

WHEN CAN ONE DETECT OVERDOMINANT SELECTION IN THE INFINITE-ALLELES MODEL?

BY PAUL JOYCE,¹ STEPHEN M. KRONE¹ AND THOMAS G. KURTZ²

University of Idaho, University of Idaho and University of Wisconsin

One of the goals of this paper is to show that the infinite-alleles model with overdominant selection “looks like” the neutral infinite-alleles model when the selection intensity and mutation rate get large together. This rather surprising behavior was noticed by Gillespie (1999) in simulations. To make rigorous and refine Gillespie’s observations, we analyze the limiting behavior of the likelihood ratio of the stationary distributions for the model under selection and neutrality, as the mutation rate and selection intensity go to ∞ together in a specified manner. In particular, we show that the likelihood ratio tends to 1 as the mutation rate goes to ∞ , provided the selection intensity is a multiple of the mutation rate raised to a power less than $3/2$. (Gillespie’s simulations correspond to the power 1.) This implies that we cannot distinguish between the two models in this setting. Conversely, if the selection intensity grows like a multiple of the mutation rate raised to a power greater than $3/2$, selection can be detected; that is, the likelihood ratio tends to 0 under neutrality and ∞ under selection. We also determine the nontrivial limit distributions in the case of the critical exponent $3/2$. We further analyze the limiting behavior when the exponent is less than $3/2$ by determining the rate at which the likelihood ratio converges to 1 and by developing results for the distributions of finite samples.

1. Preliminaries and main results. In this paper we attempt to explain and refine some observations of Gillespie (1999), which were based on simulations of an infinite-alleles model with selective overdominance (as well as other models). Gillespie made the surprising observation that, if the selection intensity and mutation rate get large together, the behavior looks like that of a neutral model. Gillespie argued that, for a K allele model, high mutation rate and high selection intensity tend to push the population toward equal frequencies, implying that the population frequencies will tend to $(1/K, \dots, 1/K)$ under both selection and neutrality. This argument does not extend to the infinite-alleles model, so simulations were used. His simulations show that various measures of gene diversity look increasingly similar as the mutation rate and selection intensity are increased. By analyzing the infinite-alleles diffusion, we are able to justify Gillespie’s claims and, in fact, to go quite a bit farther.

Received December 2001; revised April 2002.

¹Supported in part by Grants DMS-96-26764, DMS-00-72198 and EPS-00-80935.

²Supported in part by NSF Grant DMS-99-71571.

AMS 2000 subject classifications. Primary 62F05; secondary 60G42, 60F15.

Key words and phrases. Infinite-alleles model, overdominant selection, homozygosity, mutation rate, likelihood ratio, Ewens sampling formula, Poisson–Dirichlet, weak limits.

Our approach is based on likelihood ratios for the stationary measures under selection and neutrality. We begin with a description of the diploid infinite-alleles model with selective overdominance (i.e., heterozygote advantage). This is a diffusion process obtained via a large-population limit as follows. Consider a diploid population of M individuals in which each of the $2M$ genes is assigned a type $x \in [0, 1]$. Suppose the genes in the current generation are labeled x_1, \dots, x_{2M} . To determine the genotype of a new individual in the next generation, a pair of parental genes is selected at random from the population. The probability that the i th and j th genes are selected is

$$\frac{w_M(x_i, x_j)}{\sum_{1 \leq l, m \leq 2M} w_M(x_l, x_m)},$$

where $w_M(x, y) \geq 0$ is a symmetric function specifying the fitness of an individual with genotype (x, y) . To model selective overdominance, we give homozygotes a fitness disadvantage by setting

$$w_M(x, y) = 1 - \frac{\sigma}{4M} \delta_{x,y},$$

where $\sigma > 0$ is a constant (the selection intensity) and $\delta_{x,y}$ is the Kronecker delta. With probability $u_M = \theta/(4M)$, one of the genes in the new individual will be selected at random to mutate, and the new mutant type will be chosen according to a uniform distribution on $[0, 1]$; with probability $1 - u_M$, it will be different from all previous types. With probability $1 - u_M$, there is no mutation and the genotype of the new individual is determined by the parental genes. Letting $M \rightarrow \infty$ and keeping track of the proportions of the different types of genes, one obtains a (measure-valued) diffusion process known as the infinite-alleles model with overdominant selection [cf. Ethier and Kurtz (1994) for more details on the above construction and process]. Our interest here is in the stationary distribution of this process.

Consider the ordered infinite simplex

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

Here, we think of x_i as representing the relative frequency of the i th most frequent allele. When these frequencies are considered to be random, we will write $\mathbf{X} = (X_1, X_2, \dots) \in \nabla$ for the random vector of allele relative frequencies in the population, in descending order. Let μ_σ be the stationary distribution for \mathbf{X} in the infinite-alleles model with selection intensity σ , and \mathbb{E}_σ the corresponding expectation. In particular, μ_0 and \mathbb{E}_0 give the stationary distribution and expectation in the neutral case, $\sigma = 0$.

In the neutral case, if we keep track of the proportions of the alleles in descending order, the stationary distribution, μ_0 , is given by the Poisson–Dirichlet

distribution with parameter θ [cf. Kingman (1977)]. This is a probability measure on ∇ and, to make explicit the dependence on θ , we write $\mathbf{X}_\theta \sim \mu_\theta$ when $\mathbf{X}_\theta \in \nabla$ has this distribution.

The form of the stationary distribution μ_σ , in the case of overdominant selection with intensity σ , follows as a special case of Theorem 4.4 in Ethier and Kurtz (1994): for integrable f ,

$$(1) \quad \int_{\nabla} f(\mathbf{x}) \mu_\sigma(d\mathbf{x}) = \int_{\nabla} f(\mathbf{x}) \frac{e^{-\sigma H(\mathbf{x})}}{\mathcal{Z}(\sigma, \theta)} \mu_0(d\mathbf{x}),$$

where

$$(2) \quad H(\mathbf{x}) = H(x_1, x_2, \dots) \equiv \sum_{i=1}^{\infty} x_i^2$$

is the *population homozygosity* when the allele frequencies are given by $\mathbf{x} = (x_1, x_2, \dots) \in \nabla$, and

$$(3) \quad \mathcal{Z}(\sigma, \theta) = \int_{\nabla} e^{-\sigma H(\mathbf{x})} \mu_0(d\mathbf{x})$$

is a normalizing constant. Thus,

$$\frac{d\mu_\sigma}{d\mu_0}(\mathbf{x}) = \frac{e^{-\sigma H(\mathbf{x})}}{\mathcal{Z}(\sigma, \theta)}$$

is the Radon–Nikodym derivative (or *likelihood ratio*) for the stationary measure under selection with respect to the stationary measure under neutrality.

We are interested in the situation where the selection intensity σ goes to ∞ like a power of θ . As we will see, the exponent $3/2$ plays a very special role. Thus, we write $\sigma = c\theta^{3/2+\gamma}$, where c is a positive constant and $\gamma \in (-\infty, \infty)$ is also held constant. We will write $\mathbf{Y}_\theta^{(\gamma)} \sim \mu_\sigma$ when $\mathbf{Y}_\theta^{(\gamma)} \in \nabla$ has this distribution. Sometimes, when dealing with the case $\gamma < 0$, we will instead write $\mathbf{Y}_\theta^{(-\alpha)} \sim \mu_\sigma$, $\sigma = c\theta^{3/2-\alpha}$, and so forth, where $\alpha > 0$. Our first results concern the asymptotic behavior of the above likelihood ratio. Here and throughout the paper, we denote weak convergence by \Rightarrow .

THEOREM 1. *Suppose $\mathbf{X}_\theta = (X_{1,\theta}, X_{2,\theta}, \dots) \sim \mu_\theta$ and $\sigma = c\theta^{3/2+\gamma}$, where $c > 0$ is a constant. Then, as $\theta \rightarrow \infty$,*

$$\frac{d\mu_\sigma}{d\mu_0}(\mathbf{X}_\theta) = \frac{\exp\{-\sigma H(\mathbf{X}_\theta)\}}{\mathbb{E}_0(\exp\{-\sigma H(\mathbf{X}_\theta)\})} \Rightarrow \begin{cases} 1, & \text{if } \gamma < 0, \\ \exp(cZ_2 - c^2), & \text{if } \gamma = 0, \\ 0, & \text{if } \gamma > 0, \end{cases}$$

where $Z_2 \sim N(0, 2)$.

THEOREM 2. Suppose $\mathbf{X}_\theta = (X_{1,\theta}, X_{2,\theta}, \dots) \sim \mu_0$, $\mathbf{Y}_\theta^{(\gamma)} = (Y_{1,\theta}^{(\gamma)}, Y_{2,\theta}^{(\gamma)}, \dots) \sim \mu_\sigma$, where $\sigma = c\theta^{3/2+\gamma}$ and $c > 0$ is a constant. Then, as $\theta \rightarrow \infty$,

$$\frac{d\mu_\sigma}{d\mu_0}(\mathbf{Y}_\theta^{(\gamma)}) = \frac{\exp\{-\sigma H(\mathbf{Y}_\theta^{(\gamma)})\}}{\mathbb{E}_0(\exp\{-\sigma H(\mathbf{X}_\theta)\})} \Rightarrow \begin{cases} 1, & \text{if } \gamma < 0, \\ \exp(cZ_2^* - c^2), & \text{if } \gamma = 0, \\ \infty, & \text{if } \gamma > 0, \end{cases}$$

where $Z_2^* \sim N(2c, 2)$.

These two results combine to say that, if the selection intensity grows like θ to a power greater than $3/2$, then, under neutrality, the likelihood ratio $(d\mu_\sigma/d\mu_0)(\mathbf{X}_\theta)$ goes to 0, while, under selection, $(d\mu_\sigma/d\mu_0)(\mathbf{Y}_\theta)$ goes to ∞ . This says that, for large θ , one can distinguish between a neutral population and a population with selection intensity very large relative to the mutation rate. If the selection intensity grows like θ to a power less than $3/2$, one cannot distinguish between the neutral and selective models when θ is large; that is, one cannot detect selective overdominance. Loosely speaking, the model with $\sigma = c\theta^\beta$ undergoes a phase transition as the exponent β crosses $3/2$. Notice that, in the critical case, $\sigma = c\theta^{3/2}$, the limiting likelihood ratio has a lognormal distribution under both neutrality and selection. Notice also that Z_2^* is larger, on average, than Z_2 . As expected, in the limit, the likelihood ratio tends to be smaller under neutrality than under selection.

When $\sigma = c\theta^{3/2+\gamma}$, with $\gamma < 0$, selection is masked by a high mutation rate. In the limit, one cannot distinguish between neutrality and selection and the likelihood ratio tends to 1, regardless of whether the population frequencies were generated from neutral or selective models. The next theorem gives the rate at which the likelihood ratio tends to 1 in this case.

THEOREM 3. Suppose $\mathbf{X}_\theta = (X_{1,\theta}, X_{2,\theta}, \dots) \sim \mu_0$, $\mathbf{Y}_\theta^{(-\alpha)} = (Y_{1,\theta}^{(-\alpha)}, Y_{2,\theta}^{(-\alpha)}, \dots) \sim \mu_\sigma$, where $\sigma = c\theta^{3/2-\alpha}$ and $\alpha > 0$ and $c > 0$ are constants. Then, as $\theta \rightarrow \infty$,

$$(4) \quad \theta^\alpha \left(\frac{d\mu_\sigma}{d\mu_0}(\mathbf{X}_\theta) - 1 \right) \Rightarrow cZ_2$$

and

$$(5) \quad \theta^\alpha \left(\frac{d\mu_\sigma}{d\mu_0}(\mathbf{Y}_\theta^{(-\alpha)}) - 1 \right) \Rightarrow cZ_2,$$

where $Z_2 \sim N(0, 2)$.

The case in which σ grows like $c\theta = c\theta^{3/2-1/2}$ in the above theorems corresponds to the simulations in Gillespie (1999). To make clear the connection between this paper and Gillespie (1999), recall that $\theta = 4Mu$, where u is the probability of mutation per individual, and $w = 1 - \sigma/(4M)$ is the fitness for

homozygotes. Therefore, as $M \rightarrow \infty$, both θ and $\sigma = (1 - w)4M$ tend to ∞ ; moreover,

$$\sigma = \frac{1 - w}{u} \theta.$$

In Gillespie's simulations, u and w were fixed as $M \rightarrow \infty$. This implies that σ increases linearly in θ , with $c = (1 - w)/u$. The above results make the point that, even if the entire population were observable, the neutral model would be indistinguishable from the selective overdominance model when the selection intensity is not sufficiently larger than θ .

In practice, one cannot observe the entire population. Instead, one must make inferences based on information in a *finite sample*. In the next two theorems, we see that, if we restrict attention to a finite sample when $\sigma = c\theta^{3/2+\gamma}$ and $\gamma < 0$, then the likelihood ratio converges to 1 faster than in the case in which the entire population is observed. To consider a finite sample of genes drawn from a population that has evolved under the infinite-alleles model, it is convenient to use the notion of a partition structure.

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 $\mathbb{P}(\mathbf{A}_n = \mathbf{a} | \mathbf{X} = \mathbf{x})$ is given by the *multinomial sampling function* [cf. Kingman (1977)]:

$$(6) \quad \mathbb{P}(\mathbf{A}_n = \mathbf{a} | \mathbf{X} = \mathbf{x}) = \phi_{\mathbf{a}}(\mathbf{x}) \equiv \frac{n!}{\prod_{i=1}^n (i!)^{a_i}} \sum_{\mathbf{v} \in \mathcal{A}_{\mathbf{a}}} x_1^{v_1} x_2^{v_2} \dots,$$

where $\mathcal{A}_{\mathbf{a}} \equiv \{\mathbf{v} = (v_1, v_2, \dots) \in \mathbb{Z}_+ \times \mathbb{Z}_+ \times \dots : \#(i : v_i = j) = a_j\}$ represents the set of allele counts that 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 ,

$$(7) \quad \mathbb{P}_0(\mathbf{A}_n = \mathbf{a}) = \int_{\nabla} \phi_{\mathbf{a}}(\mathbf{x}) \mu_0(d\mathbf{x}) = \mathbb{E}_0(\phi_{\mathbf{a}}(\mathbf{X}_\theta)),$$

is given by the *Ewens sampling formula*

$$(8) \quad \mathbb{P}_0(\mathbf{A}_n = \mathbf{a}) = \text{ESF}(\theta, \mathbf{a}) \equiv \frac{n!}{\theta_{(n)}} \prod_{j=1}^n \left(\frac{\theta}{j}\right)^{a_j} \frac{1}{a_j!},$$

where $\theta_{(n)} = \theta(\theta + 1) \dots (\theta + n - 1)$. In the selective case, $\sigma > 0$, the stationary sampling distribution is given by

$$(9) \quad \mathbb{P}_\sigma(\mathbf{A}_n = \mathbf{a}) = \int_{\nabla} \phi_{\mathbf{a}}(\mathbf{x}) \mu_\sigma(d\mathbf{x}) = \mathbb{E}_0\left(\phi_{\mathbf{a}}(\mathbf{X}_\theta) \frac{d\mu_\sigma}{d\mu_0}(\mathbf{X}_\theta)\right).$$

THEOREM 4. *Let $c > 0$ and $\alpha \geq 0$ be constants and suppose $\sigma = c\theta^{3/2-\alpha}$. If $\mathbf{a} = (a_1, a_2, \dots, a_n)$ is an allelic partition for a sample of size n with $a_1 \neq n$, then*

$$(10) \quad \lim_{\theta \rightarrow \infty} \theta^{\alpha+1/2} \left| \frac{\mathbb{P}_\sigma(\mathbf{A}_n = \mathbf{a})}{\mathbb{P}_0(\mathbf{A}_n = \mathbf{a})} - 1 \right| = 2c \sum_{i=2}^n a_i \binom{i}{2}.$$

REMARKS. Note that the sum $\sum_{i=2}^n a_i \binom{i}{2}$ appearing in the limit gives the number of homozygous pairs that can be formed from the n alleles in the sample. The above theorem says that, if we restrict attention to a finite sample, the convergence is of order $\sqrt{\theta}$ faster than that given in Theorem 3. It also says that, unlike the situation in Theorem 3, the likelihood ratio converges to one in the case of the critical exponent $3/2$.

Again, the case with $\alpha = 1/2$ covers the Gillespie conjecture in the context of random samples. Returning to Gillespie's notation, let $\theta = 4Mu$ and $\sigma = 4M(1-w)$. As the effective population size M becomes large, the likelihood ratio of a sample under selection relative to a sample under neutrality converges to 1 at a rate proportional to $1/M$.

The case $\mathbf{a} = (n, 0, \dots, 0)$, not covered by Theorem 4, corresponds to all alleles in the sample being distinct. This is the sample configuration with the lowest possible homozygosity. Since large values of θ and σ in both the neutral and selective models give increasingly higher probability to samples for which $a_1 = n$, it is natural to expect that the two models will look even more similar at this most probable sample configuration; cf. (8). This suggests that the rate of convergence of $\mathbb{P}_\sigma(\mathbf{A}_n = \mathbf{a})/\mathbb{P}_0(\mathbf{A}_n = \mathbf{a})$ to 1 is even faster in this case. This is the content of the next theorem.

THEOREM 5. *Let $\alpha > 0$ and $c > 0$ be constants. Suppose $\sigma = c\theta^{3/2-\alpha}$ and let $\mathbf{a}_1 = (n, 0, \dots, 0)$ be the allelic partition corresponding to a sample of n distinct alleles. Then*

$$(11) \quad \lim_{\theta \rightarrow \infty} \theta^{\alpha+3/2} \left| \frac{\mathbb{P}_\sigma(\mathbf{A}_n = \mathbf{a}_1)}{\mathbb{P}_0(\mathbf{A}_n = \mathbf{a}_1)} - 1 \right| = 2c.$$

2. Some normal limits. Most of the results in this paper depend on normal limit theorems for population homozygosities and sampling functions. In this section, we recall some such results from Joyce, Krone and Kurtz (2001) and prove some needed extensions.

If the decreasing allele frequencies in the population are given by $\mathbf{x} = (x_1, x_2, \dots)$, define, for $m \geq 2$, the m th-order population homozygosity by

$$(12) \quad H_m(\mathbf{x}) \equiv \sum_{i=1}^{\infty} x_i^m.$$

This can be thought of as the probability that all individuals in a sample of size m are of the same type, given that the vector of population allele frequencies is \mathbf{x} . Thus, (12) is a special case of (6). Of course, $H_2(\mathbf{x}) = H(\mathbf{x})$ is the usual population homozygosity.

Next, define, for $m \geq 2$, the following scaled versions of the m th-order population homozygosities when the ordered allele frequencies are distributed according to the stationary distributions under neutrality and selection, respectively. If $\mathbf{X}_\theta \sim \mu_0$, put

$$(13) \quad Z_{m,\theta} \equiv \sqrt{\theta} \left(\frac{\theta^{m-1} H_m(\mathbf{X}_\theta) - (m-1)!}{(m-1)!} \right),$$

and if $\mathbf{Y}_\theta^{(\gamma)} \sim \mu_\sigma$, where $\sigma = c\theta^{3/2+\gamma}$, put

$$(14) \quad Z_{m,\theta}^{(\gamma)} \equiv \sqrt{\theta} \left(\frac{\theta^{m-1} H_m(\mathbf{Y}_\theta^{(\gamma)}) - (m-1)!}{(m-1)!} \right).$$

Note that, instead of centering by the means and dividing by the standard deviations in the definitions of $Z_{m,\theta}$ and $Z_{m,\theta}^{(\gamma)}$, we scale by quantities that are asymptotically the same (when θ is large), but easier to work with.

It is easy to see that, for any $\gamma \in (-\infty, \infty)$ and $\sigma = c\theta^{3/2+\gamma}$,

$$(15) \quad \frac{d\mu_\sigma}{d\mu_0}(\mathbf{X}_\theta) = \frac{\exp\{-\sigma H(\mathbf{X}_\theta)\}}{\mathbb{E}_0(\exp\{-\sigma H(\mathbf{X}_\theta)\})} = \frac{\exp\{-c\theta^\gamma Z_{2,\theta}\}}{\mathbb{E}_0(\exp\{-c\theta^\gamma Z_{2,\theta}\})}$$

and

$$(16) \quad \frac{d\mu_\sigma}{d\mu_0}(\mathbf{Y}_\theta^{(\gamma)}) = \frac{\exp\{-\sigma H(\mathbf{Y}_\theta^{(\gamma)})\}}{\mathbb{E}_0(\exp\{-\sigma H(\mathbf{X}_\theta)\})} = \frac{\exp\{-c\theta^\gamma Z_{2,\theta}^{(\gamma)}\}}{\mathbb{E}_0(\exp\{-c\theta^\gamma Z_{2,\theta}^{(\gamma)}\})}.$$

In the next lemma, we show that $Z_{m,\theta}$ and $Z_{m,\theta}^{(\gamma)}$ have normal limit laws as $\theta \rightarrow \infty$. These weak limits will play a crucial role in the proofs of the theorems described above.

LEMMA 1. *Let (Z_2, Z_3, \dots, Z_n) be multivariate normal with mean $\mathbf{0}$ and*

$$(17) \quad \text{Cov}(Z_i, Z_j) = \frac{(i+j-1)! - i!j!}{(i-1)!(j-1)!}, \quad i, j \in \{2, 3, \dots, n\},$$

and let $Z_{m,\theta}$ and $Z_{m,\theta}^{(\gamma)}$ be defined as in (13) and (14). Then, as $\theta \rightarrow \infty$, we have the following limits:

(i) *If $\mathbf{X}_\theta \sim \mu_0$, then*

$$(18) \quad (Z_{2,\theta}, Z_{3,\theta}, \dots, Z_{n,\theta}) \Rightarrow (Z_2, Z_3, \dots, Z_n).$$

(ii) If $\mathbf{Y}_\theta^{(-\alpha)} \sim \mu_\sigma$, with $\sigma = c\theta^{3/2-\alpha}$, then

$$(19) \quad (Z_{2,\theta}^{(-\alpha)}, Z_{3,\theta}^{(-\alpha)}, \dots, Z_{n,\theta}^{(-\alpha)}) \Rightarrow \begin{cases} (Z_2^*, Z_3^*, \dots, Z_n^*), & \text{if } \alpha = 0, \\ (Z_2, Z_3, \dots, Z_n), & \text{if } \alpha > 0, \end{cases}$$

where $\mathbf{Z}^* = \mathbf{Z} - c\mathbf{b}$ and $\mathbf{b} = (2, 3 \cdot 2, \dots, n(n-1))$.

PROOF. The neutral case, given by (18), is proved in Joyce, Krone and Kurtz (2001). To establish (19) in the selective case with $\sigma = c\theta^{3/2}$, we can use (18) together with (15), with $\gamma = 0$, to conclude for any $\mathbf{z}^* \in \mathbb{R}^{n-1}$ that

$$\begin{aligned} \mathbb{P}_\sigma(\mathbf{Z}_\theta^{(0)} \leq \mathbf{z}^*) &= \mathbb{E}_0\left(\frac{d\mu_\sigma}{d\mu_0}(\mathbf{X}_\theta)\mathbf{1}\{\mathbf{Z}_\theta \leq \mathbf{z}^*\}\right) \\ &= \mathbb{E}_0\left(\frac{\exp\{-cZ_{2,\theta}\}}{\mathbb{E}_0(\exp\{-cZ_{2,\theta}\})}\mathbf{1}\{\mathbf{Z}_\theta \leq \mathbf{z}^*\}\right) \\ &\rightarrow \mathbb{E}\left(\frac{\exp\{-cZ_2\}}{\mathbb{E}_0(\exp\{-cZ_2\})}\mathbf{1}\{\mathbf{Z} \leq \mathbf{z}^*\}\right) \end{aligned}$$

as $\theta \rightarrow \infty$. To get convergence in the last step, it is not enough that $\mathbf{Z}_\theta \Rightarrow \mathbf{Z}$ as $\theta \rightarrow \infty$. In addition, we need to show that

$$(20) \quad \mathbb{E}_0(\exp\{-cZ_{2,\theta}\}) \Rightarrow \mathbb{E}(\exp\{-cZ_2\}).$$

This requires a uniform integrability argument, which will be given in Lemma 9 in Section 4.

Assuming this to hold, the rest of the proof in the case $\alpha = 0$ follows from standard algebraic manipulations involving multivariate normal distributions. Indeed, let Σ be the covariance matrix for \mathbf{Z} and write $\mathbf{c}' = (c, 0, \dots, 0)$, the prime denoting transpose. Then

$$\begin{aligned} &\mathbb{E}\left(\frac{\exp\{-cZ_2\}}{\mathbb{E}_0(\exp\{-cZ_2\})}\mathbf{1}\{\mathbf{Z} \leq \mathbf{z}^*\}\right) \\ &= \frac{\exp\{-c^2\}}{(2\pi)^{n/2}\sqrt{\det \Sigma}} \int_{\mathbf{z} \leq \mathbf{z}^*} \exp\{-cz_2\} \exp\left\{\frac{-\mathbf{z}'\Sigma^{-1}\mathbf{z}}{2}\right\} d\mathbf{z} \\ &= \frac{\exp\{-c^2\}}{(2\pi)^{n/2}\sqrt{\det \Sigma}} \int_{\mathbf{z} \leq \mathbf{z}^*} \exp\{-\mathbf{c}'\mathbf{z}\} \exp\left\{\frac{-\mathbf{z}'\Sigma^{-1}\mathbf{z}}{2}\right\} d\mathbf{z} \\ &= \frac{1}{(2\pi)^{n/2}\sqrt{\det \Sigma}} \int_{\mathbf{z} \leq \mathbf{z}^*} \exp\left\{-\frac{1}{2}(2\mathbf{c}'\mathbf{z} + \mathbf{z}'\Sigma^{-1}\mathbf{z} + \mathbf{c}'\Sigma\mathbf{c})\right\} d\mathbf{z} \\ &= \frac{1}{(2\pi)^{n/2}\sqrt{\det \Sigma}} \int_{\mathbf{z} \leq \mathbf{z}^*} \exp\left\{-\frac{1}{2}(\mathbf{z}' + \mathbf{c}'\Sigma)\Sigma^{-1}(\mathbf{z} + \Sigma\mathbf{c})\right\} d\mathbf{z} \\ &= \mathbb{P}(\mathbf{Z}^* \leq \mathbf{z}^*). \end{aligned}$$

The remaining case is much simpler. Use $\gamma = -\alpha < 0$ in (15) and compute the limit as above with the change

$$(21) \quad \frac{d\mu_\sigma}{d\mu_0}(\mathbf{X}_\theta) = \frac{\exp\{-c\theta^{-\alpha}Z_{2,\theta}\}}{\mathbb{E}_0(\exp\{-c\theta^{-\alpha}Z_{2,\theta}\})} \Rightarrow \frac{e^0}{\mathbb{E}e^0} = 1$$

as $\theta \rightarrow \infty$. \square

The next lemma is a restatement of Theorem 2 in Joyce, Krone and Kurtz (2001). It is essentially the analogue of (18) for finite sampling distributions.

LEMMA 2. *Suppose $\mathbf{X}_\theta \sim \mu_0$ and (Z_2, \dots, Z_n) is as in Lemma 1 and let \mathbf{a} be a partition corresponding to a sample of size n drawn from a population with ordered frequencies \mathbf{X}_θ .*

(i) *If $\mathbf{a} \neq (n, 0, \dots, 0)$, then, as $\theta \rightarrow \infty$,*

$$(22) \quad \sqrt{\theta} \left(\frac{\phi_{\mathbf{a}}(\mathbf{X}_\theta) - \text{ESF}(\theta, \mathbf{a})}{\text{ESF}(\theta, \mathbf{a})} \right) \Rightarrow \sum_{i=2}^n a_i Z_i.$$

(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, as $\theta \rightarrow \infty$,*

$$(23) \quad \sqrt{\theta} \left(\frac{\phi_{\mathbf{a}_1}(\mathbf{X}_\theta) - \text{ESF}(\theta, \mathbf{a}_1)}{\text{ESF}(\theta, \mathbf{a}_2)} \right) \Rightarrow -Z_2.$$

3. Two basic lemmas. The following lemmas, which may apply in a variety of settings, will be used to prove Theorems 1 and 2. More specifically, Lemma 3 is used to prove the last part of Lemma 4, and the latter provides the form of the limits in Theorems 1 and 2.

LEMMA 3. *Suppose (Y_n) is a sequence of random variables satisfying $Y_n \Rightarrow Y$, where Y is a continuous random variable whose density is strictly positive on all of $(-\infty, \infty)$. Suppose farther that $\mathbb{E}(e^{-sY_n}) < \infty$ for all $s \in (0, \infty)$. Then, for any $\alpha > 0$,*

$$(24) \quad \lim_{n \rightarrow \infty} \frac{\mathbb{E}(\sqrt{e^{-n^\alpha Y_n}})}{\sqrt{\mathbb{E}(e^{-n^\alpha Y_n})}} = 0.$$

PROOF. We begin by noting that the exponential moment condition in the hypotheses is simply to ensure that all the expected values in (24) are finite. Let $F_n(x) = \mathbb{P}(Y_n \leq x)$ and define $F_n^{-1}(y) = \min\{x : F_n(x) \geq y\}$ for $y \in [0, 1]$. It is easy to see that $\mathbb{P}(Y_n \leq F_n^{-1}(y)) \geq y$. Defining $-y_n = F_n^{-1}(1/n)$, it is clear that

$-y_n \rightarrow -\infty$ as $n \rightarrow \infty$ and that $\mathbb{P}(Y_n \leq -y_n) \geq 1/n$. Therefore, by the Cauchy-Schwarz inequality,

$$\begin{aligned}
& \frac{\mathbb{E}(\exp\{-n^\alpha Y_n/2\})}{\sqrt{\mathbb{E}(\exp\{-n^\alpha Y_n\})}} \\
&= \frac{\mathbb{E}(\exp\{-n^\alpha Y_n/2\}\mathbf{1}\{|Y_n| \leq \sqrt{y_n}\})}{\sqrt{\mathbb{E}(\exp\{-n^\alpha Y_n\})}} + \frac{\mathbb{E}(\exp\{-n^\alpha Y_n/2\}\mathbf{1}\{|Y_n| > \sqrt{y_n}\})}{\sqrt{\mathbb{E}(\exp\{-n^\alpha Y_n\})}} \\
&\leq \frac{\exp(n^\alpha \sqrt{y_n}/2)}{\sqrt{\mathbb{E}(\exp\{-n^\alpha Y_n\})}} + \sqrt{\mathbb{P}(|Y_n| > \sqrt{y_n})} \\
&\leq \frac{\exp(n^\alpha \sqrt{y_n}/2)}{\sqrt{\mathbb{E}(\exp\{-n^\alpha Y_n\}\mathbf{1}\{Y_n \leq -y_n\})}} + \sqrt{\mathbb{P}(|Y_n| > \sqrt{y_n})} \\
&\leq \frac{\exp(n^\alpha \sqrt{y_n}/2)}{\sqrt{\exp(n^\alpha y_n)\mathbb{P}(Y_n \leq -y_n)}} + \sqrt{\mathbb{P}(|Y_n| > \sqrt{y_n})} \\
&\leq \frac{\exp(n^\alpha \sqrt{y_n}/2)}{\exp(n^\alpha y_n/2)\sqrt{(1/n)}} + \sqrt{\mathbb{P}(|Y_n| > \sqrt{y_n})} \\
&\leq \sqrt{n} \exp(-n^\alpha \sqrt{y_n}/2) + \sqrt{\mathbb{P}(|Y_n| > \sqrt{y_n})} \rightarrow 0
\end{aligned}$$

as $n \rightarrow \infty$. \square

LEMMA 4. *Suppose (Y_n) is a sequence of random variables satisfying $Y_n \Rightarrow Y$, where Y is a continuous random variable whose density is strictly positive on all of $(-\infty, \infty)$. Suppose further that $\mathbb{E}(e^{-sY_n}) < \infty$ for all $s \in (0, \infty)$ and that $\sup_n \mathbb{E}(e^{-tY_n}) < \infty$ for some $t > 1$. Then, as $n \rightarrow \infty$,*

$$(25) \quad \frac{e^{-n^\gamma Y_n}}{\mathbb{E}(e^{-n^\gamma Y_n})} \Rightarrow \begin{cases} 1, & \text{if } \gamma < 0, \\ \frac{e^{-Y}}{\mathbb{E}(e^{-Y})}, & \text{if } \gamma = 0, \\ 0, & \text{if } \gamma > 0. \end{cases}$$

PROOF. Starting with the case $\gamma = 0$, note that $e^{-Y_n} \Rightarrow e^{-Y}$ by the continuous mapping theorem. The hypothesis, $\sup_n \mathbb{E}(e^{-tY_n}) < \infty$ for some $t > 1$, implies uniform integrability of $\{e^{-Y_n}\}$, and this implies $\lim_{n \rightarrow \infty} \mathbb{E}(e^{-Y_n}) = \mathbb{E}(e^{-Y})$. If $\gamma < 0$, then $n^\gamma \rightarrow 0$, and hence $n^\gamma Y_n \Rightarrow 0$. The continuous mapping theorem implies $\exp(-n^\gamma Y_n) \Rightarrow 1$. In addition, $\gamma < 0$ implies $\exp(-n^\gamma Y_n) \leq \max\{\exp(-Y_n), 1\}$. So, $\mathbb{E}[\exp(-n^\gamma Y_n)] \rightarrow 1$ follows from the dominated convergence theorem. Finally, in the case $\gamma > 0$, set $U_n = e^{-n^\gamma Y_n} / \mathbb{E}(e^{-n^\gamma Y_n})$. An appeal to Lemma 3 yields $\mathbb{E}(\sqrt{U_n}) \rightarrow 0$ as $n \rightarrow \infty$. Therefore, $\sqrt{U_n} \Rightarrow 0$, implying $U_n \Rightarrow 0$. \square

4. Uniform integrability. This section contains uniform integrability results that are essential for proving the main theorems in Section 1. Since this is a long and technical section, it is probably a good idea for the reader to simply look at the statements of Lemmas 5 and 9 and then move on to the rest of the paper to see how they are used before coming back to digest the proofs.

To help motivate this section, we begin by alerting the reader that the condition, $\sup_n \mathbb{E}(e^{-tY_n}) < \infty$ for some $t > 1$, in Lemma 4, is an important one. To illustrate this point for the case $\gamma = 0$, consider the following example for which this condition does not hold. Let Y_n be defined via its density

$$f_n(y) = \left(1 - \frac{1}{n}\right) \frac{e^{-y^2/2}}{\sqrt{2\pi}} + \frac{1}{n} \frac{e^{-(y+\log n)^2/2}}{\sqrt{2\pi}}, \quad y \in (-\infty, \infty).$$

Then Y_n converges in distribution to the standard normal Z , and hence e^{-Y_n} converges to the lognormal e^{-Z} . Now, using moment generating functions of normal random variables, we can see that

$$\mathbb{E}(e^{-tY_n}) = \left(1 - \frac{1}{n}\right) e^{t^2/2} + \frac{1}{n} e^{t \log n + t^2/2} = \left(1 - \frac{1}{n}\right) e^{t^2/2} + n^{t-1} e^{t^2/2}.$$

So, even though $\mathbb{E}(Y_n^r) \rightarrow \mathbb{E}(Z^r)$ for all $r \geq 1$, we have $\mathbb{E}(e^{-Y_n}) \rightarrow 2e^{1/2} \neq \mathbb{E}(e^{-Z})$. The missing ingredient, which would have allowed for the interchange of limit with expectation, is uniform integrability of the family $\{e^{-Y_n}\}$; this would have followed if $\sup_n \mathbb{E}(e^{-tY_n}) < \infty$ for some $t > 1$. However, the above calculation shows that $\mathbb{E}(e^{-tY_n}) \rightarrow \infty$, as $n \rightarrow \infty$, for all $t > 1$.

Thus motivated, we seek to establish uniform integrability for certain functions of the $Z_{m,\theta}$'s, the uniformity being in θ . First, we show $\sup_{\theta > 1} \mathbb{E}(Z_{m,\theta}^r) < \infty$ for all $r \geq 1$ (Lemma 5) and then use this to prove $\sup_{\theta > \theta_0} \mathbb{E}(\exp(-tZ_{m,\theta})) < \infty$ for some $\theta_0 = \theta_0(t)$, when $t > 1$ is fixed (Lemma 9). This will allow us to apply Lemma 4 in the proofs of Theorems 1 and 2. Note that Lemma 9 was also needed in the proof of Lemma 1. In the next section, we will use these results to prove uniform integrability for the conditional sampling function $\phi_{\mathbf{a}}(\mathbf{X}_\theta)$ (Lemma 11). This will be used, together with Lemma 2, to prove Theorems 4 and 5.

It is well known that, if V_1, V_2, \dots are i.i.d. Beta(1, θ) random variables with density given by

$$f(v) = \theta(1-v)^{\theta-1} \mathbf{1}_{(0,1)}(v),$$

then a ‘‘size-biased’’ version of the Poisson–Dirichlet distribution with parameter θ is given by

$$\begin{aligned} W_{1,\theta} &= V_1, \\ W_{k,\theta} &= (1 - V_1)(1 - V_2) \cdots (1 - V_{k-1}) V_k \quad \text{for } k \geq 2. \end{aligned}$$

The $W_{k,\theta}$'s are just a relabeling of the $X_{k,\theta}$'s, and so $H_m \equiv H_m(\mathbf{X}_\theta) = \sum_k X_{k,\theta}^m = \sum_k W_{k,\theta}^m$. It is easy to see, and was established by Griffiths (1988), that one can write

$$H_m \stackrel{\mathcal{L}}{=} V^m + (1 - V)^m \tilde{H}_m,$$

where $V \sim \text{Beta}(1, \theta)$ and \tilde{H}_m is a random variable that is independent of V and has the same distribution as H_m . A version of this *recursion* for $Z_{m,\theta}$ is

$$\begin{aligned} (m-1)!Z_{m,\theta} &= \sqrt{\theta}(\theta^{m-1}H_m - (m-1)!) \\ &\stackrel{\mathcal{L}}{=} \sqrt{\theta}(\theta^{m-1}[V^m + (1-V)^m\tilde{H}_m] - (m-1)!) \\ (26) \quad &= \sqrt{\theta}(\theta^{m-1}V^m + (m-1)!(1-V)^m - (m-1)!) \\ &\quad + (1-V)^m\sqrt{\theta}(\theta^{m-1}\tilde{H}_m - (m-1)!) \\ &= a_m(V) + (1-V)^m(m-1)!\tilde{Z}_{m,\theta}, \end{aligned}$$

where

$$(27) \quad a_m(V) \equiv \sqrt{\theta}(\theta^{m-1}V^m - (m-1)![1 - (1-V)^m]),$$

$V \sim \text{Beta}(1, \theta)$ and $\tilde{Z}_{m,\theta}$ is a random variable that is independent of V and has the same distribution as $Z_{m,\theta}$.

We record here a simple moment formula that will be used a number of times in the paper. If $V \sim \text{Beta}(1, \theta)$, then, for any nonnegative integers m and n ,

$$\begin{aligned} (28) \quad \mathbb{E}[V^m(1-V)^n] &= \theta \frac{\Gamma(m+1)\Gamma(n+\theta)}{\Gamma(m+n+\theta+1)} \\ &= \frac{m!\theta}{(n+\theta)(n+\theta+1)\cdots(n+\theta+m)}. \end{aligned}$$

We will also need the following bounds for $a_m(V)$:

$$(29) \quad \frac{1}{\sqrt{\theta}}(\theta^m V^m - m!2^m \theta V) \leq a_m(V) \leq \frac{1}{\sqrt{\theta}}(\theta^m V^m + m!2^m \theta V).$$

The second inequality follows from the fact that $0 < V < 1$ with probability 1, and hence

$$(30) \quad 0 < 1 - (1 - V)^m \leq 2^m V.$$

The first inequality follows similarly. Also, (29) implies

$$(31) \quad |a_m(V)| \leq \frac{1}{\sqrt{\theta}}(\theta^m V^m + m!2^m \theta V)$$

and so, for $k \geq 1$,

$$\begin{aligned}
 \mathbb{E}|a_m(V)|^k &\leq \theta^{-k/2} \mathbb{E}(\theta^m V^m + m! 2^m \theta V)^k \\
 &\leq \theta^{-k/2} 2^k [\mathbb{E}(\theta V)^{mk} + (m! 2^m)^k \mathbb{E}(\theta V)^k] \\
 (32) \quad &\leq (2/\sqrt{\theta})^k [(mk)! + (m! 2^m)^k k!] \\
 &\sim O(\theta^{-k/2}),
 \end{aligned}$$

where the last inequality follows from

$$(33) \quad 0 \leq \mathbb{E}(\theta V)^j = \frac{j! \theta^j}{(\theta + 1) \cdots (\theta + j)} \leq j!, \quad j \geq 1.$$

It is worth noting that one of the reasons for writing the bounds in (29) and (31) the way we did is that θV converges in distribution to an $\text{Exp}(1)$ random variable as $\theta \rightarrow \infty$. Note also that we are suppressing the dependence on θ in the notation for $V \sim \text{Beta}(1, \theta)$.

LEMMA 5. *Let $Z_{m,\theta}$ be defined by (13). Then, for all integers $r \geq 1$ and $m \geq 2$,*

$$(34) \quad \sup_{\theta > 1} \mathbb{E}_0(|Z_{m,\theta}|^r) < \infty.$$

PROOF. Since $\mathbb{E}(|Z_{m,\theta}|^r) \leq \sqrt{\mathbb{E}_0(Z_{m,\theta}^{2r})}$, the lemma will follow if we prove

$$(35) \quad \sup_{\theta > 1} \mathbb{E}_0(Z_{m,\theta}^r) < \infty$$

for all integers $r \geq 1$. It is easy to see, using the Ewens sampling formula, that

$$\begin{aligned}
 \mathbb{E}_0(Z_{m,\theta}) &= \sqrt{\theta} \left[\frac{\theta^{m-1}}{(\theta + 1) \cdots (\theta + m - 1)} - 1 \right] \\
 (36) \quad &\sim O\left(\frac{1}{\sqrt{\theta}}\right)
 \end{aligned}$$

and hence (35) holds when $r = 1$. In the last step, we used the fact that one can express

$$(37) \quad \frac{\theta^{m-1}}{(\theta + 1) \cdots (\theta + m - 1)} = 1 - \phi_m(\theta),$$

where $|\phi_m(\theta)| < \binom{m+1}{2} \theta^{-1} + O(\theta^{-2})$.

We now proceed by induction on r (with m fixed), assuming that $\sup_{\theta > 1} \mathbb{E}_0(Z_{m,\theta}^k) < \infty$ holds for $1 \leq k \leq r-1$. Using (26), we can write

$$\begin{aligned} [(m-1)!Z_{m,\theta}]^r &\stackrel{\mathcal{L}}{=} (a_m(V) + (1-V)^m(m-1)!\tilde{Z}_{m,\theta})^r \\ &= \sum_{k=0}^r \binom{r}{k} a_m^k(V)(1-V)^{m(r-k)} [(m-1)!\tilde{Z}_{m,\theta}]^{r-k} \\ &= (1-V)^{mr} [(m-1)!\tilde{Z}_{m,\theta}]^r \\ &\quad + \sum_{k=1}^r \binom{r}{k} a_m^k(V)(1-V)^{m(r-k)} [(m-1)!\tilde{Z}_{m,\theta}]^{r-k}. \end{aligned}$$

Taking expectation on both sides, using the independence of V and $\tilde{Z}_{m,\theta}$ and setting

$$(38) \quad \mu_{r,\theta} \equiv ((m-1)!)^r \mathbb{E}_0(Z_{m,\theta}^r),$$

we get

$$\mu_{r,\theta} = \mathbb{E}((1-V)^{mr})\mu_{r,\theta} + \sum_{k=1}^r \binom{r}{k} \mathbb{E}(a_m^k(V)(1-V)^{m(r-k)})\mu_{r-k,\theta}.$$

This implies

$$\mu_{r,\theta} \left(1 - \frac{\theta}{mr + \theta}\right) = \sum_{k=1}^r \binom{r}{k} \mathbb{E}(a_m^k(V)(1-V)^{m(r-k)})\mu_{r-k,\theta},$$

and therefore we have the recursion

$$(39) \quad \mu_{r,\theta} = \frac{mr + \theta}{mr} \sum_{k=1}^r \binom{r}{k} \mathbb{E}(a_m^k(V)(1-V)^{m(r-k)})\mu_{r-k,\theta}.$$

So the lemma will follow if we can establish

$$(40) \quad \sup_{\theta > 1} \theta |\mathbb{E}(a_m^k(V)(1-V)^{m(r-k)})| < \infty$$

for $k = 1, \dots, r$, where $V \sim \text{Beta}(1, \theta)$. The cases $k = 1$ and $k \geq 2$ must be treated separately.

We begin with $k = 1$:

$$\begin{aligned} &\theta |\mathbb{E}(a_m(V)(1-V)^{m(r-1)})| \\ &= \sqrt{\theta} |\mathbb{E}\{\theta^m V^m (1-V)^{m(r-1)} - (m-1)!\theta [1 - (1-V)^m] (1-V)^{m(r-1)}\}| \\ &= \sqrt{\theta} \left| \frac{m!\theta^{m+1}}{(\theta + m(r-1)) \cdots (\theta + m(r-1) + m)} \right. \\ &\quad \left. - (m-1)!\theta \left(\frac{\theta}{\theta + m(r-1)} - \frac{\theta}{\theta + mr} \right) \right| \\ &= \sqrt{\theta} |m!(1 - \phi(\theta)) - m!(1 - \psi(\theta))|, \end{aligned}$$

where $|\phi(\theta)| \sim O(\theta^{-1})$ and $|\psi(\theta)| \sim O(\theta^{-1})$. Thus,

$$(41) \quad \theta |\mathbb{E}(a_m(V)(1-V)^{m(r-1)})| \leq \sqrt{\theta} m! (|\phi(\theta)| + |\psi(\theta)|) \sim O(1/\sqrt{\theta})$$

and (40) is verified for $k = 1$.

To handle the case $2 \leq k \leq r$, use (32) to see

$$\begin{aligned} & \sup_{\theta > 1} \theta |\mathbb{E}(a_m^k(V)(1-V)^{m(r-k)})| \\ & \leq \sup_{\theta > 1} \theta \mathbb{E}|a_m(V)|^k \leq \sup_{\theta > 1} 2^k \theta^{1-k/2} [(mk)! + (m!2^m)^k k!] < \infty. \end{aligned}$$

Thus, (40) is satisfied when $k \geq 2$, and this completes the proof. \square

The next three lemmas provide estimates that are used in the proof of the crucial uniform integrability result in Lemma 9. To obtain the desired uniform integrability, we will need to bound (uniformly in θ) an alternating series related to $\exp\{-tZ_{m,\theta}\}$. The following three lemmas (essentially calculus exercises that were not so obvious that we could omit the proofs) will aid us in the calculations involving the aforementioned series.

LEMMA 6. *Let $\sum_j A_j x^j$ and $\sum_j B_j x^j$ be two absolutely convergent series. Then, for any $1 \leq k \leq l$,*

$$\left| \sum_{i=l}^{\infty} \sum_{j=k}^i A_j B_{i-j} x^i \right| \leq \left| \sum_{j=k}^{\infty} A_j x^j \sum_{i=0}^{\infty} B_i x^i \right| + |x|^k C_{kl}(|x|),$$

where

$$(42) \quad C_{kl}(u) \equiv \sum_{j=k}^{l-1} |A_j| u^{j-k} \sum_{i=0}^{l-j-1} |B_i| u^i$$

is an increasing function of $u \in [0, \infty)$.

PROOF. Because each series is absolutely convergent, Fubini's theorem implies, for $k \leq l$,

$$\begin{aligned} (43) \quad \sum_{i=l}^{\infty} \sum_{j=k}^i A_j B_{i-j} x^i &= \sum_{j=k}^{l-1} \sum_{i=l}^{\infty} A_j x^j B_{i-j} x^{i-j} + \sum_{j=l}^{\infty} \sum_{i=j}^{\infty} A_j x^j B_{i-j} x^{i-j} \\ &= \sum_{j=k}^{l-1} A_j x^j \sum_{q=l-j}^{\infty} B_q x^q + \sum_{j=l}^{\infty} A_j x^j \sum_{q=0}^{\infty} B_q x^q \\ &= \sum_{j=k}^{\infty} A_j x^j \sum_{q=0}^{\infty} B_q x^q - \sum_{j=k}^{l-1} A_j x^j \sum_{q=0}^{l-j-1} B_q x^q. \end{aligned}$$

An application of the triangle inequality now yields the desired result. \square

LEMMA 7. For $k = 1, 2$ and positive integers $l > k$, define $h_{l,k} : \mathbb{R} \rightarrow \mathbb{R}$ by

$$h_{l,k}(z) \equiv \sum_{r=l-k}^{\infty} \frac{(-z)^r}{(r+k)r!}.$$

Then, for $l > k + 1$:

(i) $h'_{l,k}(z) = -h_{l,k+1}(z)$;

Furthermore, if x lies between z and 0 , then

(ii) $|h_{l,k}(x)| \leq e^{-z}/l + 2 \sum_{r=0}^{l-1} \frac{|z|^r}{r!}$; and

(iii) $|z||h_{l,k}(x)| \leq e^{-z} + 2 \sum_{r=0}^{l-1} \frac{|z|^r}{r!}$.

PROOF. The equality in (i) follows immediately from an interchange of summation and differentiation. The conditions that justify this interchange are easily verified in this case.

To prove (ii), begin by defining, for $k = 1, 2$, the continuous functions

$$(44) \quad h_k(z) \equiv \begin{cases} \sum_{r=0}^{\infty} \frac{(-z)^r}{(r+k)r!}, & \text{if } z \neq 0, \\ \frac{1}{k}, & \text{if } z = 0. \end{cases}$$

It is easy to see that, when $z \neq 0$,

$$h_1(z) = \frac{1 - e^{-z}}{z} \quad \text{and} \quad h_2(z) = -\frac{d}{dz} h_1(z) = \frac{1 - e^{-z} - ze^{-z}}{z^2}.$$

It is also easy to check that $h_1(z)$ and $h_2(z)$ are nonnegative, decreasing functions of $z \in \mathbb{R}$. We now proceed to prove (ii) by considering separately the cases $0 \leq x \leq z$ and $z \leq x \leq 0$.

In the first case, $h_k(z) \leq h_k(0)$ and so, for $k = 1, 2$ and $l > k$,

$$(45) \quad \begin{aligned} |h_{l,k}(x)| &= \left| h_k(x) - \sum_{r=0}^{l-k-1} \frac{(-x)^r}{(r+k)r!} \right| \\ &\leq h_k(x) + \sum_{r=0}^{l-2} \frac{|x|^r}{(r+1)!} \leq h_k(0) + \sum_{r=0}^{l-2} \frac{|z|^r}{(r+1)!} \leq 2 \sum_{r=0}^{l-1} \frac{|z|^r}{r!}. \end{aligned}$$

Now suppose $z \leq x \leq 0$. Then $0 \leq -x \leq -z$, and hence

$$(46) \quad |h_{l,k}(x)| = \sum_{r=l-k}^{\infty} \frac{(-x)^r}{(r+k)r!} \leq \sum_{r=l-k}^{\infty} \frac{(-z)^r}{(r+k)r!} < \frac{1}{l} \sum_{r=l-k}^{\infty} \frac{(-z)^r}{r!} \leq \frac{e^{-z}}{l}.$$

So (ii) follows from (45) and (46).

To prove (iii), we begin by assuming $0 \leq x \leq z$ as before. Arguing as in (45), we obtain

$$(47) \quad |z||h_{l,k}(x)| \leq |z|h_k(0) + \sum_{r=0}^{l-2} \frac{|z|^{r+1}}{(r+1)!} \leq 2 \sum_{r=0}^{l-1} \frac{|z|^r}{r!}.$$

In the case $z \leq x \leq 0$, we have

$$(48) \quad |z| |h_{l,k}(x)| = -z \sum_{r=l-k}^{\infty} \frac{(-x)^r}{(r+k)r!} \leq \sum_{r=l-k}^{\infty} \frac{(-z)^{r+1}}{(r+1)r!} < \sum_{r=0}^{\infty} \frac{(-z)^r}{r!} = e^{-z}.$$

Therefore, (iii) follows from (47) and (48). \square

LEMMA 8. For $m \geq 2$, let $a_m(V)$ be defined as in (27), with $V \sim \text{Beta}(1, \theta)$ and $\theta > 1$. Then, for each $j \geq 1$, there is a positive, increasing, deterministic function $\gamma_{m,j}(t)$ such that, for all $t > 0$,

$$(49) \quad \left| \sum_{k=j}^{\infty} \frac{(-ta_m(V))^k}{k!} \right| \leq t^j \gamma_{m,j}(t) |a_m(V)|^j.$$

PROOF. First, note that

$$\left| \sum_{k=j}^{\infty} \frac{(-ta_m(V))^k}{k!} \right| = |ta_m(V)|^j \left| \sum_{k=j}^{\infty} \frac{1}{k!} (-ta_m(V))^{k-j} \right| = t^j |a_m(V)|^j g_j(ta_m(V)),$$

where

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

To prove the desired inequality, it suffices to show that $g_j(ta_m(V))$ is a.s. bounded above as a function of $\theta > 1$. A simple calculus exercise shows that $g_j(x)$ is a decreasing function of $x \in (-\infty, \infty)$. Now define $f(u) = u^m - m!2^m u$. Note that f is bounded below on $[0, \infty)$ and let $u^* = ((m-1)!2^m)^{1/(m-1)}$ be the point in $[0, \infty)$ at which f achieves this minimum. Using (29) and the fact that $\theta V > 0$, we have

$$a_m(V) \geq \frac{f(\theta V)}{\sqrt{\theta}} \geq \frac{f(u^*)}{\sqrt{\theta}}.$$

Also, it is easy to see that $f(u^*) < 0$, and hence $f(u^*)/\sqrt{\theta} > f(u^*)$ for $\theta > 1$. Therefore, for all $t > 0$,

$$g_j(ta_m(V)) \leq g_j\left(\frac{tf(u^*)}{\sqrt{\theta}}\right) \leq g_j(tf(u^*)) \equiv \gamma_{m,j}(t) < \infty.$$

This complete the proof. \square

LEMMA 9. Let $Z_{m,\theta}$ be defined by (13) and suppose $m \geq 2$ and $t > 0$ are fixed. Then there exists $\theta_0 = \theta_0(t) < \infty$ such that

$$\sup_{\theta > \theta_0} \mathbb{E}_0(e^{-tZ_{m,\theta}}) < \infty.$$

PROOF. Since m is fixed throughout, to simplify notation, we will write

$$(50) \quad \bar{Z}_\theta \equiv (m-1)! \tilde{Z}_{m,\theta},$$

where \bar{Z}_θ is independent of V as in (26). As before, let $\mu_{r,\theta} = ((m-1)!)^r \times \mathbb{E}_0(Z'_{m,\theta}) = \mathbb{E}_0(\bar{Z}'_\theta)$. By (34), it is enough to show, for $t > 0$, that there are finite values $l = l(t)$ and $\theta_0 = \theta_0(l, t)$ such that

$$(51) \quad \sup_{\theta > \theta_0} \left| \sum_{r=l}^{\infty} \frac{(-t)^r}{r!} \mu_{r,\theta} \right| < \infty.$$

(This is not true if $-t$ is replaced with t , and it is this reliance on the alternating nature of the sum that makes the following argument rather delicate.)

First, use (39) to write

$$(52) \quad \left| \sum_{r=l}^{\infty} \frac{(-t)^r}{r!} \mu_{r,\theta} \right| = \left| \sum_{r=l}^{\infty} \frac{mr + \theta}{mr} \frac{(-t)^r}{r!} \sum_{k=1}^r \binom{r}{k} \mu_{r-k,\theta} \mathbb{E}(a_m^k(V) (1-V)^{m(r-k)}) \right|,$$

where $a_m(V)$ is defined in (27). Motivated by Lemma 6, we define the following random variables:

$$(53) \quad A_k \equiv \frac{(a_m(V))^k}{k!}, \quad B_q \equiv \frac{(\bar{Z}_\theta (1-V)^m)^q}{q!}.$$

Next, rewrite (52) as

$$(54) \quad \begin{aligned} \left| \sum_{r=l}^{\infty} \frac{(-t)^r}{r!} \mu_{r,\theta} \right| &= \left| \sum_{r=l}^{\infty} \sum_{k=1}^r \left(1 + \frac{\theta}{mr}\right) \mathbb{E}(A_k B_{r-k}) (-t)^r \right| \\ &= \left| \mathbb{E} \left(\sum_{r=l}^{\infty} \sum_{k=1}^r \left(1 + \frac{\theta}{mr}\right) A_k B_{r-k} (-t)^r \right) \right| \\ &\leq \left| \mathbb{E} \left(\sum_{r=l}^{\infty} \sum_{k=1}^r A_k B_{r-k} (-t)^r \right) \right| \\ &\quad + \left| \mathbb{E} \left(\sum_{r=l}^{\infty} \sum_{k=1}^r \frac{\theta}{mr} A_k B_{r-k} (-t)^r \right) \right| \\ &\equiv |\mathbb{E}(I_1)| + |\mathbb{E}(I_2)|. \end{aligned}$$

Note that, for fixed θ , the random variables V , $a_m(V)$ and \bar{Z}_θ are bounded. Thus, there are positive constants, c_θ and d_θ , such that $|A_k| \leq c_\theta^k/k!$ and $|B_q| \leq d_\theta^q/q!$ hold almost surely. Therefore, the interchange of expectation with summation in (54) is justified by the dominated convergence theorem.

Our strategy is now to bound each of the two terms in the last line of (54). Indeed, we will demonstrate that, for θ_0 and l chosen sufficiently large and all $\theta > \theta_0$,

$$(55) \quad |\mathbb{E}(I_1)| \leq \frac{1}{4} \sum_{r=0}^{\infty} \frac{(-t)^r}{r!} \mu_{r,\theta} + \alpha_1(t),$$

$$(56) \quad |\mathbb{E}(I_2)| \leq \frac{1}{4} \sum_{r=0}^{\infty} \frac{(-t)^r}{r!} \mu_{r,\theta} + \alpha_2(t),$$

where $\alpha_1(t)$ and $\alpha_2(t)$ are suitably chosen deterministic functions of t . Once the bounds in (55) and (56) have been established, we will have

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

Now collect all terms involving the common infinite series and put them on the left-hand side to get

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

Considering separately the cases where $\sum_{r=l}^{\infty} (-t)^r \mu_{r,\theta} / r!$ is positive and negative, it is easy to see that

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

Since the right-hand side is a sum of a finite number of terms, Lemma 5 can be used to conclude that

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

Thus, the proof of Lemma 9 reduces to establishing the inequalities in (55) and (56).

(I_1). Recall that A_k and B_k , defined by (53), are bounded random variables. Therefore, we can apply Lemma 6 to I_1 , with $k = 1$ and $x = -t$, to get

$$(57) \quad |I_1| \leq \left| \sum_{j=1}^{\infty} (-t)^j A_j \sum_{q=0}^{\infty} (-t)^q B_q \right| + tC_{1l}(t).$$

By Lemma 8, we have

$$\left| \sum_{j=1}^{\infty} (-t)^j A_j \right| = \left| \sum_{j=1}^{\infty} \frac{(-ta_m(V))^j}{j!} \right| \leq t\gamma_{m,1}(t) |a_m(V)|.$$

In addition,

$$(58) \quad \begin{aligned} \left| \sum_{q=0}^{\infty} B_q(-t)^q \right| &= \exp\{-t\bar{Z}_\theta(1-V)^m\} \\ &\leq \max\{1, \exp\{-t\bar{Z}_\theta\}\} \\ &\leq 1 + \exp\{-t\bar{Z}_\theta\}, \end{aligned}$$

and so

$$(59) \quad |I_1| \leq t\gamma_{m,1}(t)|a_m(V)|(1 + e^{-t\bar{Z}_\theta}) + tC_{1l}(t).$$

Therefore,

$$\begin{aligned} |\mathbb{E}(I_1)| &\leq \mathbb{E}|I_1| \\ &\leq t\gamma_{m,1}(t)\mathbb{E}|a_m(V)|\mathbb{E}(e^{-t\bar{Z}_\theta}) + t\gamma_{m,1}(t)\mathbb{E}|a_m(V)| + t\mathbb{E}(C_{1l}(t)). \end{aligned}$$

Since $C_{1l}(t)$ is a sum of a finite number of random variables with means uniformly bounded in θ , it follows that $\sup_{\theta>1} \mathbb{E}(C_{1l}(t)) < \infty$. Since $\mathbb{E}|a_m(V)| \sim O(\theta^{-1/2})$, there exists $\theta_0 = \theta_0(t) < \infty$ such that $t\gamma_{m,1}(t)\mathbb{E}|a_m(V)| < 1/4$ for all $\theta > \theta_0(t)$. Define $\alpha_1(t) \equiv \sup_{\theta>1} \mathbb{E}(C_{1l}(t)) + 1/4$. Then, for $\theta > \theta_0(t)$,

$$|\mathbb{E}(I_1)| \leq \frac{1}{4}\mathbb{E}(e^{-t\bar{Z}_\theta}) + \alpha_1(t) = \frac{1}{4} \sum_{r=0}^{\infty} \frac{(-t)^r}{r!} \mu_{r,\theta} + \alpha_1(t),$$

and this completes the proof of (55).

(I_2). Note that the factor $\theta/(mr)$ in I_2 does not appear in I_1 . Because of this, the proof of (56) requires a more delicate argument. We will need to consider separately the cases $k = 1$, $k = 2$ and $k \geq 3$. Begin by writing

$$\begin{aligned} I_2 &= \sum_{r=l}^{\infty} \sum_{k=1}^r \frac{\theta}{mr} A_k B_{r-k}(-t)^r \\ &= \sum_{r=l}^{\infty} \frac{\theta}{mr} A_1 B_{r-1}(-t)^r + \sum_{r=l}^{\infty} \frac{\theta}{mr} A_2 B_{r-2}(-t)^r + \sum_{r=l}^{\infty} \sum_{k=3}^r \frac{\theta}{mr} A_k B_{r-k}(-t)^r \\ &\equiv J_1 + J_2 + J_3. \end{aligned}$$

To prove (56), we now proceed to bound the means of the last three terms.

(J_1). Defining $\bar{X}_\theta \equiv (1-V)^m \bar{Z}_\theta$ and using the notation in Lemma 7, we can rewrite J_1 as

$$J_1 = -\frac{t\theta}{m} a_m(V) \sum_{r=l}^{\infty} \frac{(-t\bar{X}_\theta)^{r-1}}{r!} = -\frac{t\theta}{m} a_m(V) h_{l,1}(t\bar{X}_\theta).$$

Now recall from Lemma 7 that $h'_{l,1} = -h_{l,2}$ and apply the mean value theorem to $h_{l,1}$ to get

$$(60) \quad \begin{aligned} J_1 &= -\frac{t\theta}{m} a_m(V) [h_{l,1}(t\bar{Z}_\theta) + t(\bar{X}_\theta - \bar{Z}_\theta)h'_{l,1}(\xi)] \\ &= -\frac{t\theta}{m} a_m(V) [h_{l,1}(t\bar{Z}_\theta) + t(\bar{Z}_\theta - \bar{X}_\theta)h_{l,2}(\xi)], \end{aligned}$$

where ξ lies between $t\bar{X}_\theta$ and $t\bar{Z}_\theta$. Since $\bar{Z}_\theta - \bar{X}_\theta = [1 - (1 - V)^m]\bar{Z}_\theta$, it follows from (30) that

$$(61) \quad |\bar{Z}_\theta - \bar{X}_\theta| \leq 2^m V |\bar{Z}_\theta|.$$

Also, \bar{X}_θ and \bar{Z}_θ have the same sign, so we can apply Lemma 7 to get

$$(62) \quad \begin{aligned} |t\bar{Z}_\theta| |h_{l,2}(\xi)| &\leq e^{-t\bar{Z}_\theta} + 2 \sum_{r=0}^{l-1} \frac{|t\bar{Z}_\theta|^r}{r!}, \\ |h_{l,1}(t\bar{Z}_\theta)| &\leq e^{-t\bar{Z}_\theta} + 2 \sum_{r=0}^{l-1} \frac{|t\bar{Z}_\theta|^r}{r!}. \end{aligned}$$

Taking expectations in (60) and applying the bounds in (61) and (62), we get

$$\begin{aligned} |\mathbb{E}(J_1)| &\leq \frac{t\theta}{m} |\mathbb{E}(a_m(V))| |\mathbb{E}|h_{l,1}(t\bar{Z}_\theta)|| + \frac{t^2\theta}{m} \mathbb{E}|a_m(V)(\bar{Z}_\theta - \bar{X}_\theta)h_{l,2}(\xi)| \\ &\leq \left(\frac{t\theta}{m} |\mathbb{E}(a_m(V))| + \frac{2^m t^2\theta}{m} \mathbb{E}|V a_m(V)| \right) \left(\mathbb{E}(e^{-t\bar{Z}_\theta}) + 2 \sum_{r=0}^l \frac{\mathbb{E}|t\bar{Z}_\theta|^r}{r!} \right) \end{aligned}$$

for each $t > 0$. It follows from (27), (28) and (37) that $\theta |\mathbb{E}(a_m(V))| \sim m! O(1/\sqrt{\theta})$, and from (31) and (33) we have $\theta \mathbb{E}|V a_m(V)| \sim O(1/\sqrt{\theta})$. Therefore, we can choose θ large enough so that

$$(63) \quad |\mathbb{E}(J_1)| \leq \frac{1}{12} \mathbb{E}(e^{-t\bar{Z}_\theta}) + \beta_1(t) = \frac{1}{12} \sum_{r=0}^{\infty} \frac{(-t)^r \mu_{r,\theta}}{r!} + \beta_1(t),$$

where $\beta_1(t) = (1/6) \sup_{\theta > 1} \sum_{r=0}^l \mathbb{E}|t\bar{Z}_\theta|^r / r!$.

(J_2). Start with

$$\begin{aligned} |J_2| &= \left| \sum_{r=l}^{\infty} \frac{\theta}{mr} A_2 B_{r-2} (-t)^r \right| \\ &= \theta a_m^2(V) \frac{t^2}{2m} \left| \sum_{r=l}^{\infty} \frac{(-t\bar{X}_\theta)^{r-2}}{r(r-2)!} \right| = \theta a_m^2(V) \frac{t^2}{2m} |h_{l,2}(t\bar{X}_\theta)|. \end{aligned}$$

Then, from part (ii) of Lemma 7,

$$|J_2| \leq \frac{\theta t^2}{2m} a_m^2(V) \left(\frac{e^{-t\bar{Z}_\theta}}{l} + 2 \sum_{r=0}^l \frac{|t\bar{Z}_\theta|^r}{r!} \right).$$

Define

$$\lambda(t) \equiv \sup_{\theta > 1} \frac{\theta t^2}{2m} \mathbb{E}(a_m^2(V)).$$

It follows from (32) that $\lambda(t) < \infty$, so choosing $l > 12\lambda(t)$ leads to

$$(64) \quad \mathbb{E}|J_2| \leq \frac{1}{12} \mathbb{E}(e^{-t\bar{Z}_\theta}) + 12\lambda(t)\beta_1(t) = \frac{1}{12} \sum_{r=0}^{\infty} \frac{(-t)^r}{r!} \mu_{r,\theta} + \beta_2(t),$$

where $\beta_2(t) \equiv 12\lambda(t)\beta_1(t)$.

(J_3). We now rewrite J_3 and again apply Lemma 6 to get

$$\begin{aligned} |J_3| &= \left| \int_0^t \frac{1}{s} \sum_{r=l}^{\infty} \sum_{k=3}^r \frac{\theta}{m} A_k B_{r-k} (-s)^r ds \right| \\ &\leq \int_0^t \left(\frac{1}{s} \left| \sum_{k=3}^{\infty} \frac{\theta}{m} A_k (-s)^k \sum_{q=0}^{\infty} B_q (-s)^q \right| + s^3 C_{3l}(s) \right) ds \\ &\leq \int_0^t \frac{1}{s} \left| \sum_{k=3}^{\infty} \frac{\theta}{m} A_k (-s)^k \right| \left(\sum_{q=0}^{\infty} B_q (-s)^q \right) ds + t^2 C_{3l}(t). \end{aligned}$$

It follows from Lemma 8 that

$$\left| \sum_{k=3}^{\infty} A_k (-s)^k \right| = \left| \sum_{k=3}^{\infty} \frac{(-s a_m(V))^k}{k!} \right| \leq s^3 \gamma_{m,3}(t) |a_m(V)|^3$$

for $0 \leq s \leq t$, and by the same argument used in (58), we have $0 < \sum_{q=0}^{\infty} B_q (-s)^q \leq 1 + \exp\{-t\bar{Z}_\theta\}$. Therefore,

$$|J_3| \leq (1 + e^{-t\bar{Z}_\theta}) \frac{\gamma_{m,3}(t)}{m} \theta |a_m(V)|^3 \int_0^t s^2 ds + t^2 C_{3l}(t),$$

and this implies, by independence of V and \bar{Z}_θ ,

$$\mathbb{E}|J_3| \leq t^3 \frac{\gamma_{m,3}(t)}{m} \theta \mathbb{E}|a_m(V)|^3 \mathbb{E}(e^{-t\bar{Z}_\theta}) + t^3 \frac{\gamma_{m,3}(t)}{m} \theta \mathbb{E}|a_m(V)|^3 + t^2 \mathbb{E}(C_{3l}(t)).$$

By (32), we can choose θ so large that

$$t^3 \frac{\gamma_{m,3}(t)}{m} \theta \mathbb{E}|a_m(V)|^3 < \frac{1}{12}.$$

Therefore, setting $\beta_3(t) \equiv 1/12 + t^2 \sup_{\theta > 1} \mathbb{E}(C_{3l}(t))$, we arrive at

$$(65) \quad \mathbb{E}|J_3| \leq \frac{1}{12} \mathbb{E}(e^{-t\bar{Z}_\theta}) + \beta_3(t) = \frac{1}{12} \sum_{r=0}^{\infty} \frac{(-t)^r \mu_{r,\theta}}{r!} + \beta_3(t),$$

when θ is sufficiently large. So (56) now follows from (63)–(65), with $\alpha_2(t) = \beta_1(t) + \beta_2(t) + \beta_3(t)$. \square

5. Rescaled likelihood ratios and sampling distributions. In this section we extend our uniform integrability and limit results to rescaled versions of the likelihood ratios and sampling distributions. These results will be used in the proofs of Theorems 3, 4 and 5.

LEMMA 10. *Suppose $\sigma = c\theta^{3/2-\alpha}$, where $\alpha > 0$ and $c > 0$ are constants.*

(i) *If $\mathbf{X}_\theta = (X_{1,\theta}, X_{2,\theta}, \dots) \sim \mu_0$, then*

$$\theta^\alpha \left(\frac{d\mu_\sigma}{d\mu_0}(\mathbf{X}_\theta) - 1 \right) = -cZ_{2,\theta} + L(Z_{2,\theta}),$$

where $\lim_{\theta \rightarrow \infty} \mathbb{E}_0 |L(Z_{2,\theta})|^r = 0$ for all $r > 0$.

(ii) *If $r > 0$, then*

$$\sup_{\theta > 1} \mathbb{E}_0 \left| \theta^\alpha \left(\frac{d\mu_\sigma}{d\mu_0}(\mathbf{X}_\theta) - 1 \right) \right|^r < \infty.$$

(iii) *If $\mathbf{Y}_\theta^{(-\alpha)} = (Y_{1,\theta}^{(-\alpha)}, Y_{2,\theta}^{(-\alpha)}, \dots) \sim \mu_\sigma$, then*

$$\theta^\alpha \left(\frac{d\mu_\sigma}{d\mu_0}(\mathbf{Y}_\theta^{(-\alpha)}) - 1 \right) = -cZ_{2,\theta}^{(-\alpha)} + L_\alpha(Z_{2,\theta}^{(-\alpha)}),$$

where $\lim_{\theta \rightarrow \infty} \mathbb{E}_\sigma |L_\alpha(Z_{2,\theta}^{(-\alpha)})|^r = 0$ for all $r > 0$.

REMARK. Lemma 10, when combined with Lemma 1, will provide a quick proof of Theorem 3. Part (ii) yields a uniform integrability condition that will be needed in the proofs of Theorems 4 and 5.

PROOF OF LEMMA 10. Beginning with (i), recall that, for $\sigma = c\theta^{3/2-\alpha}$,

$$\frac{d\mu_\sigma}{d\mu_0}(\mathbf{X}_\theta) = \frac{e^{-c\theta^{-\alpha}Z_{2,\theta}}}{\mathbb{E}_0(e^{-c\theta^{-\alpha}Z_{2,\theta}})}.$$

Using the Taylor decomposition, $e^{-x} = 1 - x + R(x)$, with $0 < R(x) < \max\{x^2/2, e^{-x}x^2/2\} \leq (x^2/2)(1 + e^{-x})$, we can write

$$e^{-c\theta^{-\alpha}Z_{2,\theta}} = 1 - c\theta^{-\alpha}Z_{2,\theta} + R(c\theta^{-\alpha}Z_{2,\theta}).$$

For notational convenience let us denote $R(c\theta^{-\alpha}Z_{2,\theta})$ by R_θ .

As a consequence of the classical result $\mathbb{E}_0(H(\mathbf{X}_\theta)) = 1/(1+\theta)$, we have

$$\begin{aligned}\mathbb{E}_0(e^{-c\theta^{-\alpha}Z_{2,\theta}}) &= 1 - \mathbb{E}_0(c\theta^{-\alpha}Z_{2,\theta}) + \mathbb{E}_0(R_\theta) \\ &= 1 + c\theta^{1/2-\alpha}\frac{1}{1+\theta} + \mathbb{E}_0(R_\theta).\end{aligned}$$

Hence, the quantity of interest can be written as

$$\begin{aligned}&\theta^\alpha\left(\frac{e^{-c\theta^{-\alpha}Z_{2,\theta}}}{\mathbb{E}_0(e^{-c\theta^{-\alpha}Z_{2,\theta}})} - 1\right) \\ &= \frac{\theta^\alpha}{\mathbb{E}_0(e^{-c\theta^{-\alpha}Z_{2,\theta}})}(1 - c\theta^{-\alpha}Z_{2,\theta} + R_\theta - \mathbb{E}_0(e^{-c\theta^{-\alpha}Z_{2,\theta}})) \\ &= \frac{\theta^\alpha}{\mathbb{E}_0(e^{-c\theta^{-\alpha}Z_{2,\theta}})}\left(-c\theta^{-\alpha}Z_{2,\theta} + R_\theta - c\frac{\theta^{1/2-\alpha}}{1+\theta} - \mathbb{E}_0(R_\theta)\right) \\ &= \frac{1}{\mathbb{E}_0(e^{-c\theta^{-\alpha}Z_{2,\theta}})}\left(-cZ_{2,\theta} - c\frac{\sqrt{\theta}}{1+\theta} + \theta^\alpha R_\theta - \mathbb{E}_0(\theta^\alpha R_\theta)\right) \\ &= -cZ_{2,\theta} + L(Z_{2,\theta}),\end{aligned}$$

where we define

$$\begin{aligned}L(Z_{2,\theta}) &\equiv \frac{1}{\mathbb{E}_0(e^{-c\theta^{-\alpha}Z_{2,\theta}})}\left(-c\frac{\sqrt{\theta}}{1+\theta} + \theta^\alpha R_\theta - \theta^\alpha\mathbb{E}_0(R_\theta)\right) \\ &\quad - cZ_{2,\theta}\left(\frac{1}{\mathbb{E}_0(e^{-c\theta^{-\alpha}Z_{2,\theta}})} - 1\right).\end{aligned}$$

Since Lemma 9 implies, for any constant $\delta > 1$,

$$(66) \quad \sup_{\theta > \theta_0(\delta)} \mathbb{E}_0(e^{-\delta\theta^{-\alpha}Z_{2,\theta}}) \leq 1 + \sup_{\theta > \theta_0(\delta)} \mathbb{E}_0(e^{-\delta Z_{2,\theta}}) < \infty,$$

we see that $\mathbb{E}_0(e^{-c\theta^{-\alpha}Z_{2,\theta}}) \rightarrow 1$. Furthermore, using the Cauchy–Schwarz inequality, it is clear that

$$\begin{aligned}0 &< \theta^{r\alpha}\mathbb{E}_0|R_\theta^r| < \theta^{r\alpha}\mathbb{E}_0((c\theta^{-\alpha}Z_{2,\theta})^{2r}(1 + \exp\{-c\theta^{-\alpha}Z_{2,\theta}\})^r) \\ &< c^{2r}\theta^{-r\alpha}\sqrt{\mathbb{E}_0(Z_{2,\theta}^{4r})\mathbb{E}_0(1 + \exp\{-c\theta^{-\alpha}Z_{2,\theta}\})^{2r}}\end{aligned}$$

holds for all $r > 0$. Thus, we can use Lemmas 5 and 9 to get $\theta^{r\alpha}\mathbb{E}_0|R_\theta^r| \rightarrow 0$ as $\theta \rightarrow \infty$. A simple argument, together with Lemmas 5 and 9, now produces

$$\lim_{\theta \rightarrow \infty} \mathbb{E}_0|L(Z_{2,\theta})|^r = 0,$$

completing the proof of (i).

Part (ii) follows immediately from part (i) and Lemma 5.

To prove part (iii), one argues as in the proof of (i), making the obvious changes. Replacing $L(Z_{2,\theta})$ will be

$$L_\alpha(Z_{2,\theta}^{(-\alpha)}) \equiv \frac{1}{\mathbb{E}_0(e^{-c\theta^{-\alpha}Z_{2,\theta}})} \left(-c \frac{\sqrt{\theta}}{1+\theta} + \theta^\alpha R_\theta^{(-\alpha)} - \theta^\alpha \mathbb{E}_0(R_\theta) \right) - c Z_{2,\theta}^{(-\alpha)} \left(\frac{1}{\mathbb{E}(e^{-c\theta^{-\alpha}Z_{2,\theta}})} - 1 \right),$$

where $R_\theta^{(-\alpha)} \equiv R(c\theta^{-\alpha}Z_{2,\theta}^{(-\alpha)})$. The last part of (iii) requires uniform integrability of $Z_{2,\theta}^{(-\alpha)}$ and $e^{-Z_{2,\theta}^{(-\alpha)}}$. These follow from Lemmas 5 and 9 together with part (ii) above. \square

Sampling distributions and homozygosities. Recall that Theorems 4 and 5 involve limits for sampling distributions. Our next goal is to establish moment bounds for certain scalings of the conditional sampling functions, $\phi_{\mathbf{a}}(\mathbf{X}_\theta)$. These will lead to the uniform integrability we need to prove the aforementioned theorems. For example, we wish to show that if the partition \mathbf{a} satisfies $a_1 \neq n$ (i.e., not all alleles in the sample are distinct), then

$$(67) \quad \sup_{\theta > 1} \mathbb{E}_0 \left| \sqrt{\theta} \left(\frac{\phi_{\mathbf{a}}(\mathbf{X}_\theta) - \text{ESF}(\theta, \mathbf{a})}{\text{ESF}(\theta, \mathbf{a})} \right) \right|^r < \infty.$$

To aid us in our derivation of (67), we will exploit a relationship between the population moments, $H_m(\mathbf{X}_\theta)$, and the conditional sampling function, $\phi_{\mathbf{a}}(\mathbf{X}_\theta)$.

This is given by the following general identity, which holds for any partition \mathbf{a} of a sample of size n :

$$(68) \quad \prod_{m=2}^n \left(\frac{\theta^{m-1}}{(m-1)!} H_m(\mathbf{X}_\theta) \right)^{a_m} = \frac{\theta^n}{\theta_{(n)}} \frac{\phi_{\mathbf{a}}(\mathbf{X}_\theta)}{\mathbb{E}_0(\phi_{\mathbf{a}}(\mathbf{X}_\theta))} + \frac{\theta^{n-k}}{\prod_{m=2}^n ((m-1)!)^{a_m}} \sum_{\mathbf{b} \in \mathcal{C}_{\mathbf{a}}} l_{\mathbf{b}} \phi_{\mathbf{b}}(\mathbf{X}_\theta),$$

where $\mathcal{C}_{\mathbf{a}}$ is defined to be the collection of all partitions formed by coalescing two or more of the classes associated with \mathbf{a} , and $l_{\mathbf{b}}$ is a nonnegative combinatorial factor whose exact form is not important here. A derivation appears in Joyce, Krone and Kurtz (2001). To understand why this identity will be useful, note that the term on the left involves population homozygosities and the last term is small when θ is large, as will be seen below. So the large- θ asymptotics of $\phi_{\mathbf{a}}(\mathbf{X}_\theta)$ will be determined by those of the various population homozygosities. This is helpful because the population homozygosities are especially simple examples of sampling distributions.

Since $\mathbb{E}_0(\phi_{\mathbf{a}}(\mathbf{X}_\theta)) = \text{ESF}(\theta, \mathbf{a})$, the above identity can be used to obtain

$$\begin{aligned} \left| \frac{\phi_{\mathbf{a}}(\mathbf{X}_\theta) - \text{ESF}(\theta, \mathbf{a})}{\text{ESF}(\theta, \mathbf{a})} \right| &= \left| \frac{\phi_{\mathbf{a}}(\mathbf{X}_\theta)}{\mathbb{E}_0(\phi_{\mathbf{a}}(\mathbf{X}_\theta))} - 1 \right| \\ &\leq \left| \frac{\theta_{(n)}}{\theta^n} \prod_{m=2}^n \left(\frac{\theta^{m-1} H_m(\mathbf{X}_\theta)}{(m-1)!} \right)^{a_m} - 1 \right| \\ &\quad + \frac{\theta_{(n)}}{\theta^n} \frac{\theta^{n-k}}{\prod_{m=2}^n ((m-1)!)^{a_m}} \sum_{\mathbf{b} \in \mathcal{C}_{\mathbf{a}}} l_{\mathbf{b}} \phi_{\mathbf{b}}(\mathbf{X}_\theta). \end{aligned}$$

Next, notice that

$$\sqrt{\theta} \left(\frac{\theta_{(n)}}{\theta^n} - 1 \right) \sim O\left(\frac{1}{\sqrt{\theta}} \right),$$

and set

$$B \equiv \sup_{\theta > 1} \sqrt{\theta} \left| \frac{\theta_{(n)}}{\theta^n} - 1 \right|.$$

Thus, the quantity whose r th moment appears in (67) has absolute value

$$\begin{aligned} &\sqrt{\theta} \left| \frac{\phi_{\mathbf{a}}(\mathbf{X}_\theta)}{\mathbb{E}_0(\phi_{\mathbf{a}}(\mathbf{X}_\theta))} - 1 \right| \\ &\leq \sqrt{\theta} \left| \prod_{m=2}^n \left(\frac{\theta^{m-1} H_m(\mathbf{X}_\theta)}{(m-1)!} \right)^{a_m} - 1 \right| + \sqrt{\theta} \left| \frac{\theta_{(n)}}{\theta^n} - 1 \right| \prod_{m=2}^n \left(\frac{\theta^{m-1} H_m(\mathbf{X}_\theta)}{(m-1)!} \right)^{a_m} \\ (69) \quad &+ \frac{\theta^{n-k}}{\prod_{m=2}^n ((m-1)!)^{a_m}} \sum_{\mathbf{b} \in \mathcal{C}_{\mathbf{a}}} l_{\mathbf{b}} \phi_{\mathbf{b}}(\mathbf{X}_\theta) \\ &\leq \sqrt{\theta} \left| \prod_{m=2}^n \left(\frac{\theta^{m-1} H_m(\mathbf{X}_\theta)}{(m-1)!} \right)^{a_m} - 1 \right| + B \prod_{m=2}^n \left| \frac{Z_{m,\theta}}{\sqrt{\theta}} + 1 \right|^{a_m} \\ &\quad + \frac{\theta^{n-k}}{\prod_{m=2}^n ((m-1)!)^{a_m}} \sum_{\mathbf{b} \in \mathcal{C}_{\mathbf{a}}} l_{\mathbf{b}} \phi_{\mathbf{b}}(\mathbf{X}_\theta). \end{aligned}$$

The first term in the last line of (69) is the most important, and we isolate it in the first part of the next lemma. The uniform integrability condition given in (67) will then be given, followed by the last part of the lemma treating the special case when $a_1 = n$.

LEMMA 11. *Suppose $\mathbf{X}_\theta \sim \mu_0$ and let \mathbf{a} be a partition corresponding to a sample of size n drawn from a population with ordered frequencies \mathbf{X}_θ .*

(i) If $a_1 \neq n$, then

$$(70) \quad \sup_{\theta > 1} \theta^{r/2} \mathbb{E}_0 \left| \prod_{m=2}^n \left(\frac{\theta^{m-1} H_m(\mathbf{X}_\theta)}{(m-1)!} \right)^{a_m} - 1 \right|^r < \infty \quad \text{for each } r > 0.$$

(ii) If $a_1 \neq n$, then

$$\sup_{\theta > 1} \theta^{r/2} \mathbb{E}_0 \left| \frac{\phi_{\mathbf{a}}(\mathbf{X}_\theta) - \text{ESF}(\theta, \mathbf{a})}{\text{ESF}(\theta, \mathbf{a})} \right|^r < \infty \quad \text{for each } r > 0.$$

(iii) 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

$$\sup_{\theta > 1} \theta^{r/2} \mathbb{E}_0 \left| \frac{\phi_{\mathbf{a}_1}(\mathbf{X}_\theta) - \text{ESF}(\theta, \mathbf{a}_1)}{\text{ESF}(\theta, \mathbf{a}_2)} \right|^r < \infty \quad \text{for each } r > 0.$$

PROOF. We prove (i) by induction on the sample size n . If $n = 2$, then the assumption, $a_1 \neq n$, implies a_2 must be 1. In this case, (70) reduces to the uniform integrability result in Lemma 5:

$$\sup_{\theta > 1} \theta^{r/2} \mathbb{E}_0 |\theta H_2(\mathbf{X}_\theta) - 1|^r = \sup_{\theta > 1} \mathbb{E}_0 |Z_{2,\theta}|^r < \infty.$$

Now assume that (70) holds for all partitions corresponding to samples of size less than or equal to n , and let $\mathbf{a} = (a_1, a_2, \dots, a_{n+1})$ be a partition corresponding to a sample of size $n+1$. For partitions with $a_{n+1} = 1$ and with $a_n = 1$, $a_1 = 1$, the quantity of interest,

$$\theta^{r/2} \left| \prod_{m=2}^{n+1} \left(\frac{\theta^{m-1} H_m(\mathbf{X}_\theta)}{(m-1)!} \right)^{a_m} - 1 \right|^r,$$

reduces to $|Z_{n+1,\theta}|^r$ and $|Z_{n,\theta}|^r$, respectively, and hence can be handled by Lemma 5. (Notice that terms associated with a_1 never appear in the product.) Since the hypothesis in (i) says $a_1 \neq n+1$, we are left to treat the situation in which $a_j \geq 1$ for some $1 < j \leq n-1$. For such a j , define $\mathbf{a}_j = (a_{j1}, \dots, a_{jn})$ to be a partition corresponding to a sample of size $n+1-j$ formed by removing one allele with j representatives. That is, $a_{jj} = a_j - 1$, and $a_{ji} = a_i$ for all $i \neq j$. Then

$$(71) \quad \begin{aligned} & \theta^{r/2} \left| \prod_{m=2}^{n+1} \left(\frac{\theta^{m-1} H_m(\mathbf{X}_\theta)}{(m-1)!} \right)^{a_m} - 1 \right|^r \\ &= \theta^{r/2} \left| \frac{\theta^{j-1} H_j(\mathbf{X}_\theta)}{(j-1)!} \left(\prod_{m=2}^{n+1-j} \left(\frac{\theta^{m-1} H_m(\mathbf{X}_\theta)}{(m-1)!} \right)^{a_{mj}} - 1 \right) \right. \\ & \quad \left. + \frac{\theta^{j-1} H_j(\mathbf{X}_\theta)}{(j-1)!} - 1 \right|^r \\ &\leq 2^r \theta^{r/2} \left| \frac{\theta^{j-1} H_j(\mathbf{X}_\theta)}{(j-1)!} \left(\prod_{m=2}^{n+1-j} \left(\frac{\theta^{m-1} H_m(\mathbf{X}_\theta)}{(m-1)!} \right)^{a_{mj}} - 1 \right) \right|^r + 2^r |Z_{j,\theta}|^r. \end{aligned}$$

Since Lemma 5 can be used to handle the mean of the last term, it remains to treat the first term in the last line of (71). An appeal to the Cauchy–Schwarz inequality allows us to write

$$(72) \quad \mathbb{E}_0 \left| \frac{\theta^{j-1} H_j(\mathbf{X}_\theta)}{(j-1)!} \left(\prod_{m=2}^{n+1-j} \left(\frac{\theta^{m-1} H_m(\mathbf{X}_\theta)}{(m-1)!} \right)^{a_{mj}} - 1 \right) \right|^r \\ \leq \sqrt{\mathbb{E}_0 \left| \frac{\theta^{j-1} H_j(\mathbf{X}_\theta)}{(j-1)!} \right|^{2r} \mathbb{E}_0 \left| \prod_{m=2}^{n+1-j} \left(\frac{\theta^{m-1} H_m(\mathbf{X}_\theta)}{(m-1)!} \right)^{a_{mj}} - 1 \right|^{2r}}.$$

Clearly, the induction hypothesis implies

$$\sup_{\theta > 0} \theta^{r/2} \mathbb{E}_0 \left| \prod_{m=2}^{n+1-j} \left(\frac{\theta^{m-1} H_m(\mathbf{X}_\theta)}{(m-1)!} \right)^{a_{mj}} - 1 \right|^{2r} < \infty,$$

and the first term under the square root is handled with Lemma 5. Combining all this with (71) and (72) completes the proof of (i).

To prove (ii), we proceed by induction on the number, k , of different alleles represented in the sample. If $k = 1$, then all the sampled individuals have the same allele; that is, $a_n = 1$. Thus, $\phi_{\mathbf{a}}(\mathbf{X}_\theta) = \sum X_{i,\theta}^n$, and (ii) follows as a special case of Lemma 5, after a simple argument to account for the fact that the scaling terms are only asymptotically the same. Next, assume (ii) holds for all partitions of n with fewer than k distinct alleles. In other words, if \mathbf{b} is such a partition and we define

$$D_{\mathbf{b}}(\mathbf{X}_\theta) \equiv \sqrt{\theta} \left(\frac{\phi_{\mathbf{b}}(\mathbf{X}_\theta)}{\mathbb{E}_0(\phi_{\mathbf{b}}(\mathbf{X}_\theta))} - 1 \right),$$

then the induction hypothesis implies $\sup_{\theta > 1} \mathbb{E}_0 |D_{\mathbf{b}}(\mathbf{X}_\theta)|^r < \infty$ for all $r > 0$. However, from the definition of $D_{\mathbf{b}}(\mathbf{X}_\theta)$, we can write

$$(73) \quad (\phi_{\mathbf{b}}(\mathbf{X}_\theta))^r = (\mathbb{E}_0(\phi_{\mathbf{b}}(\mathbf{X}_\theta)))^r \left(\frac{D_{\mathbf{b}}(\mathbf{X}_\theta)}{\sqrt{\theta}} + 1 \right)^r.$$

To apply this to a partition \mathbf{a} with k distinct alleles, recall the inequality in (69). Clearly, to complete the induction argument, it suffices to show that the r th moment of each of the terms in the last line of (69) is uniformly bounded in $\theta > 1$. For the first term in (69), just apply part (i) above. For the last term in (69), using the fact that the number of distinct alleles represented in any partition $\mathbf{b} \in \mathcal{C}_{\mathbf{a}}$ is less than or equal to $k - 1$, we can infer from the Ewens sampling formula that $\mathbb{E}_0(\phi_{\mathbf{b}}(\mathbf{X}_\theta)) \sim O(\theta^{k-1-n})$, and so, taking expectations on both sides of (73), we get

$$(74) \quad \sup_{\theta > 1} \theta^{r(n-k)} \mathbb{E}_0(\phi_{\mathbf{b}}(\mathbf{X}_\theta))^r < \infty.$$

From this, it follows immediately that the r th moment of the last term in (69) is bounded in $\theta > 1$. Finally, the middle term in (69) is easy to handle; just use Lemma 5 and Hölder's inequality to see that

$$(75) \quad \sup_{\theta > 1} \mathbb{E}_0 \left(\prod_{m=2}^n \left| \frac{Z_{m,\theta}}{\sqrt{\theta}} + 1 \right|^{ra_m} \right) < \infty.$$

To prove (iii), we note that $\phi_{\mathbf{a}_1}(\mathbf{X}_\theta) = 1 - \sum_{\mathbf{a} \neq \mathbf{a}_1} \phi_{\mathbf{a}}(\mathbf{X}_\theta)$. Therefore,

$$(76) \quad \begin{aligned} & \frac{\phi_{\mathbf{a}_1}(\mathbf{X}_\theta) - \mathbb{E}_0(\phi_{\mathbf{a}_1}(\mathbf{X}_\theta))}{\mathbb{E}_0(\phi_{\mathbf{a}_2}(\mathbf{X}_\theta))} \\ &= \frac{\mathbb{E}_0(\sum_{\mathbf{a} \neq \mathbf{a}_1} \phi_{\mathbf{a}}(\mathbf{X}_\theta)) - \sum_{\mathbf{a} \neq \mathbf{a}_1} \phi_{\mathbf{a}}(\mathbf{X}_\theta)}{\mathbb{E}_0(\phi_{\mathbf{a}_2}(\mathbf{X}_\theta))} \\ &= \frac{\mathbb{E}_0(\phi_{\mathbf{a}_2}(\mathbf{X}_\theta)) - \phi_{\mathbf{a}_2}(\mathbf{X}_\theta)}{\mathbb{E}_0(\phi_{\mathbf{a}_2}(\mathbf{X}_\theta))} + \sum_{\mathbf{a} \neq \mathbf{a}_1, \mathbf{a}_2} \frac{\mathbb{E}_0(\phi_{\mathbf{a}}(\mathbf{X}_\theta)) - \phi_{\mathbf{a}}(\mathbf{X}_\theta)}{\mathbb{E}_0(\phi_{\mathbf{a}}(\mathbf{X}_\theta))} \frac{\mathbb{E}_0(\phi_{\mathbf{a}}(\mathbf{X}_\theta))}{\mathbb{E}_0(\phi_{\mathbf{a}_2}(\mathbf{X}_\theta))}. \end{aligned}$$

It follows from direct calculation using the Ewens sampling formula (8) that

$$\lim_{\theta \rightarrow \infty} \frac{\mathbb{E}_0(\phi_{\mathbf{a}}(\mathbf{X}_\theta))}{\mathbb{E}_0(\phi_{\mathbf{a}_2}(\mathbf{X}_\theta))} = 0$$

if $\mathbf{a} \neq \mathbf{a}_1, \mathbf{a}_2$. Thus,

$$(77) \quad \sup_{\theta > 1} \frac{\mathbb{E}_0(\phi_{\mathbf{a}}(\mathbf{X}_\theta))}{\mathbb{E}_0(\phi_{\mathbf{a}_2}(\mathbf{X}_\theta))} < \infty.$$

Therefore,

$$\sup_{\theta > 1} \theta^{r/2} \mathbb{E}_0 \left| \frac{\phi_{\mathbf{a}_1}(\mathbf{X}_\theta) - \text{ESF}(\theta, \mathbf{a}_1)}{\text{ESF}(\theta, \mathbf{a}_2)} \right|^r < \infty$$

follows from part (ii) and (76) and (77). \square

6. Proofs of the main results.

PROOF OF THEOREM 1. It follows from (18) that $cZ_{2,\theta} \Rightarrow cZ_2$, where $Z_2 \sim N(0, 2)$. So, in the case $\sigma = c\theta^{3/2}$, we can use (15) and Lemmas 4 and 9, with Y_n replaced by $cZ_{2,\theta}$ and n with θ , to get

$$(78) \quad \frac{d\mu_\sigma}{d\mu_0}(\mathbf{X}_\theta) = \frac{\exp\{-cZ_{2,\theta}\}}{\mathbb{E}_0(\exp\{-cZ_{2,\theta}\})} \Rightarrow \frac{\exp\{cZ_2\}}{\mathbb{E}(\exp\{-cZ_2\})} = \exp\{cZ_2 - c^2\}.$$

Here we have used $\mathbb{E}(e^{-cZ_2}) = e^{-c^2}$ and the fact that Z_2 is equal in distribution to $-Z_2$. (Note that the t in Lemma 9 can be chosen so $ct > 1$.) The cases $\sigma = c\theta^{3/2+\gamma}$ with $\gamma < 0$ and $\gamma > 0$ of Theorem 1 also follow by applying Lemma 4. \square

PROOF OF THEOREM 2. Suppose $\mathbf{Y}_\theta^{(\gamma)} \sim \mu_\sigma$, where $\sigma = c\theta^{3/2+\gamma}$. First, consider the case $\gamma \leq 0$. By (16), we have

$$(79) \quad \frac{d\mu_\sigma}{d\mu_0}(\mathbf{Y}_\theta^{(\gamma)}) = \frac{\exp\{-c\theta^\gamma Z_{2,\theta}^{(\gamma)}\}}{\mathbb{E}_0(\exp\{-c\theta^\gamma Z_{2,\theta}\})},$$

and so the desired result follows from the continuous mapping theorem for weak limits and the uniform integrability obtained from (66) and Lemma 9.

Now assume that $\sigma = c\theta^{3/2+\gamma}$, with $\gamma > 0$, and let $g \equiv d\mu_\sigma/d\mu_0$. Then

$$\mathbb{E}_\sigma\left(\frac{1}{\sqrt{g(\mathbf{Y}_\theta^{(\gamma)})}}\right) = \mathbb{E}_0\left(\frac{g(\mathbf{X}_\theta)}{\sqrt{g(\mathbf{X}_\theta)}}\right) = \mathbb{E}_0(\sqrt{g(\mathbf{X}_\theta)}).$$

It follows from Lemma 3 and (15) that $\lim_{\theta \rightarrow \infty} \mathbb{E}_0(\sqrt{g(\mathbf{X}_\theta)}) = 0$. Thus, we conclude that $\lim_{\theta \rightarrow \infty} \mathbb{E}_\sigma(1/\sqrt{g(\mathbf{Y}_\theta^{(\gamma)})}) = 0$, implying $g(\mathbf{Y}_\theta^{(\gamma)}) \Rightarrow \infty$. \square

PROOF OF THEOREM 3. The result follows immediately from Lemmas 1 and 10. \square

PROOF OF THEOREM 4. Suppose $\sigma = c\theta^{3/2-\alpha}$ and $\mathbf{a} \neq (n, 0, \dots, 0)$ is a partition for a sample of size n . Treating first the case $\alpha > 0$, we have

$$\begin{aligned} & \left| \theta^{\alpha+1/2} \frac{\mathbb{P}_\sigma(\mathbf{A}_n = \mathbf{a})}{\mathbb{P}_0(\mathbf{A}_n = \mathbf{a})} - 1 \right| \\ &= \theta^{\alpha+1/2} \left| \frac{\mathbb{P}_\sigma(\mathbf{A}_n = \mathbf{a}) - \mathbb{P}_0(\mathbf{A}_n = \mathbf{a})}{\mathbb{P}_0(\mathbf{A}_n = \mathbf{a})} \right| \\ &= \theta^{\alpha+1/2} \left| \int \frac{d\mu_\sigma}{d\mu_0}(\mathbf{x}) \left(\frac{\phi_{\mathbf{a}}(\mathbf{X}_\theta) - \text{ESF}(\theta, \mathbf{a})}{\text{ESF}(\theta, \mathbf{a})} \right) \mu_0(d\mathbf{x}) \right| \\ &= \left| \mathbb{E}_0 \left(\sqrt{\theta} \left(\frac{\phi_{\mathbf{a}}(\mathbf{X}_\theta) - \text{ESF}(\theta, \mathbf{a})}{\text{ESF}(\theta, \mathbf{a})} \right) \theta^\alpha \frac{d\mu_\sigma}{d\mu_0}(\mathbf{X}_\theta) \right) \right| \\ &= \left| \mathbb{E}_0 \left(\sqrt{\theta} \left(\frac{\phi_{\mathbf{a}}(\mathbf{X}_\theta) - \text{ESF}(\theta, \mathbf{a})}{\text{ESF}(\theta, \mathbf{a})} \right) \theta^\alpha \frac{d\mu_\sigma}{d\mu_0}(\mathbf{X}_\theta) - \theta^\alpha + \theta^\alpha \right) \right| \\ &= \left| \mathbb{E}_0 \left(\sqrt{\theta} \left(\frac{\phi_{\mathbf{a}}(\mathbf{X}_\theta) - \text{ESF}(\theta, \mathbf{a})}{\text{ESF}(\theta, \mathbf{a})} \right) \theta^\alpha \left(\frac{d\mu_\sigma}{d\mu_0}(\mathbf{X}_\theta) - 1 \right) \right) \right|. \end{aligned}$$

Note that Lemma 11 and part (ii) of Lemma 10, together with the Cauchy–Schwarz inequality, give uniform integrability of the last expression. Therefore, we can bring limits inside the expectation and use part (i) of Lemma 2, together with Theorem 3, to get

$$\lim_{\theta \rightarrow \infty} \theta^{\alpha+1/2} \left| \frac{\mathbb{P}_\sigma(\mathbf{A}_n = \mathbf{a})}{\mathbb{P}_0(\mathbf{A}_n = \mathbf{a})} - 1 \right| = c \left| \sum_{i=2}^n a_i \text{Cov}(Z_i, Z_2) \right| = 2c \sum_{i=2}^n a_i \binom{i}{2}.$$

In the case $\alpha = 0$, applying Lemma 2 and noting that uniform integrability follows by Lemmas 9 and 11, we have

$$\begin{aligned} & \lim_{\theta \rightarrow \infty} \theta^{1/2} \left| \frac{\mathbb{P}_\sigma(\mathbf{A}_n = \mathbf{a})}{\mathbb{P}_0(\mathbf{A}_n = \mathbf{a})} - 1 \right| \\ &= \lim_{\theta \rightarrow \infty} \left| \mathbb{E}_0 \left(\sqrt{\theta} \left(\frac{\phi_{\mathbf{a}}(\mathbf{X}_\theta) - \text{ESF}(\theta, \mathbf{a})}{\text{ESF}(\theta, \mathbf{a})} \right) \frac{e^{-cZ_{2,\theta}}}{\mathbb{E}_0(e^{-cZ_{2,\theta}})} \right) \right| \\ &= \sum_{i=2}^n a_i \frac{|\mathbb{E}(Z_i e^{-cZ_2})|}{\mathbb{E}(e^{-cZ_2})}. \end{aligned}$$

It follows from standard properties of bivariate normal distributions that

$$\begin{aligned} (80) \quad \mathbb{E}(e^{-cZ_2} Z_i) &= \mathbb{E}(e^{-cZ_2} \mathbb{E}(Z_i | Z_2)) \\ &= \frac{\text{Cov}(Z_2, Z_i)}{\text{Var } Z_2} \mathbb{E}(e^{-cZ_2} Z_2) = \frac{\text{Cov}(Z_2, Z_i)}{2} \mathbb{E}(e^{-cZ_2} Z_2) \end{aligned}$$

and

$$\begin{aligned} (81) \quad \mathbb{E}(e^{-cZ_2} Z_2) &= \int_{-\infty}^{\infty} \frac{1}{2\sqrt{\pi}} z e^{-z^2/4} e^{-cz} dz \\ &= e^{c^2} \int_{-\infty}^{\infty} \frac{1}{2\sqrt{\pi}} z e^{-(z+2c)^2/4} dz = -2ce^{c^2}. \end{aligned}$$

Therefore,

$$\begin{aligned} \lim_{\theta \rightarrow \infty} \theta^{1/2} \left| \frac{\mathbb{P}_\sigma(\mathbf{A}_n = \mathbf{a})}{\mathbb{P}_0(\mathbf{A}_n = \mathbf{a})} - 1 \right| &= e^{-c^2} \sum_{i=2}^n a_i |\mathbb{E}(Z_i e^{-cZ_2})| \\ &= e^{-c^2} \sum_{i=2}^n a_i \frac{\text{Cov}(Z_2, Z_i)}{2} (2ce^{c^2}) \\ &= c \sum_{i=2}^n a_i \text{Cov}(Z_2, Z_i) = 2c \sum_{i=2}^n a_i \binom{i}{2}. \quad \square \end{aligned}$$

PROOF OF THEOREM 5. Suppose $\mathbf{a}_1 = (n, 0, \dots, 0)$ and $\mathbf{a}_2 = (n - 2, 1, 0, \dots, 0)$ are the allelic partitions with n singletons and $n - 2$ singletons, respectively. Then, for $\alpha > 0$,

$$\begin{aligned} & \theta^{\alpha+3/2} \left| \frac{\mathbb{P}_\sigma(\mathbf{A}_n = \mathbf{a}_1)}{\mathbb{P}_0(\mathbf{A}_n = \mathbf{a}_1)} - 1 \right| \\ &= \frac{\theta^{\alpha+1}}{\mathbb{E}_0(e^{-c\theta^{-\alpha} Z_{2,\theta}})} \left| \mathbb{E}_0 \left(\sqrt{\theta} \left(\frac{\phi_{\mathbf{a}_1}(\mathbf{X}_\theta) - \text{ESF}(\theta, \mathbf{a}_1)}{\text{ESF}(\theta, \mathbf{a}_1)} \right) \frac{d\mu_\sigma}{d\mu_0}(\mathbf{X}_\theta) \right) \right| \\ &= \frac{\theta^{\alpha+1}}{\mathbb{E}_0(e^{-c\theta^{-\alpha} Z_{2,\theta}})} \frac{\text{ESF}(\theta, \mathbf{a}_2)}{\text{ESF}(\theta, \mathbf{a}_1)} \end{aligned}$$

$$\begin{aligned} & \times \left| \mathbb{E}_0 \left(\sqrt{\theta} \left(\frac{\phi_{\mathbf{a}_1}(\mathbf{X}_\theta) - \text{ESF}(\theta, \mathbf{a}_1)}{\text{ESF}(\theta, \mathbf{a}_2)} \right) \left(\frac{d\mu_\sigma}{d\mu_0}(\mathbf{X}_\theta) - 1 \right) \right) \right| \\ &= \frac{\theta}{\mathbb{E}_0(e^{-c\theta^{-\alpha} Z_{2,\theta}})} \frac{\text{ESF}(\theta, \mathbf{a}_2)}{\text{ESF}(\theta, \mathbf{a}_1)} \\ & \times \left| \mathbb{E}_0 \left(\sqrt{\theta} \left(\frac{\phi_{\mathbf{a}_1}(\mathbf{X}_\theta) - \text{ESF}(\theta, \mathbf{a}_1)}{\text{ESF}(\theta, \mathbf{a}_2)} \right) \theta^\alpha \left(\frac{d\mu_\sigma}{d\mu_0}(\mathbf{X}_\theta) - 1 \right) \right) \right|. \end{aligned}$$

Note that $\theta \text{ESF}(\theta, \mathbf{a}_2)/\text{ESF}(\theta, \mathbf{a}_1) \rightarrow 1$, and $\mathbb{E}_0(e^{-c\theta^{-\alpha} Z_{2,\theta}}) \rightarrow 1$ as $\theta \rightarrow \infty$. Once again, the uniform integrability established in Lemmas 10 and 11 allows us to bring the limit inside the last expected value. It then follows from part (ii) of Lemma 2 and Theorem 3 that

$$\lim_{\theta \rightarrow \infty} \theta^{\alpha+3/2} \left| \frac{\mathbb{P}_\sigma(\mathbf{A}_n = \mathbf{a}_1)}{\mathbb{P}_0(\mathbf{A}_n = \mathbf{a}_1)} - 1 \right| = c |\mathbb{E}_0(-Z_2 Z_2)| = c \text{Var}(Z_2) = 2c. \quad \square$$

Acknowledgments. We thank John Gillespie for bringing this problem to our attention. We also thank Bob Griffiths for suggesting the recursion satisfied by $Z_{m,\theta}$.

REFERENCES

- ETHIER, S. N. and KURTZ, T. G. (1994). Convergence to Fleming–Viot processes in the weak atomic topology. *Stochastic Process. Appl.* **54** 1–27.
- GILLESPIE, J. H. (1999). The role of population size in molecular evolution. *Theoretical Population Biology* **55** 145–156.
- GRIFFITHS, R. C. (1979). On the distribution of allele frequencies in a diffusion model. *Theoretical Population Biology* **15** 140–158.
- GRIFFITHS, R. C. (1988). Distribution of F in the infinitely-many-alleles model. Statistics Research Report 183, Dept. Mathematics, Monash Univ.
- JOYCE, P., KRONE, S. M. and KURTZ, T. G. (2002). Gaussian limits associated with the Poisson–Dirichlet distribution and the Ewens sampling formula. *Ann. Appl. Probab.* **12** 101–124.
- KINGMAN, J. F. C. (1977). The population structure associated with the Ewens sampling formula. *Theoretical Population Biology* **11** 274–283.

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

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

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