{1,\theta}(t)\} = \prod_{i=1}^{[\theta t]}(1-U_i)^m,
$$

and for $2 \le m \le n$,

$$W_{m,\theta}(t) = \frac{1}{\sqrt{\theta}} \sum_{i=1}^{[\theta t]} [(\theta U_i)^m - E(\theta U_i)^m]. \tag{29}$$

Recall that $U_i \sim \text{Beta}(1, \theta)$, and hence

$$E[(\theta U_i)^m] = m! \frac{\theta^{m+1}}{\theta_{(m+1)}}. \tag{30}$$

Thus, for $2 \le m \le n$,

$$\begin{aligned} H_{m,\theta}(t) &= \theta^{m-1/2} \sum_{k=1}^{[\theta t]} U_k^m \prod_{i=1}^{k-1} (1-U_i)^m - m! \frac{\theta^{m+1}}{\theta_{(m+1)}} \sqrt{\theta} \int_0^t e^{-ms}\, ds \\ &= \sqrt{\theta} \int_0^t \exp\{-m Y_{1,\theta}(s-)\} dY_{m,\theta}(s) - m! \frac{\theta^{m+1}}{\theta_{(m+1)}} \sqrt{\theta} \int_0^t e^{-ms}\, ds \\ &= \sqrt{\theta} \int_0^t e^{-ms} dY_{m,\theta}(s) - m! \frac{\theta^{m+1}}{\theta_{(m+1)}} \sqrt{\theta} \int_0^t e^{-ms}\, ds \\ &\quad + \sqrt{\theta} \int_0^t \left(\exp\{-m Y_{1,\theta}(s-)\} - e^{-ms}\right) dY_{m,\theta}(s). \end{aligned} \tag{31}$$

Note that the product in the first line of (31) is defined to be 1 when $k = 1$.

We write $\mathbf{H}_\theta = \mathbf{H}_\theta^{(1)} + \mathbf{H}_\theta^{(2)}$ and $\mathbf{H} = \mathbf{H}^{(1)} + \mathbf{H}^{(2)}$ where, for $2 \le m \le n$,

$$H_{m,\theta}^{(1)}(t) \equiv \sqrt{\theta} \int_0^t e^{-ms} dY_{m,\theta}(s) - m! \frac{\theta^{m+1}}{\theta_{(m+1)}} \sqrt{\theta} \int_0^t e^{-ms}\, ds, \tag{32}$$

$$H_{m,\theta}^{(2)}(t) \equiv \sqrt{\theta} \int_0^t \left(\exp\{-m Y_{1,\theta}(s-)\} - e^{-ms}\right) dY_{m,\theta}(s), \tag{33}$$

$$H_m^{(1)}(t) \equiv \int_0^t e^{-ms}\, dW_m(s), \tag{34}$$

$$H_m^{(2)}(t) \equiv -m \cdot m! \int_0^t W_1(s) e^{-ms}\, ds. \tag{35}$$

To prove $\mathbf{H}_\theta \Rightarrow \mathbf{H}$ as $\theta \to \infty$, we show that

$$(\mathbf{H}_\theta^{(1)}, \mathbf{H}_\theta^{(2)}) \Rightarrow (\mathbf{H}^{(1)}, \mathbf{H}^{(2)}) \tag{36}$$

in $D_{\mathbb{R}^{2n-2}}[0, \infty)$.

Note that, for $2 \le m \le n$,

$$H^{(1)}_{m,\theta}(t) = \sqrt{\theta}\sum_{k=1}^{[\theta t]} \frac{1}{\theta}(\theta U_k)^m e^{-mk/\theta} - m!\frac{\theta^{m+1}}{\theta_{(m+1)}}\sqrt{\theta}\int_0^t e^{-ms}\,ds$$
$$= \frac{1}{\sqrt{\theta}}\sum_{k=1}^{[\theta t]}\left[(\theta U_k)^m - m!\frac{\theta^{m+1}}{\theta_{(m+1)}}\right]e^{-mk/\theta}$$
$$+ m!\frac{\theta^{m+1}}{\theta_{(m+1)}}\sqrt{\theta}\left[\sum_{k=1}^{[\theta t]} e^{-mk/\theta}\frac{1}{\theta} - \int_0^t e^{-ms}\,ds\right],$$

and hence it follows from equations (29) and (30) that

$$H^{(1)}_{m,\theta}(t) = \int_0^t e^{-ms}\,dW_{m,\theta}(s) + m!\frac{\theta^{m+1}}{\theta_{(m+1)}}\sqrt{\theta}\left[\sum_{k=1}^{[\theta t]} e^{-mk/\theta}\frac{1}{\theta} - \int_0^t e^{-ms}\,ds\right]. \tag{37}$$

It is easy to see that $\sum_{k=1}^{[\theta t]} e^{-mk/\theta}\frac{1}{\theta}$ is a Riemann sum approximation of $\int_0^t e^{-ms}\,ds$ and converges at rate $1/\theta$ as $\theta \to \infty$. Therefore

$$m!\frac{\theta^{m+1}}{\theta_{(m+1)}}\sqrt{\theta}\left[\sum_{k=1}^{[\theta t]} e^{-mk/\theta}\frac{1}{\theta} - \int_0^t e^{-ms}\,ds\right] \to 0. \tag{38}$$

We now establish the analogue of equation (37) for $\mathbf{H}^{(2)}_\theta(t)$. Recall that $J_{1,\theta}(i) \equiv -\theta\log(1-U_i)$ has an exponential distribution with mean 1 and $Y_{1,\theta}(t) = \frac{1}{\theta}\sum_{i=1}^{[\theta t]} J_{1,\theta}(i)$. For some ξ between s and t, we have the Taylor expansion

$$e^{-mt} = e^{-ms} - me^{-ms}(t-s) + r_m(\xi, s, t),$$

where $r_m(\xi, s, t) = \frac{m^2 e^{-m\xi}}{2}(t-s)^2$. Substituting $Y_{1,\theta}(s)$ for t yields

$$e^{-mY_{1,\theta}(s)} = e^{-ms} - me^{-ms}(Y_{1,\theta}(s) - s) + r_m(\xi, s, Y_{1,\theta}(s)). \tag{39}$$

Note that

$$\sqrt{\theta}\, r_m(\xi, s, Y_{1,\theta}(s)) \le \frac{m^2}{2}\sqrt{\theta}\,(Y_{1,\theta}(s) - s)^2. \tag{40}$$

Furthermore, it follows from Doob's submartingale inequality that

$$E\left[\sup_{0\le s\le t}\sqrt{\theta}\,(Y_{1,\theta}(s) - s)^2\right] = \sqrt{\theta}\,E\left[\sup_{0\le s\le t}\left(\frac{1}{\theta}\sum_{i=1}^{[\theta s]}(J_{1,\theta}(i) - 1)\right)^2\right]$$
$$\le \frac{\sqrt{\theta}}{\theta^2}E\left(\sum_{i=1}^{[\theta t]}(J_{1,\theta}(i) - 1)\right)^2$$
$$= \theta^{-3/2}[\theta t] \le t/\sqrt{\theta}.$$

This, together with (40), yields

$$\sup_{0\le s\le t} \sqrt{\theta}\, r_m(\xi, s, Y_{1,\theta}(s)) \xrightarrow{P} 0 \tag{41}$$

as $\theta \to \infty$.

It follows from (39) that

$$\begin{aligned} H^{(2)}_{m,\theta}(t) &= \sqrt{\theta}\int_0^t \left[\exp\{-mY_{1,\theta}(s-)\} - e^{-ms}\right] dY_{m,\theta}(s) \\ &= \int_0^t \left[-\sqrt{\theta}\,(Y_{1,\theta}(s-) - s)\, me^{-ms} + \sqrt{\theta}\, r_m(\xi, s, Y_{1,\theta}(s-))\right] dY_{m,\theta}(s). \end{aligned}$$

So, since $W_{1,\theta}(t) = \sqrt{\theta}\,(Y_{1,\theta}(t) - [\theta t]/\theta)$,

$$\begin{aligned} H^{(2)}_{m,\theta}(t) = \int_0^t \Big[&-W_{1,\theta}(s-)\, me^{-ms} \\ &+\sqrt{\theta}\,\left\{(s - [\theta s]/\theta)me^{-ms} + r_m(\xi, s, Y_{1,\theta}(s-))\right\}\Big] dY_{m,\theta}(s). \end{aligned} \tag{42}$$

Note that the second term in the integrand goes to 0 in probability. By Lemma 3 we have $\mathbf{Y}_\theta(\cdot) \xrightarrow{P} \mathbf{y}(\cdot)$ as $\theta \to \infty$ in $D_{\mathbb{R}^n}[0,\infty)$. Since this last function is deterministic, it follows from Lemma 2 that

$$(\mathbf{W}_\theta, \mathbf{Y}_\theta) \Rightarrow (\mathbf{W}, \mathbf{y}) \ \text{ in } D_{\mathbb{R}^{2n}}[0,\infty)$$

[cf. Billingsley (1968)].

Let $\mathbb{M}^{nn}$ denote the space of real-valued $n \times n$ matrices. To apply the theorem of Kurtz and Protter on convergence of stochastic integrals, define the following $\mathbb{M}^{nn}$–valued processes:

$$\mathbf{\Phi}(s) = \operatorname{diag}(e^{-s}, e^{-2s}, \ldots, e^{-ns}),$$

$$\mathbf{\Psi}_\theta(s) = W_{1,\theta}(s)\operatorname{diag}(e^{-s}, 2e^{-2s}, \ldots, ne^{-ns})$$

and

$$\mathbf{\Psi}(s) = W_1(s)\operatorname{diag}(e^{-s}, 2e^{-2s}, \ldots, ne^{-ns}).$$

We then have

$$(\mathbf{\Phi}, \mathbf{W}_\theta) \Rightarrow (\mathbf{\Phi}, \mathbf{W}) \ \text{ and } \ (\mathbf{\Psi}_\theta, \mathbf{Y}_\theta) \Rightarrow (\mathbf{\Psi}, \mathbf{y})$$

in $D_{\mathbb{M}^{nn}\times\mathbb{R}^n}[0,\infty)$. Applying Theorem 2.2 of Kurtz and Protter (1991), we obtain

$$\int_0^\cdot \mathbf{\Phi}(s-)\, d\mathbf{W}_\theta(s) \Rightarrow \int_0^\cdot \mathbf{\Phi}(s-)\, d\mathbf{W}(s)$$

and

$$\int_0^\cdot \mathbf{\Psi}_\theta(s-) d\mathbf{Y}_\theta(s) \Rightarrow \int_0^\cdot \mathbf{\Psi}(s-)\, d\mathbf{y}(s),$$

where the integrators should be thought of as column vectors. It is an easy exercise to see that these actually converge jointly in $D_{\mathbb{R}^{2n}}[0,\infty)$.

Now use this together with (37) and (42) and the limits established in (38) and (41) to get $(\mathbf{H}_\theta^{(1)}, \mathbf{H}_\theta^{(2)}) \Rightarrow (\mathbf{H}^{(1)}, \mathbf{H}^{(2)})$, and therefore $\mathbf{H}_\theta = \mathbf{H}_\theta^{(1)} + \mathbf{H}_\theta^{(2)} \Rightarrow \mathbf{H}^{(1)} + \mathbf{H}^{(2)} = \mathbf{H}$ in $D_{\mathbb{R}^{n-1}}[0,\infty)$. $\square$

LEMMA 4. *Given $\varepsilon > 0$, there exists $t_0 = t_0(\varepsilon) < \infty$ such that*

$$P\left(\left|\theta^{m-1/2}\sum_{k=[\theta t_0]+1}^{\infty} V_{k,\theta}^m - (m-1)!\frac{\theta^{m+1}}{\theta_{(m+1)}}\sqrt{\theta}e^{-mt_0}\right| > \varepsilon\right) < \frac{\varepsilon}{2} \tag{43}$$

for all $\theta \geq m$.

REMARK. Note that t_0 depends only on ε and not on θ. Therefore, the above lemma shows that the tail of the series converges in probability to zero uniformly in θ.

PROOF OF LEMMA 4. By the same argument used to derive (31) we have

$$\begin{aligned}
&\theta^{m-1/2}\sum_{k=[\theta t_0]+1}^{\infty} V_{k,\theta}^m - (m-1)!\frac{\theta^{m+1}}{\theta_{(m+1)}}\sqrt{\theta}e^{-mt}\\
&\quad = \theta^{m-1/2}\sum_{k=[\theta t]+1}^{\infty} U_k^m \prod_{i=1}^{k-1}(1-U_i)^m - m!\frac{\theta^{m+1}}{\theta_{(m+1)}}\sqrt{\theta}\int_t^\infty e^{-ms}\,ds\\
&\quad = \sqrt{\theta}\int_t^\infty \left(\exp\{-mY_{1,\theta}(s-)\} - e^{-ms}\right) dY_{m,\theta}(s)\\
&\qquad + \sqrt{\theta}\int_t^\infty e^{-ms}\,dY_{m,\theta}(s) - m!\frac{\theta^{m+1}}{\theta_{(m+1)}}\sqrt{\theta}\int_t^\infty e^{-ms}\,ds.
\end{aligned} \tag{44}$$

So to prove Lemma 4 we need only show that

$$\sqrt{\theta}\int_t^\infty e^{-ms} dY_{m,\theta}(s) - m!\frac{\theta^{m+1}}{\theta_{(m+1)}}\sqrt{\theta}\int_t^\infty e^{-ms} ds \xrightarrow{P} 0 \tag{45}$$

and

$$\sqrt{\theta}\int_t^\infty \left(\exp\{-mY_{1,\theta}(s-)\} - e^{-ms}\right) dY_{m,\theta}(s) \xrightarrow{P} 0 \tag{46}$$

as $t \to \infty$, uniformly in θ.

To show (45), we note that

$$
\begin{aligned}
&\left|\sqrt{\theta}\int_t^\infty e^{-ms}\,dY_{m,\theta}(s) - m!\frac{\theta^{m+1}}{\theta_{(m+1)}}\sqrt{\theta}\int_t^\infty e^{-ms}\,ds\right| \\
&\quad = \left|\sqrt{\theta}\sum_{k=[\theta t]+1}^{\infty}\frac{1}{\theta}(\theta U_k)^m e^{-mk/\theta} - m!\frac{\theta^{m+1}}{\theta_{(m+1)}}\sqrt{\theta}\int_t^\infty e^{-ms}\,ds\right| \\
&\quad \le \left|\sum_{k=[\theta t]+1}^{\infty}\big((\theta U_k)^m - E(\theta U_k)^m\big)\frac{e^{-mk/\theta}}{\sqrt{\theta}}\right| \\
&\qquad + m!\frac{\theta^{m+1}}{\theta_{(m+1)}}\sqrt{\theta}\left|\sum_{k=[\theta t]+1}^{\infty} e^{-mk/\theta}\frac{1}{\theta} - \int_t^\infty e^{-ms}\,ds\right|.
\end{aligned}
\tag{47}
$$

We now treat each of the last two quantities separately.

Since $\mathrm{Var}[(\theta U)^m] \le (2m)!$ when $U \sim \mathrm{Beta}\,(1,\theta)$, Chebyshev's inequality and independence of the U_k's implies

$$
\begin{aligned}
P\left(\left|\sum_{k=[\theta t]+1}^{\infty}\big((\theta U_k)^m - E(\theta U_k)^m\big)\frac{e^{-mk/\theta}}{\sqrt{\theta}}\right| > \varepsilon\right) &\le \frac{(2m)!}{\varepsilon^2}\sum_{k=[\theta t]+1}^{\infty}\frac{e^{-2mk/\theta}}{\theta} \\
&\le \frac{(2m)!}{\varepsilon^2}\int_t^\infty e^{-2ms}\,ds \\
&= \frac{(2m-1)!}{\varepsilon^2}e^{-2mt}.
\end{aligned}
\tag{48}
$$

So the first term in (47) converges in probability to 0 as $t \to \infty$, uniformly in θ. The other term is non-random, so we just need to bound it by quantities that go to 0 as $t \to \infty$, uniformly in θ.

Approximating the integral by upper and lower Riemann sums,

$$
\sum_{k=[\theta t]+1}^{\infty} e^{-mk/\theta}\frac{1}{\theta} \le \int_t^\infty e^{-ms}\,ds \le \sum_{k=[\theta t]}^{\infty} e^{-mk/\theta}\frac{1}{\theta},
$$

allows us to bound the quantity at the end of (47):

$$
\begin{aligned}
&m!\frac{\theta^{m+1}}{\theta_{(m+1)}}\sqrt{\theta}\left|\sum_{k=[\theta t]+1}^{\infty} e^{-mk/\theta}\frac{1}{\theta} - \int_t^\infty e^{-ms}\,ds\right| \\
&\quad \le m!\frac{\theta^{m+1}}{\theta_{(m+1)}}\sqrt{\theta}\left[\sum_{k=[\theta t]}^{\infty} e^{-mk/\theta}\,\frac{1}{\theta}\right] \\
&\qquad - m!\frac{\theta^{m+1}}{\theta_{(m+1)}}\sqrt{\theta}\left[\sum_{k=[\theta t]+1}^{\infty} e^{-mk/\theta}\,\frac{1}{\theta}\right] \\
&\quad = m!\frac{\theta^{m+1}}{\theta_{(m+1)}\sqrt{\theta}}\,e^{-m[\theta t]/\theta} \\
&\quad \le m!\,e^{-m(t-1)}.
\end{aligned}
\tag{49}
$$

Therefore, (45) follows from (48) and (49).

To show (46), begin with a simple Laplace transform calculation to get

$$
\begin{aligned}
E\left[\exp\left(-mY_{1,\theta}(t)\right)\right] &= E\left[\exp\left\{-\frac{m}{\theta}\sum_{i=1}^{[\theta t]} J_{1,\theta}(i)\right\}\right] \\
&= (1+m/\theta)^{-[\theta t]} \\
&= \exp\left\{-[\theta t]\log(1+m/\theta)\right\} \\
&\le \exp\left\{-\frac{[\theta t]m}{2\theta}\right\},
\end{aligned} \tag{50}
$$

provided $\theta \ge m$. The last line of (50) follows from the fact that $\log(1+x) \ge x/(x+1) \ge x/2$ when $0 \le x \le 1$.

Now,

$$
|e^{-my} - e^{-mx}| = \left|\int_x^y me^{-mu}\,du\right| \le me^{-m\min\{x,y\}}|y-x|,
$$

so defining $R_\theta(s) = \min\{Y_{1,\theta}(s), s\}$, we have

$$
\begin{aligned}
&E\left|\sqrt{\theta}\int_t^\infty \left(\exp\{-mY_{1,\theta}(s-)\} - e^{-ms}\right)\,dY_{m,\theta}(s)\right| \\
&\quad\le \sqrt{\theta}E\int_t^\infty m\,\exp\left\{-mR_\theta(s)\right\}\left|Y_{1,\theta}(s-) - s\right|\,dY_{m,\theta}(s) \\
&\quad\le \sum_{k=[\theta t]+1}^{\infty} E\left[m\exp\left\{-mR_\theta(k/\theta)\right\}\frac{\left|\sum_{i=1}^k J_{1,\theta}(i) - k\right|}{\sqrt{\theta}}\,\frac{(\theta U_k)^m}{\theta}\right] \\
&\quad= \sum_{k=[\theta t]+1}^{\infty} E[Q_1(k)Q_2(k)Q_3(k)],
\end{aligned}
$$

where

$$
Q_1(k) = m\exp\left\{-mR_\theta(k/\theta)\right\},
$$

$$
Q_2(k) = \frac{\left|\sum_{i=1}^k J_{1,\theta}(i) - k\right|}{\sqrt{\theta}}
$$

and

$$
Q_3(k) = \frac{(\theta U_k)^m}{\theta}.
$$

The generalized version of Hölder's inequality yields

$$
E[Q_1Q_2Q_3] \le \left[E(Q_1^4)\right]^{1/4}\left[E(Q_2^2)\right]^{1/2}\left[E(Q_3^4)\right]^{1/4}.
$$

It follows from equation (50) that

$$\begin{aligned} E(Q_1^4) &= m^4 E\left[\exp\{-4mR_\theta(k/\theta)\}\right] \\ &= m^4 E\left[\exp\{-4m(Y_{1,\theta}(k/\theta)\wedge(k/\theta)\}\right] \\ &\le m^4 E\left[\exp\{-4m(Y_{1,\theta}(k/\theta)\}\right] + m^4\exp\{-4mk/\theta\} \\ &\le m^4\exp\{-2mk/\theta\} + m^4\exp\{-4mk/\theta\} \\ &\le 2m^4\exp\{-2mk/\theta\}. \end{aligned}$$

Therefore,

$$E[(Q_1^4)]^{1/4} \le 2^{1/4} m\ \exp\{-mk/(2\theta)\}$$

Since $J_{1,\theta}(1), J_{1,\theta}(2), \ldots$ are independent exponentials with mean one, we have

$$[E(Q_2^2)]^{1/2} = \left[\frac{1}{\theta}\operatorname{Var}\left(\sum_{i=1}^{k} J_{1,\theta}(i)\right)\right]^{1/2} = \sqrt{\frac{k}{\theta}}.$$

Using (30), we have

$$\left[E(Q_3^4)\right]^{1/4} \le \frac{1}{\theta}[(4m)!]^{1/4}.$$

Therefore,

$$\begin{aligned} &E\left|\sqrt{\theta}\int_t^\infty \left(\exp\{-mY_{1,\theta}(s-)\} - e^{-ms}\right) dY_{m,\theta}(s)\right| \\ &\le \frac{2^{1/4}m[(4m)!]^{1/4}}{\theta}\sum_{k=[\theta t]+1}^{\infty}\sqrt{\frac{k}{\theta}}\exp\{-mk/(2\theta)\} \qquad (51) \\ &\le m2^{1/4}[(4m)!]^{1/4}\int_t^\infty \sqrt{s}e^{-(ms)/2}\,ds, \end{aligned}$$

the last line following by a Riemann sum approximation. Thus we have shown

$$E\left|\sqrt{\theta}\int_t^\infty \left(\exp\{-mY_{1,\theta}(s-)\} - e^{-ms}\right) dY_{m,\theta}(s)\right| \to 0,$$

uniformly in θ. This implies

$$\sqrt{\theta}\int_t^\infty \left(\exp\{-mY_{1,\theta}(s-)\} - e^{-ms}\right) dY_{m,\theta}(s) \xrightarrow{P} 0,$$

uniformly in θ, as desired. *Box*

LEMMA 5. *Let* $\mathbf{H}(\cdot)$ *be the* $D_{\mathbb{R}^{n-1}}[0,\infty)$*−valued process defined in Theorem 3. Then* $\mathbf{H}(\infty) \equiv \lim_{t\to\infty}\mathbf{H}(t)$ *exists and has a multivariate normal distribution with mean* $\mathbf{0}$ *and covariance matrix given by*

$$\operatorname{Cov}(\mathbf{H}_i(\infty), \mathbf{H}_j(\infty)) = (i+j-1)! - i!\,j! \qquad (i, j = 2, \ldots, n). \qquad (52)$$

PROOF. We first note that the integration by parts formula for stochastic integrals [cf. Karatzas and Shreve (1988)] implies

$$\int_0^t e^{-ms} dW_1(s) = e^{-mt} W_1(t) + m \int_0^t e^{-ms} W_1(s)\, ds. \tag{53}$$

So

$$\begin{aligned} H_m(t) &= \int_0^t e^{-ms}\, dW_m(s) - m \cdot m! \int_0^t e^{-ms} W_1(s)\, ds \\ &= \int_0^t e^{-ms}\, dW_m(s) - m! \int_0^t e^{-ms}\, dW_1(s) + m! e^{-mt} W_1(t). \end{aligned} \tag{54}$$

Note that

$$\left(\int_0^t e^{-2s}\, dW_2(s), \ldots, \int_0^t e^{-ns}\, dW_n(s), \int_0^t e^{-2s} dW_1(s), \ldots, \int_0^t e^{-ns} dW_1(s) \right)$$

defines a uniformly integrable martingale (bounded second moments) in $\mathbb{R}^{2n-2}$, and hence the limit exists as $t \to \infty$, with probability one. Denote the limit by $(\mathbf{E}, \mathbf{G})$, where

$$E_m = \int_0^\infty e^{-ms}\, dW_m(s) \quad \text{and} \quad G_m = \int_0^\infty e^{-ms}\, dW_1(s) \qquad (m = 2, \ldots, n).$$

Thus $\mathbf{H}(\infty) = (H_2(\infty), \ldots, H_n(\infty))$, where

$$H_m(\infty) = \int_0^\infty e^{-ms}\, dW_m(s) - m! \int_0^\infty e^{-ms}\, dW_1(s) = E_m - m! G_m. \tag{55}$$

Now, $(\mathbf{E}, \mathbf{G})$ is multivariate normal (due to the deterministic integrands) with mean $\mathbf{0}$ and its covariance matrix is determined by the covariance structure of $\mathbf{W}$, given in Lemma 2. Therefore $\mathbf{H}(\infty)$ is also multivariate normal with mean $\mathbf{0}$ and covariance matrix to be computed next. We begin by calculating, for $i, j = 2, \ldots, n$,

$$\operatorname{Cov}(E_i, E_j) = \int_0^\infty e^{-(i+j)s} C_{ij}\, ds = \frac{C_{ij}}{i+j},$$

$$\operatorname{Cov}(E_i, G_j) = \frac{C_{i1}}{i+j} \quad \text{and} \quad \operatorname{Cov}(G_i, G_j) = \frac{C_{11}}{i+j}.$$

Hence, by (55),

$$\begin{aligned} \operatorname{Cov}(H_i(\infty), H_j(\infty)) &= \frac{C_{ij} - i! C_{j1} - j! C_{i1} + i! j! C_{11}}{i+j} \\ &= (i+j-1)! - i! j!, \end{aligned}$$

for $i, j = 2, \ldots, n$. This completes the proof of Lemma 5. □

PROOF OF THEOREM 1. It follows from standard results for product spaces [Ethier and Kurtz (1986)] that $\mathbf{Z}_\theta \Rightarrow \mathbf{Z}$ as $\theta \to \infty$ if and only if $(Z_{2,\theta}, Z_{3,\theta}\dots, Z_{n,\theta})$ converges in distribution to $(Z_2, Z_3, \dots, Z_n)$ as $\theta \to \infty$ for any integer n. It follows from Lemma 5 and the definition of $\mathbf{Z}$ in Theorem 1 that

$$\mathbf{H}(\infty) \overset{d}{=} (Z_2, 2!Z_3, \dots, (n-1)!Z_n), \tag{56}$$

since both have multivariate normal distribution with mean $\mathbf{0}$ and covariance as in (52). Define $\mathbf{H}_\theta(\infty) = \lim_{t\to\infty} \mathbf{H}_\theta(t)$. Then

$$\begin{aligned} H_{m,\theta}(\infty) &= \theta^{m-1/2}\sum_{i=1}^{\infty} V_{i,\theta}^m - (m-1)!\sqrt{\theta}\frac{\theta^{m+1}}{\theta_{(m+1)}} \\ &= (m-1)!Z_{m,\theta} - (m-1)!\sqrt{\theta}\left(\frac{\theta^{m+1}}{\theta_{(m+1)}} - 1\right). \end{aligned} \tag{57}$$

Since the last term is $O(1/\sqrt{\theta}) \to 0$, (56) and (57) imply that $\mathbf{Z}_\theta \Rightarrow \mathbf{Z}$ will follow if we can show that $\mathbf{H}_\theta(\infty) \Rightarrow \mathbf{H}(\infty)$. From Theorem 3 and Lemma 5 we know that $\lim_{t\to\infty}\lim_{\theta\to\infty}\mathbf{H}_\theta(t) = \lim_{t\to\infty}\mathbf{H}(t) = \mathbf{H}(\infty)$. So to prove that $\mathbf{H}_\theta(\infty) \Rightarrow \mathbf{H}(\infty)$ we must justify the interchange of limits.

Given $\varepsilon > 0$, define $\boldsymbol{\varepsilon} = (\varepsilon, \dots, \varepsilon) \in \mathbb{R}^{n-1}$. By Lemma 4 we can choose t_0 large enough (and independent of $\theta \geq m$) so that

$$P\left(\left|\theta^{m-1/2}\sum_{i=[\theta t_0]+1}^{\infty} V_{i,\theta}^m - (m-1)!\frac{\theta^{m+1}}{\theta_{(m+1)}}\sqrt{\theta}e^{-mt_0}\right| > \varepsilon\right) < \frac{\varepsilon}{2} \tag{58}$$

for all $m = 2, \dots, n$. Taking t_0 even larger if necessary, we also have

$$\begin{aligned} P(\mathbf{H}(\infty) \leq \mathbf{x} - \boldsymbol{\varepsilon}) - \varepsilon/2 &\leq P(\mathbf{H}(t_0) \leq \mathbf{x} - \boldsymbol{\varepsilon}), \\ P(\mathbf{H}(t_0) \leq \mathbf{x} + \boldsymbol{\varepsilon}) &\leq P(\mathbf{H}(\infty) \leq \mathbf{x} + \boldsymbol{\varepsilon}) + \varepsilon/2. \end{aligned} \tag{59}$$

It follows from equation (58) that for all $m = 2, \dots, n$

$$P(|H_{m,\theta}(\infty) - H_{m,\theta}(t_0)| > \varepsilon) < \varepsilon/2. \tag{60}$$

Equation (60) implies that

$$P(\mathbf{H}_\theta(t_0) \leq \mathbf{x} - \boldsymbol{\varepsilon}) - \varepsilon/2 \leq P(\mathbf{H}_\theta(\infty) \leq \mathbf{x}) \leq P(\mathbf{H}_\theta(t_0) \leq \mathbf{x} + \boldsymbol{\varepsilon}) + \varepsilon/2. \tag{61}$$

Taking the limsup as $\theta \to \infty$ in equation (61) gives

$$\begin{aligned} P(\mathbf{H}(t_0) \leq \mathbf{x} - \boldsymbol{\varepsilon}) - \varepsilon/2 &\leq \limsup_{\theta\to\infty} P(\mathbf{H}_\theta(\infty) \leq \mathbf{x}) \\ &\leq P(\mathbf{H}(t_0) \leq \mathbf{x} + \boldsymbol{\varepsilon}) + \varepsilon/2. \end{aligned} \tag{62}$$

Applying equation (59) to equation (62) we get

$$P(\mathbf{H}(\infty) \le \mathbf{x} - \boldsymbol{\varepsilon})) - \varepsilon \le \limsup_{\theta\to\infty} P(\mathbf{H}_\theta(\infty) \le \mathbf{x}) \le P(\mathbf{H}(\infty) \le \mathbf{x} + \boldsymbol{\varepsilon}) + \varepsilon.$$

Since $\mathbf{H}(\infty)$ has a continuous distribution, we take the limit as $\varepsilon \to 0$ above to get

$$\limsup_{\theta\to\infty} P(\mathbf{H}_\theta(\infty) \le \mathbf{x}) = P(\mathbf{H}(\infty) \le \mathbf{x}).$$

Theorem 1 follows after a similar argument for the lim inf. □

2.1. *Proof of Lemma 1.* The proof of Lemma 1 requires several approximations. In order to show that the remainder term, denoted by $R_{\mathbf{a}}(\theta)$, goes to zero as $\theta \to \infty$ appropriately, we will need to prove the following lemma.

LEMMA 6. *Suppose* $\mathbf{X} \sim \mathscr{P}D(\theta)$ *and let* $\mathbf{a}$ *be a partition of a sample of size* n *drawn from a population with frequencies* $\mathbf{X}$. *Define* $F_m = \sum_{i=1}^{\infty} \theta^{m-1} X_i^m$. *Then for all* $\varepsilon > 0$, *as* $\theta \to \infty$,

$$\theta^{1/2-\varepsilon}\left(1 - \prod_{m=2}^{n}\left(\frac{F_m}{(m-1)!}\right)^{a_m}\right) \Rightarrow 0. \tag{63}$$

PROOF. Note that if $a_1 = n$ then $a_i = 0$ for all $i > 1$. In this case the left hand side of (63) is identically 0 and so the result follows trivially in this case. Therefore, we will assume that $a_1 \ne n$. We proceed by induction on the sample size n. If $n = 2$, then by our assumption, a_2 must be 1 and hence, using (4), equation (63) reduces to

$$\theta^{1/2-\varepsilon}(1 - F_2) = -\theta^{-\varepsilon} Z_{2,\theta} \stackrel{d}{=} \theta^{-\varepsilon} Z_{2,\theta}.$$

and the result follows by Theorem 1. The induction hypothesis is that (63) is true for all partitions of the integer n. Let $\mathbf{a} = (a_1, a_2, \ldots, a_{n+1})$ be a partition of $n + 1$. Since $a_1 \ne n + 1$ there exist an $a_j \ge 1$ for some $j > 1$. Define $\mathbf{a}_j = (a_{j1}, \ldots, a_{jn})$ to be a partition of $n+1-j$ formed by removing one allele with j representatives. That is, $a_{jj} = a_j - 1$ and $a_{ij} = a_i$ for all $i \ne j$. Then

$$\left(1 - \prod_{m=2}^{n+1}\left(\frac{F_m}{(m-1)!}\right)^{a_m}\right) = \frac{F_j}{(j-1)!}\left(1 - \prod_{m=2}^{n}\left(\frac{F_m}{(m-1)!}\right)^{a_{mj}}\right) + 1 - \tfrac{F_j}{(j-1)!} \tag{64}$$

By Theorem 1, as $\theta \to \infty$, $F_j/(j-1)! \Rightarrow 1$ and $\theta^{1/2-\varepsilon}(1 - F_j/(j-1)!) \Rightarrow 0$. By the induction hypothesis, as $\theta \to \infty$

$$\theta^{1/2-\varepsilon}\left(1 - \prod_{m=2}^{n}\left(\frac{F_m}{(m-1)!}\right)^{a_{mj}}\right) \Rightarrow 0.$$

Therefore the result follows by multiplying equation (64) by $\theta^{1/2-\varepsilon}$ and taking the limit as $\theta \to \infty$. □

PROOF OF LEMMA 1. Let $n_1, n_2, \dots, n_k$ be the allele frequencies, listed in descending order, from a sample of size n (with k distinct alleles) drawn from a population with allele relative frequencies $\mathbf{x} = (x_1, x_2, \dots) \in \nabla$; so $n_1 \geq n_2 \geq \cdots \geq n_k > 0$ and $n_1 + \cdots + n_k = n$. Now write

$$\prod_{i=1}^{k} \sum_{j=1}^{\infty} x_j^{n_i} = \sum_{(i_1,\dots,i_k)\in D} x_{i_1}^{n_1} \cdots x_{i_k}^{n_k} + \sum_{(i_1,\dots,i_k)\in R} x_{i_1}^{n_1} \dots x_{i_k}^{n_k}, \tag{65}$$

where D is the set distinct indices (i.e., $i_j \neq i_m$ for all $j \neq m$) and every vector in R has at least two entries that are the same. We will now rewrite equation (65) using the partition structure notation established by Kingman (1977). For a given partition $\mathbf{a}$, define $\mathscr{C}_{\mathbf{a}}$ to be the collection of all partitions formed by coalescing two or more of the classes associated with $\mathbf{a}$. (For example, suppose $n = 5$, $a_1 = 1$ and $a_2 = 2$. Think of this as two red balls, two white balls, and one green. If we coalesce the red and white balls into one class, we form a new partition $\mathbf{b}$, where $b_1 = 1$, and $b_4 = 1$. Note that $\mathbf{b}$ is one of the elements in $\mathscr{C}_{\mathbf{a}}$.)

Recall the multinomial probability function $\phi_{\mathbf{a}}(\mathbf{x})$ defined by equation (1). It follows that, for $\nu \in \mathscr{A}_{\mathbf{a}}$,

$$\sum_{(i_1,\dots,i_k)\in D} x_{i_1}^{n_1} \cdots x_{i_k}^{n_k} = \prod_{i=1}^{n} a_i! \sum_{\nu\in\mathscr{A}_{\mathbf{a}}} x_1^{\nu_1} x_2^{\nu_2} \cdots = \frac{\prod_{i=1}^{n} a_i!(i!)^{a_i}}{n!} \phi_{\mathbf{a}}(\mathbf{x}) \tag{66}$$

where the non-zero entries of $\nu = (\nu_1, \nu_2, \dots,)$ are given by $n_j = \nu_{i_j}$, and

$$\sum_{(i_1,\dots,i_k)\in R} x_{i_1}^{n_1} \cdots x_{i_k}^{n_k} = \sum_{\mathbf{b}\in\mathscr{C}_{\mathbf{a}}} l_{\mathbf{b}} \phi_{\mathbf{b}}(\mathbf{x}), \tag{67}$$

where $l_{\mathbf{b}}$ is a combinatorial factor that we could in principle determine. Fortunately, the proof does not depend on knowing the explicit form of this constant.

An alternative way to express the left side of (65) is

$$\prod_{i=1}^{n} \left(\sum_{j=1}^{\infty} x_j^i \right)^{a_i} = \prod_{i=1}^{k} \sum_{j=1}^{\infty} x_j^{n_i}$$

Equations (66) and (67) give an alternative way to express the right side of equation (65). Therefore,

$$\prod_{i=1}^{n} \left(\sum_{j=1}^{\infty} x_j^i \right)^{a_i} = \frac{\prod_{i=1}^{n} a_i!(i!)^{a_i}}{n!} \phi_{\mathbf{a}}(\mathbf{x}) + \sum_{\mathbf{b}\in\mathscr{C}_{\mathbf{a}}} l_{\mathbf{b}} \phi_{\mathbf{b}}(\mathbf{x}). \tag{68}$$

Note that if $a_1 = n$ then (68) degenerates to the equation 1=1. So, again, we can assume that $a_1 \neq n$. Let $k = \sum a_i$. Evaluate equation (68) at $\mathbf{X} \sim \mathscr{P}D(\theta)$ and multiply both sides of the equation by $\frac{\theta^{n-k}}{\prod_{i=2}^{n}((i-1)!)^{a_i}}$ and recall that $\sum_{i=1}^{n} i a_i = n$

and $\sum_{i=1}^n a_i = k$ to get

$$\begin{aligned}\prod_{i=2}^{n}\left(\sum_{j=1}^{\infty}\frac{\theta^{i-1}X_j^i}{(i-1)!}\right)^{a_i} &= \theta^{n-k}\frac{\prod_{i=1}^n a_i! i^{a_i}}{n!}\phi_{\mathbf{a}}(\mathbf{X}) \\ &\quad + \frac{\theta^{n-k}}{\prod_{i=2}^n((i-1)!)^{a_i}}\sum_{\mathbf{b}\in\mathscr{C}_{\mathbf{a}}} l_{\mathbf{b}}\phi_{\mathbf{b}}(\mathbf{X}) \\ &= \frac{\theta^n}{\theta_{(n)}}\frac{\phi_{\mathbf{a}}(\mathbf{X})}{E(\phi_{\mathbf{a}}(\mathbf{X}))} \\ &\quad + \frac{\theta^{n-k}}{\prod_{i=2}^n((i-1)!)^{a_i}}\sum_{\mathbf{b}\in\mathscr{C}_{\mathbf{a}}} l_{\mathbf{b}}\phi_{\mathbf{b}}(\mathbf{X}).\end{aligned} \tag{69}$$

Recall that $F_i = \sum_{j=1}^{\infty}\theta^{i-1}X_j^i$. It is a simple consequence of the Ewens sampling formula (3) that $E(\phi_{\mathbf{a}}(\mathbf{x})) = O(\theta^{k-n})$. Since $\mathbf{b}$ has fewer than k alleles we see that $\theta^{n-k}E\left(\sum_{\mathbf{b}\in\mathscr{C}_{\mathbf{a}}}\phi_{\mathbf{b}}(\mathbf{X})\right) = O(1/\theta)$. Note also that $\theta^n/\theta_{(n)} - 1 = O(1/\theta)$. Now take the log of both sides of (69) to get

$$\sum_{i=2}^{n} a_i \log\left(\frac{F_i}{(i-1)!}\right) = \log\left(\frac{\phi_{\mathbf{a}}(\mathbf{X})}{E(\phi_{\mathbf{a}}(\mathbf{X}))} + O(1/\theta)\right). \tag{70}$$

Note that, for any real number y, we have $(y-1)^2/y \le \log y \le y-1$. Therefore,

$$0 \le y - 1 - \log y \le (y-1)^2/y.$$

So the linear approximation to $\log y$ is

$$\log y = y - 1 - r(y) \tag{71}$$

where $0 \le r(y) \le (y-1)^2/y$. Applying (71) to both sides of (70) we get

$$\log\left(\frac{F_i}{(i-1)!}\right) = \frac{F_i}{(i-1)!} - 1 - r\left(\frac{F_i}{(i-1)!}\right).$$

and

$$\log\left(\frac{\phi_{\mathbf{a}}(\mathbf{X})}{E(\phi_{\mathbf{a}}(\mathbf{X}))} + O(1/\theta)\right) = \frac{\phi_{\mathbf{a}}(\mathbf{X})}{E(\phi_{\mathbf{a}}(\mathbf{X}))} - 1 + O(1/\theta) - r\left(\frac{\phi_{\mathbf{a}}(\mathbf{X})}{E(\phi_{\mathbf{a}}(\mathbf{X}))}\right).$$

It follows from Theorem 1 that for any $\varepsilon > 0$

$$\theta^{1-2\varepsilon} r\left(\frac{F_i}{(i-1)!}\right) \le \theta^{-2\varepsilon}\left(\sqrt{\theta}(1 - \frac{F_i}{(i-1)!})\right)^2 \Big/ \frac{F_i}{(i-1)!} \Rightarrow 0 \tag{72}$$

as $\theta \to \infty$. It follows from Lemma 6 and equation (69) that we have the following limit in distribution:

$$\lim_{\theta\to\infty}\theta^{1/2-\varepsilon}\left(1 - \frac{\phi_{\mathbf{a}}(\mathbf{X})}{E(\phi_{\mathbf{a}}(\mathbf{X}))}\right) = \lim_{\theta\to\infty}\theta^{1/2-\varepsilon}\left(1 - \prod_{i=2}^{n}\left(\frac{F_i}{(i-1)!}\right)^{a_i}\right) = 0$$

implying

$$\theta^{1-2\varepsilon} r\left(\frac{\phi_{\mathbf{a}}(\mathbf{X})}{E(\phi_{\mathbf{a}}(\mathbf{X}))}\right) \Rightarrow 0 \tag{73}$$

as $\theta \to \infty$.

So we can now apply the linear approximations to both sides of of (70) to get

$$\sum_{i=2}^{n} a_i\left(\frac{F_i}{(i-1)!}-1\right) = \frac{\phi_{\mathbf{a}}(\mathbf{X})}{E(\phi_{\mathbf{a}}(\mathbf{X}))} - 1 + O(1/\theta) - r\left(\frac{\phi_{\mathbf{a}}(\mathbf{X})}{E(\phi_{\mathbf{a}}(\mathbf{X}))}\right) + \sum_{i=1}^{n} a_i r\left(\frac{F_i}{(i-1)!}\right)$$

Multiply both sides of the above equation by $\sqrt{\theta}$ and define $R_{\mathbf{a}}(\theta)$ to be $\sqrt{\theta}$ times the last three terms on right side of the above equation. That is,

$$R_{\mathbf{a}}(\theta) = O(1/\sqrt{\theta}) - \sqrt{\theta} r\left(\frac{\phi_{\mathbf{a}}(\mathbf{X})}{E(\phi_{\mathbf{a}}(\mathbf{X}))}\right) + \sqrt{\theta}\sum_{i=1}^{n} a_i r\left(\frac{F_i}{(i-1)!}\right). \tag{74}$$

Therefore,

$$\sum_{i=2}^{n} a_i Z_{i,\theta} = \sqrt{\theta}\sum_{i=2}^{n} a_i\left(\frac{F_i}{(i-1)!}-1\right) = \sqrt{\theta}\frac{\phi_{\mathbf{a}}(\mathbf{X}) - E(\phi_{\mathbf{a}}(\mathbf{X}))}{E(\phi_{\mathbf{a}}(\mathbf{X}))} + R_{\mathbf{a}}(\theta).$$

It follows from (72), (73) and (74) that $\lim_{\theta\to\infty} \theta^{1/2-\varepsilon} R_{\mathbf{a}}(\theta) = 0$. □

2.2. *Proof of Theorem* 2. As previously stated, if $a_1 \neq n$ then equation (8) follows immediately from Lemma 1 and Theorem 1. We need only show (9). Note that

$$\phi_{\mathbf{a}_1}(\mathbf{X}) = 1 - \sum_{\mathbf{a}\neq\mathbf{a}_1} \phi_{\mathbf{a}}(\mathbf{X}).$$

Therefore

$$\begin{aligned}\phi_{\mathbf{a}_1}(\mathbf{X}) - E\left(\phi_{\mathbf{a}_1}(\mathbf{X})\right) &= E\left(\sum_{\mathbf{a}\neq\mathbf{a}_1} \phi_{\mathbf{a}}(\mathbf{X})\right) - \sum_{\mathbf{a}\neq\mathbf{a}_1} \phi_{\mathbf{a}}(\mathbf{X}) \\ &= \left(E(\phi_{\mathbf{a}_2}(\mathbf{X})) - \phi_{\mathbf{a}_2}(\mathbf{X})\right) + \sum_{\mathbf{a}\neq\mathbf{a}_1,\mathbf{a}_2} \left(E(\phi_{\mathbf{a}}(\mathbf{X})) - \phi_{\mathbf{a}}(\mathbf{X})\right).\end{aligned}$$

It follows from equation (8) that if $\sum a_i = |\mathbf{a}|$ then

$$\theta^{n-|\mathbf{a}|}\left(\phi_{\mathbf{a}}(\mathbf{X}) - E(\phi_{\mathbf{a}}(\mathbf{X}))\right) \Rightarrow 0 \tag{75}$$

in probability as $\theta \to \infty$. Note that if $\mathbf{a} \neq \mathbf{a}_1, \mathbf{a}_2$ then $n - |\mathbf{a}| \geq 2$. Note also that $\mathrm{ESF}(\theta, \mathbf{a}_2) = O(\theta^{-3/2})$. Therefore, it follows from equation (75) that

$$\sum_{\mathbf{a}\neq\mathbf{a}_1,\mathbf{a}_2} \frac{\phi_{\mathbf{a}}(\mathbf{X}) - E(\phi_{\mathbf{a}}(\mathbf{X}))}{E(\phi_{\mathbf{a}_2}(\mathbf{X}))} \Rightarrow 0 \tag{76}$$

as $\theta \to \infty$. This gives the limit in distribution

$$\lim_{\theta\to\infty} \frac{\phi_{\mathbf{a}_1}(\mathbf{X}) - E\left(\phi_{\mathbf{a}_1}(\mathbf{X})\right)}{E\left(\phi_{\mathbf{a}_2}(\mathbf{X})\right)} = \lim_{\theta\to\infty} \frac{E\left(\phi_{\mathbf{a}_2}(\mathbf{X})\right) - \phi_{\mathbf{a}_2}(\mathbf{X})}{E\left(\phi_{\mathbf{a}_2}\mathbf{X}\right)} = -Z_2 \stackrel{d}{=} Z_2. \qquad \square$$

REFERENCES

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. Wiley, New York.

ETHIER, S. N. and KURTZ, T. G. (1994). Convergence to Fleming–Viot processes in the weak atomic topology. *Stochastic Process. Appl.* **54** 1–27.

EWENS, W. J. (1979). *Mathematical Population Genetics*. Springer, Berlin.

DONNELLY, P. and JOYCE, P. (1989). Continuity and weak convergence of ranked and size-biased permutations on the infinite simplex. *Stochastic Process. Appl.* **31** 89–103.

GILLESPIE, J. H. (1999). The role of population size in molecular evolution. *Theoret. Population Biology* **55** 145–156.

GRIFFITHS, R. C. (1979). On the distribution of allele frequencies in a diffusion model. *Theoret. Population Biology* **15** 140–158.

JOYCE, P. (1998). Partition structures and sufficient statistics. *J. Appl. Probab.* **35** 622–632.

JOYCE, P., KRONE, S. and KURTZ, T. G. (2000). A ratio limit theorem for the infinite alleles model with overdominant selection. Unpublished manuscript.

KARATZAS, I. and SHREVE, S. (1988). *Brownian Motion and Stochastic Calculus*. Springer, New York.

KINGMAN, J. F. C. (1977). The population structure associated with the Ewens sampling formula. *Theoret. Population Biology* **11** 274–283.

KURTZ, T. G. and PROTTER, P. (1991). Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.* **19** 1035–1070.

WATTERSON, G. A. and GUESS, H. A. (1977) Is the most frequent allele the oldest? *Theoret. Population Biology* **11** 141–160.

P. JOYCE
S. M. KRONE
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF IDAHO
MOSCOW, IDAHO 83844
E-MAIL: joyce@uidaho.edu
krone@uidaho.edu

T. G. KURTZ
DEPARTMENTS OF MATHEMATICS
AND STATISTICS
UNIVERSITY OF WISCONSIN
MADISON, WISCONSIN 53706–1388
E-MAIL: kurtz@math.wisc.edu