Coalescent Processes with Skewed Offspring Distributions and non-Equilibrium Demography

Sebastian Matuszewski[a,c,1], Marcel E. Hildebrandt[b,1], Jeffrey D. Jensen[a,c,d,1,*]

[a]*School of Life Sciences, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland*
[b]*School of Basic Sciences, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland*
[c]*Swiss Institute of Bioinformatics (SIB), Lausanne, Switzerland*
[d]*Center for Evolution & Medicine, School of Life Sciences, Arizona State University, Tempe, AZ, USA*

Abstract Non-equilibrium demography impacts coalescent genealogies leaving detectable, well-studied signatures of variation. However, similar genomic footprints are also expected under models of large reproductive skew, posing a serious problem when trying to make inference. Furthermore, current approaches consider only one of the two processes at a time, neglecting any genomic signal that could arise from their simultaneous effects, preventing the possibility of jointly inferring parameters relating to both offspring distribution and population history. Here, we develop an extended Moran model with exponential population growth, and demonstrate that the underlying ancestral process converges to a time-inhomogeneous psi-coalescent. However, by applying a non-linear change of time scale – analogous to the Kingman coalescent – we find that the ancestral process can be rescaled to its time-homogeneous analogue, allowing the process to be simulated quickly and efficiently. Furthermore, we derive analytical expressions for the expected site-frequency spectrum under the time-inhomogeneous psi-coalescent and develop an approximate-likelihood framework for the joint estimation of the coalescent and growth parameters. By means of extensive simulation, we demonstrate that both can be estimated accurately given linkage equilibrium, while linkage disequilibrium systematically biases growth rate estimates. In addition, not accounting for demography can lead to serious biases in the inferred coalescent model, with broad implications for genomic studies ranging from ecology to conservation biology. Finally, we use our method to analyze sequence data from Japanese sardine populations and find evidence of high variation in individual reproductive success, but few signs of a recent demographic expansion.

***Key words*:** Coalescent Theory, Multiple Mergers, Population Growth, Maximum Likelihood, Site-frequency Spectrum

*To whom correspondence should be addressed. E-mail: jeffrey.d.jensen@asu.edu
[1]contributed equally

May 12, 2017

Introduction

The origins of the coalescent in the early 1970s mark a milestone for evolutionary theory (Kingman 2000). More than 45 years after Kingman formally proved the existence of the "n-coalescent" (Kingman 1982a,b,c), the so-called Kingman has gradually become the key theoretical tool to study the complex interplay of mutation, genetic drift, gene flow and selection. Closely linked to its underlying forward-in-time population model – e.g., the Wright-Fisher (WF; Fisher 1930; Wright 1931) and the Moran model (Moran 1958, 1962) –, the Kingman coalescent has been used to derive expected levels of neutral variation, including the number of segregating sites s, the average number of pairwise differences π, and the distribution of the allele frequencies in a population $\boldsymbol{\eta}$ (i.e., the site frequency spectrum; SFS). In fact, these predictions not only apply to the WF and Moran model, but extend to a large class of Cannings exchangeable population models (Cannings 1974) that all converge to the Kingman coalescent in the ancestral limit (Möhle and Sagitov 2001). Furthermore, the Kingman coalescent forms the basis for many population genetic statistics – such as Tajima's D (Tajima 1989), Fay and Wu's H (Fay and Wu 2000) or more generally any SFS-based test statistic (Achaz 2009; Ferretti *et al.* 2010) – and subsequent inferences (Irwin *et al.* 2016) to detect deviations from the assumption of a neutrally evolving, constant-sized, panmictic population (Wakeley 2009).

While the Kingman coalescent has been shown to be robust to violations of its assumptions (Möhle 1998, 1999), such as constant population size, random mating, and non-overlapping generations, and has been extended to accommodate selection, migration, and population structure (Neuhauser and Krone 1997; Nordborg 1997; Wilkinson-Herbots 1998), it breaks down in the presence of skewed offspring distributions (Eldon and Wakeley 2006), strong positive selection (Neher and Hallatschek 2013), recurrent selective sweeps (Durrett and Schweinsberg 2004, 2005), and large sample sizes (Wakeley and Takahashi 2003; Bhaskar *et al.* 2014). In particular, all of these effects can cause more than two lineages to coalesce at a time, resulting in so-called multiple mergers. Hence, the underlying coalescent topology (i.e., the gene genealogy) is no longer represented by a bifurcating tree as in the "standard" Kingman case, but can take more complex tree shapes that can also feature several simultaneous mergers. Taking these points into account, a more general class of models, so-called multiple-merger coalescent (MMC) models, have been developed (e.g., Bolthausen and Sznitman 1998; Pitman 1999; Sagitov 1999; Schweinsberg 2000; Möhle and Sagitov 2001; reviewed in Tellier and Lemaire 2014), aiming to generalize the Kingman coalescent model (Wakeley 2013). As for the latter, these MMC models can often be derived from Moran models, generalized to allow multiple offspring per individual (Eldon and Wakeley 2006; Huillet and Möhle 2013; Irwin *et al.* 2016).

Starting from such an extended Moran model, Eldon and Wakeley (2006) proved that the underlying ancestral process converges to a psi-coalescent (sometimes also called Dirac coalescent; Eldon *et al.* 2015), and that population genetic parameters inferred from genetic data from Pacific oysters (*Crassostrea gigas*) under this model vastly differ from those inferred assuming the Kingman coalescent. Their study – being the first to link MMC models to actual biological questions, molecular data and population genetic inferences – highlighted that high variation in individual reproductive success drastically affect both genealogical history and subsequent analyses; this has been observed in many marine organisms such

Atlantic cod (*Gadus morhua*) and Japanese sardines (*Sardinops melanostictus*), but should also occur more generally in any species with type III survivorship curves that undergo so-called sweepstake-reproductive events (Hedgecock 1994; Hedgecock and Pudovkin 2011), Fundamentally, the problem is that an excess of low-frequency alleles (i.e., singletons), a ubiquitous characteristic of many marine species (Niwa *et al.* 2016), could be explained by either models of recent population growth or skewed offspring distributions when analyzed under the Kingman coalescent assuming neutrality which can result in serious mis-inference (e.g., a vast overestimation of population growth).

In developing a SFS-based maximum likelihood framework, Eldon *et al.* (2015) demonstrated that multiple merger coalescents and population growth can be distinguished from their genomic footprints in the higher-frequency classes of the SFS with high statistical power (see also Spence *et al.* 2016). However, there is currently neither a modelling framework that considers the genomic signal arising from the joint action of both reproductive skew and population growth nor is there any *a priori* reason to believe that the two could not act simultaneously.

Here, we develop an extension of the standard Moran model that accounts for both reproductive skewness and exponential population growth, and prove that its underlying ancestral process converges to a time-inhomogeneous psi-coalescent. By (non-linearly) rescaling branch lengths this process can – analogous to the Kingman coalescent (Griffiths and Tavaré 1998) – be transformed into its time-homogeneous analogue allowing efficient large-scale simulations. Furthermore, we derive analytical formulae for the expected site-frequency spectrum under the time-inhomogeneous psi-coalescent and develop an approximate-likelihood framework for the joint estimation of the coalescent and growth parameters. We then perform extensive validation of our inference framework on simulated data and show that both the coalescent parameter and the growth rate can be estimated accurately provided that sites are in linkage equilibrium, while complete linkage systematically biases growth rate estimates. In addition, we demonstrate that when demography is not accounted for, the inferred coalescent model can be seriously biased with broad implications for genomic studies ranging from ecology to conservation biology (e.g., due to its effects on effective population size or diversity estimates). Finally, using our joint estimation method we re-analyze mtDNA from Japanese sardine (*Sardinops melanostictus*) populations and find evidence for considerable reproductive skew, but only limited support for a recent demographic expansion.

Model and Methods

Here we will first present an extended, discrete-time Moran model (Moran 1958, 1962; Eldon and Wakeley 2006) with exponential population growth which will serve as the forward-in-time population genetic model underlying the ancestral limit process. We will then give a brief overview of coalescent models with special focus on the psi-coalescent (Eldon and Wakeley 2006), before revisiting SFS-based maximum likelihood methods to infer coalescent parameters and population growth rates.

Table 1 – Summary of notation and definitions.

U_N	Number of offspring of a reproductive event in an extended Moran model with population size N
$\boldsymbol{\nu}$	Vector of family sizes
Λ	Probability measure on $[0,1]$
$\lambda_{i,x}$	Coalescent rate for x out of i active lineages
$G_{i,x}$	Probability of an $x-$merger among i active lineages
c_N	Coalescence probability
$\left(\mathcal{A}_{n,k}^{\psi,\rho}\right)_{n\in\mathbb{N}}$	Ancestral process of the extended Moran model sweepstake parameter ψ ($\psi = 0$ implying Kingman's coalescent) and exponential population growth at rate ρ for a sample of size k
$\left(\Pi_{t,k}^{\psi,\rho}\right)_{t\geq 0}$	$\psi-$coalescent ($\psi = 0$ implying Kingman's coalescent) with exponential growth at rate ρ and sample of size k
$\mathcal{G}(\cdot)$	Time-change function
t_{k-i+1}	Time when the ith coalescent event occurs in a sample of size k
τ_i	Time span during which the coalescent tree consists of i branches
$T_{\text{MRCA}}^{(k)}$	Time until the MRCA for a sample of size k
T_i	Sum of the length of all branches with i descendants
T_{tot}	Total branch length of the coalescent tree
$\boldsymbol{\eta}^{(k)} = (\eta_1^{(k)}, \ldots, \eta_{k-1}^{(k)})$	Site frequency spectrum (SFS) for a sample of size k
$\boldsymbol{\varphi}^{(k)} = (\varphi_1^{(k)}, \ldots, \varphi_{k-1}^{(k)})$	Expected normalized SFS for a sample of size k
$\mathfrak{c}_k = (\mathfrak{c}_{2,2}, \ldots, \mathfrak{c}_{k,k})$	Expected time to the first coalescence for a sample of size $i \in \{2, \ldots, k\}$

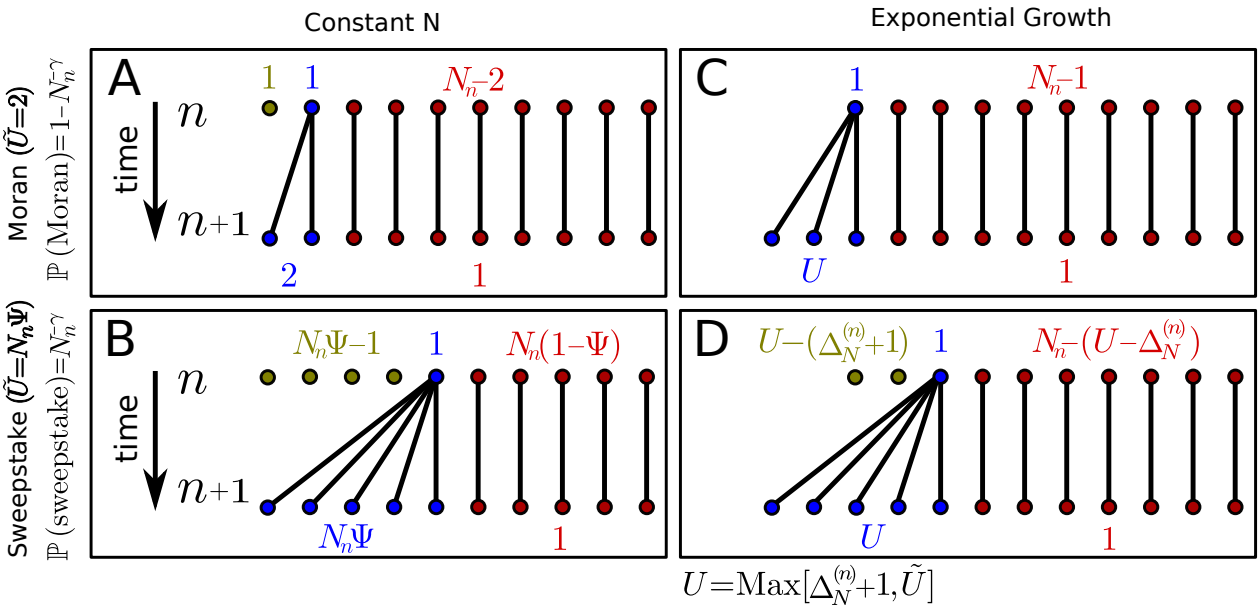

Figure 1 – Illustration of the extend Moran model with exponential growth. Shown are the four different scenarios of population transition within a single discrete time step. **A** The population size remains constant and a single individual produces exactly two offspring ('Moran-type' reproductive event). **B** The population size remains constant and a single individual produces ψN_n offspring ('sweepstake' reproductive event). **C** The population size increases by $\Delta_N^{(n)}$ individuals and a single individual produces exactly $\text{Max}[\Delta_N^{(n)}+1, 2]$ offspring. **D** The population size increases by $\Delta_N^{(n)}$ individuals and a single individual produces exactly $\text{Max}[\Delta_N^{(n)}+1, \psi N_n]$ offspring.

We consider the idealized, discrete-time model with variable population size shown generally in Figure 1. Furthermore, let $N_n \in \mathbb{N}$ be the deterministic and time-dependent population size $n \in \mathbb{N}$ time steps in the past, where by definition $N = N_0$ denotes the present population size. In particular, defining $\boldsymbol{\nu}(n)$ as the exchangeable vector of family sizes – with components $\nu_i(n)$ indicating the number of descendants of the ith individual – the (variable) population size can be expressed as

$$N_{n-1} = \sum_{i=1}^{N_n} \nu_i(n) \quad \text{with} \quad (\nu_1(n), \nu_2(n), \ldots, \nu_N(n)) \in \mathbb{N}^{N_n}. \tag{1}$$

Furthermore, we assume that the reproductive mechanism follows that of an extended Moran model (Eldon and Wakeley 2006; Huillet and Möhle 2013). In particular, as in the original Moran model, at any given point in time $n \in \mathbb{N}$ only a single individual reproduces and leaves $U_N(n)-1$ offspring. Formally, the number of offspring can be written as a sequence of random variables $(U_N(n))_{n\in\mathbb{N}}$ (where each $U_N(n)$ is supported on $\{0, 1, \ldots, N_{n-1}\}$), such that $\boldsymbol{\nu}(n)$ – up to reordering – is given by

$$\nu_i(n) := \begin{cases} 0 & \text{if } i < U_N(n) \\ U_N(n) & \text{if } i = U_N(n) \\ 1 & \text{otherwise.} \end{cases} \tag{2}$$

However, since population size varies over time, the sequence $(U_N(n))_{n\in\mathbb{N}}$ is generally not

identically distributed. On a technical note though, we require that the (U_n) are independently distributed which ensures that the corresponding backwards process satisfies the Markov property.

An illustration of our model and the four different scenarios for forming the next generation (i.e., within a single discrete time step) is shown in Figure 1. Generally we differentiate between two possible reproductive events: a classic 'Moran-type' reproductive event (Fig. 1**A**,**C**) and a 'sweepstake' reproductive event (Fig. 1**B**,**D**) occurring with probabilities $1 - N_n^{-\gamma}$ and $N_n^{-\gamma}$, respectively. If the population size remains constant between consecutive generations (Fig. 1**A**,**B**), we re-obtain the extended Moran model introduced by Eldon and Wakeley (2006) in which a single randomly chosen individual either leaves exactly two offspring (including itself) and replaces one randomly chosen individual (Moran-type) or replaces a fixed proportion $\psi \in (0, 1]$ of the population (of size N_n). Note that throughout, without loss of generality, we assume that $N_n\psi$ is integer-valued. In both reproductive scenarios the remaining individuals persist. However, if the population size increases between consecutive generations (Fig. 1**C**,**D**), the reproductive mechanism needs to be adjusted accordingly. Let

$$\Delta_N^{(n)} := N_{n-1} - N_n \tag{3}$$

denote the increment in population size between two consecutive time points. Then the number of offspring at time n is given by

$$U_N(n) := \mathrm{Max}[\Delta_N^{(n)} + 1, \tilde{U}_N(n)] \tag{4}$$

where $\tilde{U}_N(n)$ denotes number of offspring for the constant-size population. Thus, independent of the type of reproductive event – i.e., Moran-type or sweepstake – and in the spirit of the original Moran model, additional individuals are always assigned to be offspring of the single reproducing individual of the previous generation.

Following Eldon and Wakeley (2006), the distribution of the number of offspring $\mathbb{P}(U_N(n) = u)$ can be written as

$$\mathbb{P}(U_N(n) = u) = \begin{cases} N_n^{-\gamma} & \text{if } u = \mathrm{Max}[\Delta_N^{(n)} + 1, N_n\psi] \\ 1 - N_n^{-\gamma} & \text{if } u = \mathrm{Max}[\Delta_N^{(n)} + 1, 2] \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

for some $\gamma > 0$ that – for a given fixed population size – determines the probability of a sweepstake reproductive event. We will here only consider the case where $1 < \gamma < 2$ such that sweepstake events happen frequently enough that the ancestral process will be characterized by multiple mergers, but not so frequently that the population is devoid of genetic variation (Eldon and Wakeley 2006). Note that while the numbers of offspring and replaced individuals are no longer (necessarily) equal when the population size increases, the general reproductive mechanism remains unaltered.

Throughout the paper, we will assume that the population is growing exponentially over time at rate ρ, and in particular that the population size, n steps in the past, is given by

$$N_n := \left\lfloor N \left(1 - \rho \frac{\psi^2}{N^\gamma}\right)^n \right\rfloor. \tag{6}$$

Multiple merger coalescents: The Psi-coalescent

Λ-coalescents are partition-valued exchangeable stochastic processes defined by a finite measure Λ on the $\{0,1\}$ interval (Donnelly and Kurtz 1999; Pitman 1999; Sagitov 1999). In particular, the rate at which x out of i active lineages merge is given by

$$\lambda_{i,x} = \binom{i}{x} \int_0^1 y^{x-2}(1-y)^{i-x}\Lambda(\mathrm{d}y). \tag{7}$$

Special instances of the Λ-coalescent are Kingman's coalescent (Kingman 1982a,b) with

$$\Lambda(\mathrm{d}y) = \delta_0(\mathrm{d}y) \tag{8}$$

and the psi-coalescent (Eldon and Wakeley 2006) with

$$\Lambda(\mathrm{d}y) = \delta_\psi(\mathrm{d}y), \tag{9}$$

where the measure Λ is entirely concentrated at 0 and ψ, respectively.

In the following, we will briefly revisit the derivation of the ancestral process and in particular Equation 9 by Eldon and Wakeley (2006) as they form the basis for our own results. Under a (constant-size) extended Moran model (corresponding to Fig. 1**A**,**B**), the probability of an $x-$merger $G_{i,x}$ – i.e., when x out of i active lineages merge to a single lineage – is given by

$$\begin{aligned} G_{i,x} &= \sum_{u=2}^{N} P_U(u) \frac{\binom{u}{x}\binom{N-u}{i-x}}{\binom{N}{i}} \\ &= \binom{i}{x} \frac{(1-N^{-\gamma})(2)_x(N-2)_{i-x}}{(N)_i} \\ &+ \frac{N^{-\gamma}(\psi N)_x(N-\psi N)_{i-x}}{(N)_i}, \end{aligned} \tag{10}$$

where $(\cdot)_j$ is the descending factorial. Thus, the coalescence probability c_N – which serves as the 'natural' time scale for the corresponding ancestral process – becomes

$$c_N := G_{2,2} = \frac{2(1-N^{-\gamma}) + \psi N^{1-\gamma}(\psi N - 1)}{N(N-1)}. \tag{11}$$

After rescaling $G_{i,x}$ by c_N, the coalescence probabilities of the ancestral process become

$$\begin{aligned} \lambda_{i,x}^{(N)} &:= \frac{G_{i,x}}{c_N} \\ &= \binom{i}{x} \mathbb{P}^N(U_N = 2) A_N(2,i,x) \\ &+ \binom{i}{x} \mathbb{P}^N(U_N = \psi N) A_N(\psi N, i, x), \end{aligned} \tag{12}$$

where $A_N(u,i,x) := (u-2)_{x-2}(N-u)_{i-x}/(N-2)_{i-2}$ and $\mathbb{P}^N$ is the rescaled distribution of U_N given by

$$\mathbb{P}(U_N = u) := \frac{\mathbb{P}(U_N = u)u(u-1)/(N(N-1))}{c_N}. \tag{13}$$

Considering the two different reproductive events and taking the limit $N \to \infty$ then yields

$$\lim_{N\to\infty} \mathbb{P}^N(U_N = 2) = \begin{cases} 0 & \text{if } 0 < \gamma < 2 \\ \frac{2}{2+\psi^2} & \text{if } \gamma = 2 \\ 1 & \text{otherwise} \end{cases} \tag{14}$$

and

$$\lim_{N\to\infty} \mathbb{P}^N(U_N = \psi N) = \begin{cases} 1 & \text{if } 0 < \gamma < 2 \\ \frac{\psi^2}{2+\psi^2} & \text{if } \gamma = 2 \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

Thus, in the infinite population size limit all coalescence events are due to sweepstake reproductive events if $0 < \gamma < 2$, whereas sweepstake events do not happen frequently enough if $\gamma > 2$ such that all (2-)mergers are due to Moran-type reproductive events. Moreover, in the latter case the ancestral process of the Moran model can be accurately described by the Kingman coalescent (when scaled appropriately). Note that for the special case $\gamma = 2$ both reproductive events happen on the same time scale.

Using $A_N(2,i,2) = 1$, $A_N(2,i,x) = 0$ for $x > 0$ and $\lim_{N\to\infty} A_N(\psi N, i, x) = \psi^{x-2}(1-\psi)^{i-x}$ the coalescence probabilities become

$$\lim_{N\to\infty} \lambda_{i,x}^{(N)} = \begin{cases} \binom{i}{x}\psi^{x-2}(1-\psi)^{i-x} & \text{if } 0 < \gamma < 2 \\ \binom{i}{x}\left(\frac{2}{2+\psi^2}\mathbb{1}_{x=2}\frac{\psi^x}{2+\psi^2}(1-\psi)^{i-x}\right) & \text{if } \gamma = 2 \\ \binom{i}{x}\mathbb{1}_{x=2} & \text{otherwise.} \end{cases} \tag{16}$$

Hence, for sample size $k \in \mathbb{N}$ the corresponding rate matrix of the ancestral process $\boldsymbol{Q}_\psi \in \mathbb{R}^{k\times k}$ is given by

$$\boldsymbol{Q}_\psi(i,x) = \begin{cases} \lambda_{i,x} & \text{if } i > x \\ -\sum_{m=2}^{i}\binom{i}{m}\lambda_{i,m} & \text{if } i = x \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

In particular, in the boundary case $\psi = 0$ we recover the rate matrix under the Kingman coalescent as

$$\boldsymbol{Q}_0(i,x) = \begin{cases} \binom{i}{2} & \text{if } x > i-1 \\ -\binom{i}{2} & \text{if } i = x \\ 0 & \text{otherwise.} \end{cases} \tag{18}$$

SFS-based maximum likelihood inference

In order to infer the coalescent model and its associated coalescent parameter and to (separately) estimate the demographic history of the population, Eldon *et al.* (2015) recently derived an (approximate) maximum likelihood framework based on the SFS. In the following we will give a concise overview of their approach which forms the basis for the joint inference of coalescent parameters and population growth rates.

First, let k denote the number of sampled (haploid) individuals (i.e., the number of leaves in the coalescent tree). Furthermore, let $\boldsymbol{\eta}^{(k)} = (\eta_1^{(k)}, \ldots, \eta_{k-1}^{(k)})$ denote the number of segregating sites with derived allele count of $i = 1, \ldots, k-1$ of all sampled individuals (i.e., the SFS), and let $s = \sum_{i=1}^{k-1} \eta_i$ be the total number of segregating sites. Provided that $s > 0$ we define the expected normalized SFS $\boldsymbol{\varphi}^{(k)} = (\varphi_1^{(k)}, \ldots, \varphi_{k-1}^{(k)})$ as

$$\varphi_i^{(k)} = \frac{\mathbb{E}\left[\eta_i^{(k)}\right]}{\sum_{i=1}^{k-1} \mathbb{E}\left[\eta_i^{(k)}\right]}, \tag{19}$$

which – given a coalescent model $\left(\Pi_{t,k}^{\psi,\rho}\right)_{t\geq 0}$ and assuming the infinite-sites model (Watterson 1975) – can be interpreted as the probability that a mutation appears i times in a sample of size k (Eldon *et al.* 2015). Furthermore, note that $\varphi_i^{(k)}$ is a function of $\left(\Pi_{t,k}^{\psi,\rho}\right)_{t\geq 0}$ (i.e., of the coalescent process and the demographic population history), but unlike $\mathbb{E}\left[\boldsymbol{\eta}^{(k)}\right]$ is not a function of the mutation rate, and should provide a good first-order approximation of the expected SFS as long as the sample size and the mutation rate are not too small (Eldon *et al.* 2015).

Then the likelihood function $\mathcal{L}\left(\left(\Pi_{t,k}^{\psi,\rho}\right)_{t\geq 0}, \tilde{\boldsymbol{\eta}}^{(k)}, s\right)$ for the observed frequency spectrum $\tilde{\boldsymbol{\eta}}^{(k)}$ and given coalescent model $\left(\Pi_{t,k}^{\psi,\rho}\right)_{t\geq 0}$ is given by

$$\begin{aligned}
\mathcal{L}\left(\left(\Pi_{t,k}^{\psi,\rho}\right)_{t\geq 0}, \tilde{\boldsymbol{\eta}}^{(k)}, s\right) &= \mathbb{P}^{\left(\Pi_{t,k}^{\psi,\rho}\right)_{t\geq 0}, s}\left(\eta_i^{(k)} = \tilde{\eta}_i^{(k)}, i \in [k-1]\right) \\
&= \mathbb{E}^{\Pi}\left[\frac{s!}{\tilde{\eta}_1^{(k)}! \ldots \tilde{\eta}_{k-1}^{(k)}!} \prod_{i=1}^{k-1} \left(\frac{T_i^{(k)}}{T_{\mathrm{tot}^{(k)}}}\right)^{\tilde{\eta}_i^{(k)}}\right] \\
&\approx \frac{s!}{\tilde{\eta}_1^{(k)}! \ldots \tilde{\eta}_{k-1}^{(k)}!} \prod_{i=1}^{k-1} \left(\varphi_i^{(k)}\right)^{\tilde{\eta}_i^{(k)}} \\
&\propto \prod_{i=1}^{k-1} \exp\left[-s\varphi_i^{(k)}\right] \frac{\left(s\varphi_i^{(k)}\right)^{\tilde{\eta}_i^{(k)}}}{\tilde{\eta}_i^{(k)}!}
\end{aligned} \tag{20}$$

(Eldon *et al.* 2015). Note that in the third line we approximated $\mathbb{E}\left[\left(\frac{T_i^{(k)}}{T_{\mathrm{tot}^{(k)}}}\right)\right] \approx \frac{\mathbb{E}\left[T_i^{(k)}\right]}{\mathbb{E}\left[T_{\mathrm{tot}^{(k)}}\right]} = \varphi_i^{(k)}$. In fact, Bhaskar *et al.* (2015) recently used a Poisson random field approximation to derive an analogous, structurally identical likelihood function for estimating demographic parameters under the Kingman coalescent. Notably though, their approximation assumes

that all sites in a given locus are completely unlinked, such that the underlying coalescent tree is independent at each site, and Equation 20 is exact.

As an alternative to the likelihood approach, we followed Eldon *et al.* (2015) and also implemented a minimal-distance statistic approach where

$$\hat{\psi}, \hat{\rho} = \underset{\psi,\,\rho}{\arg\min}\, d_p(\tilde{\boldsymbol{\eta}}^{(k)}, \boldsymbol{\eta}^{(k)}), \tag{21}$$

where d_p is some metric on $\mathbb{R}^{p-1}$.

Note though that both the likelihood and the distance-based approach require expressions for the expected normalized SFS $\boldsymbol{\varphi}^{(k)}$. Instead of performing Monte Carlo simulations to obtain these quantities we adapted an approach recently proposed by Spence *et al.* (2016), who derived analytical formulas for the expected SFS under a given (general) coalescent model $\left(\Pi_{t,k}^{\psi,\rho}\right)_{t\geq 0}$ and an intensity measure $\xi(t) : \mathrm{R}_{\geq 0} \to \mathrm{R}_{>0}$. In particular, the authors showed that

$$\mathbb{E}\left[\boldsymbol{\eta}^{(k)}\right] = \frac{\theta}{2}\mathbf{BCL}\mathfrak{c}_k, \tag{22}$$

where $\mathbf{B} \in \mathbb{R}^{k-1\times k-1}$ and $\mathbf{C} \in \mathbb{R}^{k-1\times k-1}$ are both $\Lambda-$independent (and thus easy to calculate) matrices, $\mathbf{L} \in \mathbb{R}^{k-1\times k-1}$ is a $\Lambda-$dependent lower triangular matrix that depends on the rate matrix $\mathbf{Q}$ and its spectral decomposition, θ is the population-scaled mutation rate, and $\mathfrak{c}_k = (\mathfrak{c}_{2,2}, \ldots, \mathfrak{c}_{k,k})$ denotes the expected time to the first coalescence for a sample of size $i \in \{2, \ldots, k\}$. Importantly, the time-inhomogeneity of the underlying coalescent process only enters through the first coalescence times $\mathfrak{c}_k$. For example, the first coalescence times for the Kingman coalescent with an exponentially growing population are given by

$$\mathfrak{c}_{i,i} = -\frac{1}{\rho}\exp\left[\frac{\binom{i}{2}}{\rho}\right]\mathrm{Ei}\left(-\frac{\binom{i}{2}}{\rho}\right), \tag{23}$$

where $\mathrm{Ei}(x) := -\int_{-x}^{\infty}(\exp\left[-t\right]/t)\mathrm{d}t$ denotes the exponential integral (Polanski *et al.* 2003; Polanski and Kimmel 2003; Bhaskar *et al.* 2015). Finally, plugging Equation 22 into Equation 19 leads to

$$\varphi_i^{(k)} = \frac{\left(\mathbf{BCL}\mathfrak{c}_k\right)_i}{\sum_{i=1}^{k-1}\left(\mathbf{BCL}\mathfrak{c}_k\right)_i}, \tag{24}$$

highlighting that θ cancels and that the likelihood function (eq. 20) is independent of the mutation rate.

To obtain the coalescent parameter ψ and population growth rate ρ that maximize the likelihood function (eq. 20) or respectively minimize the distance function (eq. 21), we used a grid search procedure over an equally-spaced two-dimensional grid with $\psi_{\text{grid}} = \{0, 0.01, \ldots, 1\}$ and $\rho_{\text{grid}} = \{0, 1, \ldots, 1024\}$, and evaluated the value of the likelihood respectively distance function at each grid point.

Data availability

The empirical raw data used have been downloaded from GenBank (accession numbers LC031518 – LC031623; data from Niwa *et al.* 2016). The empirical SFS can be downloaded

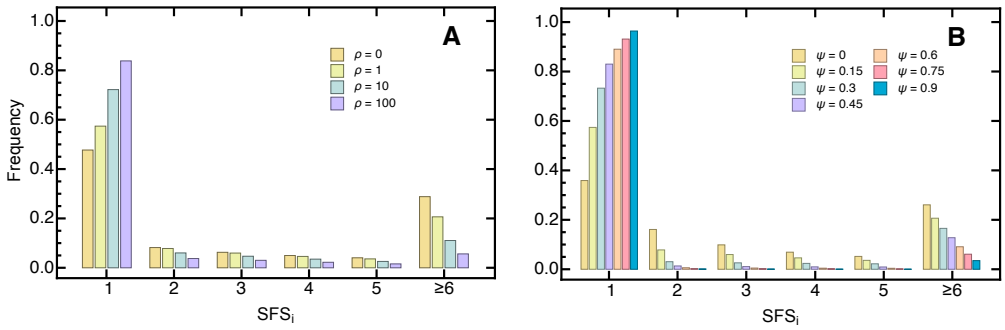

Figure 2 – The expected normalized (lumped) SFS for the psi-coalescent for an exponentially growing population (eq. 24) with sample size $k = 20$ (**A**) for different values of ψ and fixed $\rho = 1$ and (**B**) for different values of ρ and fixed $\psi = 0.15$. The sixth entry in the SFS contains the aggregate of the higher frequency classes.

from Supporting Information (Supporting Files: File E1). The simulation program and the inference program were written in C++ and are available upon request.

Results and Discussion

The aim of this work is to derive the ancestral process for an exponentially expanding population that undergoes sweepstake reproductive events. We first derive the time-inhomogeneous Markovian ancestral process that underlies the extended Moran model, and show that, analogous to the Kingman coalescent, it can be described by a time-homogeneous Markov chain on a non-linear time scale. In particular, we derive the coalescent rates and the time-change function, and prove convergence to a $\Lambda-$coalescent with Dirac measure at ψ. Detailed derivations of the results, which in the main text have been abbreviated to keep formulas concise, can be found in Supporting Information (Detailed derivation of results). On the basis of these results, we derive a maximum likelihood inference framework for the joint inference of the coalescent parameter and the population growth rate, and assess its accuracy and performance through large-scale simulations. Furthermore, we quantify the bias of coalescent and population growth parameter estimates when mistakenly neglecting population demography or reproductive skew. Finally, we apply our approach to mtDNA from Japanese sardine (*Sardinops melanostictus*) populations where patterns of sequence variation were shown to be more consistent with sole influence from sweepstake reproductive events, again highlighting the potential mis-inference of growth if reproductive skew is not properly accounted for (Niwa *et al.* 2016; Grant *et al.* 2016).

Derivation of the ancestral limit process

Unlike in the case of a constant-size population, the sequence of the number of offspring $(U_N(n))_{n\in\mathbb{N}}$ changes along with the (time-dependent) population size. Thus, the ancestral process is characterized by an inhomogeneous Markov chain with transition probabilities

$$G_{i,x}^{(n)} = \binom{i}{x} \sum_{u=2}^{N_n} \frac{\mathbb{P}(U_N(n) = u)(u)_x (N_n - u)_{i-x}}{(N_n)_i} \tag{25}$$

and time-dependent coalescence probability

$$c_N^{(n)} := G_{2,2}^{(n)} = \frac{\mathbb{E}\left[U_N(n)^2 - U_N(n)\right]}{N_n(N_n-1)} \tag{26}$$

for all $n \in \mathbb{N}$. Plugging Equation 5 into Equations 25 and 26, and using Equation 4 then

$$G_{i,x}^{(n)} = \binom{i}{x} \frac{1}{N_n\left(N_n-1\right)} \tag{27}$$

yields

$$\Big((1-N_n^{\gamma})\left(\mathrm{Max}[2,\Delta_N^{(n)}+1]\right)_x \left(N_n - \mathrm{Max}[2,\Delta_N^{(n)}+1]\right)_{i-x} + N_n^{-\gamma}(\mathrm{Max}[\psi N_n,\Delta_N^{(n)}+1])_x \left(N_n - \mathrm{Max}[\psi N_n,\Delta_N^{(n)}+1]\right)_{i-x}\Big)$$

and

$$c_N^{(n)} = \frac{(1-N_n^{\gamma})\left(\mathrm{Max}[2,\Delta_N^{(n)}+1]\right)_2 + N_n^{-\gamma}(\mathrm{Max}[\psi N_n,\Delta_N^{(n)}+1])_2}{N_n\left(N_n-1\right)} \tag{28}$$

respectively. Note, that Equation 28 is the weighted sum of the number of offspring for the two different reproductive events. Furthermore, taking the limit $N \to \infty$ in Equation 3

$$\begin{aligned} \lim_{N\to\infty} \Delta_N^{(n)} &= \lim_{N\to\infty} \left(N_n - N_{n-1}\right) \\ &\leq \lim_{N\to\infty} N\rho\frac{\psi^2}{N^{\gamma}} \\ &= 0 \end{aligned} \tag{29}$$

shows that $\Delta_N^{(n)}$ is bounded for all $n \in \mathbb{N}$ and thus allows dropping of the maxima condition in Equations 27 and 28. Furthermore, for sufficiently large N Equation 28 becomes

$$c_N^{(n)} = \frac{(1-N_n^{\gamma})\,2 + \psi N_n^{1-\gamma}\left(\psi N_n - 1\right)}{N_n\left(N_n-1\right)} \sim \frac{\psi^2}{N^{\gamma}}. \tag{30}$$

To prove that the time-scaled ancestral process of the underlying extended Moran model converges to a continuous-time Markov chain as the initial population size approaches infinity, we apply *Theorem 2.2* in Möhle (2002), which requires the following definitions: First, consider a step function $F_N : [0,\infty) \to [0,\infty)$ given by

$$F_N(s) := \sum_{n=1}^{\lfloor s \rfloor} c_N^{(n)}. \tag{31}$$

Furthermore, let $\mathcal{G}_N^{-1}$ denote a modification of the right-continuous inverse of F_N

$$\mathcal{G}_N^{-1}(t) := \inf\{s > 0 \mid F_N(s) > t\} - 1 \tag{32}$$

which will constitute the time-change function in the following. Since by assumption $\lim_{s\to\infty} F_N(s) = \infty$ for all $s \in [0,\infty)$ it follows that $\mathcal{G}_N^{-1}(t)$ is finite for all $t \in [0,\infty)$. Finally, *Theorem 2.2* (Möhle 2002) requires that for all $t \in [0\infty)$

$$\lim_{N\to\infty} \inf_{1\le r\le \mathcal{G}_N^{-1}(t)} N_r = \infty \tag{33}$$

and

$$\lim_{N\to\infty} \sup_{1\le r\le \mathcal{G}_N^{-1}(t)} c_N^{(r)} = 0 \tag{34}$$

holds, i.e., that – on the new time scale – the population size remains large while the coalescent probabilities become small.

Then, let $\left(\mathcal{A}_{n,k}^{\psi,\rho}\right)_{n\in\mathbb{N}}$ denote the ancestral process of the extended Moran model with exponential growth (see Model and Methods) and let ϕ and ξ denote two partitions of $[k]$ with $\xi \subset \phi$ of size a and $b = b_1 + \ldots + b_a$ (where $b_1 \ge b_2 \ge \ldots \ge b_a \ge 1$), respectively. The transition probability of $\left(\mathcal{A}_{n,k}^{\psi,\rho}\right)_{n\in\mathbb{N}}$ at time $n \in \mathbb{N}$ is given by

$$\Phi(n; b_1, b_2, \ldots, b_a) := \frac{1}{(N_{n-1})_b} \sum_{\substack{i_1,\ldots,1_a=1 \\ \text{all distinct}}}^{N_n} \mathrm{E}\left((\nu_{i_1}(n))_{b_1} \cdots (\nu_{i_a}(n))_{b_a}\right). \tag{35}$$

Thus, for the extended Moran model with exponential growth *Theorem 2.2* in Möhle (2002) states:

Theorem 1 (*Theorem 2.2*; Möehle 2002)**.** *Assume that Equations (33) and (34) hold and for all $t \in \mathbb{R}_{>0}$ the limit*

$$\pi\left((0,t]; b_1, \ldots, b_a\right) := \lim_{N\to\infty} \sum_{n=1}^{h_N(t)} \Phi(n; b_1, b_2, \ldots, b_a) \tag{36}$$

exists. Then for each sample size $k \in \mathbb{N}$ the ancestral process $\left(\mathcal{A}_{\mathcal{G}_N^{-1}(t),k}^{\psi,\rho}\right)_{t\ge 0}$ converges as N tends to infinity to a time-continuous and in general a time-inhomogeneous Markov chain $\left(\Pi_{t,k}^{\psi,\rho}\right)_{t\ge 0}$.

Note though that in its general form Theorem 1 was derived for any generic Cannings model as well as any kind of population size change (Möhle 2002).

We will now derive our first main result and show that the ancestral limiting process $\lim_{N\to\infty}\left(\mathcal{A}_{n,k}^{\psi,\rho}\right)_{t\ge 0}$ converges to a $\Lambda - k-$coalescent on a non-linear time scale. First, we derive the time-change function $\mathcal{G}_N^{-1}(t)$ for the ancestral process by considering the step function (eq. 31)

$$\begin{aligned} F_N(s) &= \sum_{n=1}^{s} c_N(n) \sim \sum_{n=1}^{s} \frac{\psi^2}{N_n^\gamma} \\ &= \frac{\psi^2}{N^\gamma}\left(1 - \rho\frac{\psi^2}{N^\gamma}\right)^{-\gamma} \frac{\left(\left(1-\rho\frac{\psi^2}{N^\gamma}\right)^{-\gamma s} - 1\right)}{\left(1-\rho\frac{\psi^2}{N^\gamma}\right)^{-\gamma} - 1}. \end{aligned} \tag{37}$$

Solving for s then gives

$$\begin{aligned}\mathcal{G}_N^{-1}(t) &= \inf\{s > 0 : F_N(\lfloor s \rfloor)\} \\ &\sim \left[\frac{\log[1+\rho\gamma t]}{\rho\gamma}\frac{N^\gamma}{\psi^2}\right], \end{aligned} \tag{38}$$

where we have used $\log\left[1 - \frac{\rho\psi^2}{N^\gamma}\right] \sim -\frac{\rho\psi^2}{N^\gamma}$ for sufficiently large N. In particular, we have

$$\mathcal{G}^{-1}(t) := \lim_{N\to\infty} \mathcal{G}_N^{-1}(t) c_N^{(0)} = \frac{\log[1+\rho\gamma t]}{\rho\gamma}. \tag{39}$$

Furthermore, Equations 33 and 34 hold since

$$\lim_{N\to\infty} \inf_{1\le n\le \mathcal{G}_N^{-1}(t)} N\left(1 - \rho\frac{\psi^2}{N_n^\gamma}\right)^n \tag{40}$$

$$\begin{aligned} &= \lim_{N\to\infty} N \inf_{1\le n\le \mathcal{G}_N^{-1}(t)} \exp\left(-\frac{\rho\psi^2}{N^\gamma} n\right) \\ &= \lim_{N\to\infty} N(1+\rho\gamma t)^{-\frac{1}{\gamma}} \\ &= \infty \quad \forall t \in (0,\infty), \end{aligned} \tag{41}$$

and by the same reasoning

$$\lim_{N\to\infty} \sup_{1\le n\le \mathcal{G}_N^{-1}(t)} \frac{\psi^2}{N\left(1-\rho\frac{\psi^2}{N^\gamma}\right)^n} = 0, \quad \forall t \in (0,\infty). \tag{42}$$

Finally, to show that Equation 36 holds, we first note that

$$\Phi_a(r; b_1, ..., b_a) = 0\,, \tag{43}$$

and for $a \ge 2$ and there are two indices $1 \le i < j \le a$ with $b_i, b_j \ge 2$

$$\pi_a((0,t]; b_1, ..., b_a) = 0, \tag{44}$$

since the extended Moran model does not allow for more than one reproductive event at a time.

Thus, $\lim_{N\to\infty}\left(\mathcal{A}^{\psi,\rho}_{\mathcal{G}_N^{-1}(t),k}\right)_{t\ge 0}$ is well defined and does not feature any simultaneous coalescent events, implying that the limiting process must be a (possibly time-inhomogeneous) $\Lambda - k-$ coalescent. Further, for $a = 1$,

$$\begin{aligned}\lim_{N\to\infty} \sum_{n=1}^{\mathcal{G}_N^{-1}(t)} \Phi_1^{(N)}(r; b) &= \lim_{N\to\infty} \sum_{n=1}^{\mathcal{G}_N^{-1}(t)} G_{b,b} \\ &= \psi^{b-2} t. \end{aligned} \tag{45}$$

Hence, Theorem 1 implies that for each sample size $k \in \mathbb{N}$ the limit of the time-scaled ancestral process $\left(\mathcal{A}^{\psi,\rho}_{\mathcal{G}_N^{-1}(t),k}\right)_{t\geq 0}$ exists and from Equation 45 it follows that $\lim_{N\to\infty}\left(\mathcal{A}^{\psi,\rho}_{\mathcal{G}_N^{-1}(t),k}\right)_{t\geq 0}$ is a time-homogeneous $\Lambda - k-$coalescent. Further,

$$t\psi^{b-2} = t\int_0^1 x^{b-2}\Lambda(dx) \tag{46}$$

holds for all $b \in \mathbb{N}$ if and only if Λ is the Dirac measure at ψ. Thus, $\lim_{N\to\infty}\left(\mathcal{A}^{\psi,\rho}_{\mathcal{G}_N^{-1}(t),k}\right)_{t\geq 0}$ is a psi$-k-$coalescent, implying that $\left(\Pi^{\psi,\rho}_{t,k}\right)_{t\in\mathbb{R}_{\geq 0}} := \lim_{N\to\infty}\left(\mathcal{A}^{\psi,\rho}_{\lfloor t/c_N^{(0)}\rfloor,k}\right)_{t\geq 0}$ such that $\left(\Pi^{\psi,\rho}_{t,k}\right)_{t\in\mathbb{R}_{\geq 0}}$ is equal (in distribution) to $\left(\Pi^{\psi,0}_{\mathcal{G}(t),k}\right)$ (i.e., a regular psi$-k-$coalescent) with time (non-linearly) rescaled by

$$\mathcal{G}(t) = \frac{\exp\left(\rho\gamma t\right) - 1}{\rho\gamma}. \tag{47}$$

Put differently, analogous to the results obtained for the Kingman coalescent (Griffiths and Tavaré 1998), the time-inhomogeneous ancestral limiting process of the extended Moran model with exponential growth can be transformed into a time-homogeneous psi-coalescent with coalescent rates given by Equation 16 with branches rescaled by Equation 47 allowing it to be simulated easily and efficiently. To highlight the duality between the two processes (i.e., the (forward in time) extended Moran model and the corresponding coalescent), key properties (e.g., the summed length of all branches with i descendants T_i and the total tree length T_{total}) are compared in the Supporting Information (Extended Moran model simulations). Finally note, that Equation 47 is – except for the additional factor γ that is proportional to the coalescent time scale – structurally identical to the time-change function in the Kingman case (Griffiths and Tavaré 1998). However, since $\mathcal{G}(t)$ depends on the product $\tilde{\rho} = \rho\gamma$, it is impossible to obtain a direct estimate of ρ (or γ) without additional information, and thus – analogous to the case of the population scaled mutation rate θ – only the compound parameter $\tilde{\rho}$ can be estimated.

Joint inference of coalescent parameters and population growth rates

In this paragraph we modify the likelihood function

$$\mathcal{L}\left(\left(\Pi^{\psi,\rho}_{t,k}\right)_{t\geq 0}, \tilde{\boldsymbol{\eta}}^{(k)}, s\right) \propto \prod_{i=1}^{k-1} \exp\left[-s\varphi_i^{(k)}\right] \frac{\left(s\varphi_i^{(k)}\right)^{\tilde{\eta}_i^{(k)}}}{\tilde{\eta}_i^{(k)}!} \tag{48}$$

derived in the Model and Methods section to jointly infer the coalescent parameter ψ and the population growth rate ρ. Note that while the general form of the likelihood function (eq. 20) is independent of the generating coalescent process, changes in ψ and ρ affect the expected normalized SFS as given by

$$\varphi_i^{(k)} = \frac{\left(\mathbf{BCL}\mathfrak{c}_k\right)_i}{\sum_{i=1}^{k-1}\left(\mathbf{BCL}\mathfrak{c}_k\right)_i}. \tag{49}$$

Recall that $\mathbf{B}$ and $\mathbf{C}$ depend neither on ψ nor ρ, and that $\mathbf{L}$ does depend on ψ but not on ρ, and that the time-inhomogeneity of the underlying coalescent process only enters through

the first coalescence times $\mathfrak{c}_k$, which are given by

$$\begin{aligned}\mathfrak{c}_{i,i} &= \int_0^\infty \mathbb{P}(\text{time of first coalescence for } i \text{ individuals} > t)dt \\ &= \int_0^\infty \exp\left((\mathbf{Q})_{i,i}\int_0^t (1/\xi(s))ds\right)dt, \end{aligned} \tag{50}$$

where $\xi(s)$ denotes the intensity measure (Polanski and Kimmel 2003; Bhaskar *et al.* 2015; Spence *et al.* 2016). For the psi-coalescent with exponential growth $\xi(t) = e^{-\rho\gamma t}$ such that Equation 50 becomes

$$\begin{aligned}\mathfrak{c}_{i,i} &= \int_0^\infty \exp\left((\mathbf{Q})_{i,i}\frac{\exp(\rho\gamma t) - 1}{\rho\gamma}\right)dt \\ &= -\frac{\exp((\mathbf{Q})_{i,i}\,\rho\gamma)}{\rho\gamma}\mathrm{Ei}(-\mathbf{Q})_{i,i}\,\rho\gamma), \end{aligned} \tag{51}$$

where $\mathrm{Ei}(x) := -\int_{-x}^{\infty}(\exp[-t]/t)\mathrm{d}t$ denotes the exponential integral. Thus, when growth rates are measured on their corresponding coalescent scale – i.e., $\tilde{\rho}$ under the psi-coalescent versus ρ under the Kingman coalescent – Equation 51 is a generalization of the Kingman coalescent result (eq. 23) derived by Polanski and Kimmel (2003). Finally, combining Equation 51 with Equation 49 allows for the *exact* computation of the expected normalized SFS $\boldsymbol{\varphi}^{(k)}$, avoiding the simulation error that would be introduced by Monte Carlo simulations.

Figure 2 shows the expected normalized SFS obtained from Equation 49, where higher frequency classes have been aggregated (i.e., lumped) for different values of ψ and ρ. In line with previous findings, both multiple mergers and population growth lead to an excess in singletons (Durrett and Schweinsberg 2005; Eldon *et al.* 2015; Niwa *et al.* 2016). Furthermore, this excess increases as sample size increases under the psi-coalescent (Fig. SI C_1), while it decreases for the Kingman coalescent independent of the presence or absence of exponential growth. These qualitative differences stem from the different footprints reproductive skew and exponential growth leave on a genealogy. While the latter is a simple rescaling of branch lengths leaving the topology unchanged, multiple-merger coalescents by definition affect the topology of the genealogical tree (Eldon *et al.* 2015). In particular, when ψ is large, adding samples will disproportionally increase the number of external branches T_1 such that the genealogy will become more star-like, rendering disproportionately more singletons.

Though the excess in singletons characterizes either process, their higher frequency classes will typically differ (Eldon *et al.* 2015). When both processes – reproductive skew and exponential growth – act simultaneously though, their joint effects on the SFS (non-trivially) combine. As expected, increasing growth under the psi-coalescent further exacerbates the excess in singletons. More generally, exponential growth leads to a systematic left shift in the SFS towards lower frequency classes that is independent of ψ. Increasing ψ on the other hand changes the SFS – and in particular the higher frequency classes – non-monotonically even if there is no population growth (Fig. SI C_2). Interestingly, for $\rho = 0$ the last entry of the expected normalized SFS $\mathbb{E}[\eta_{k-1}]$ initially increases with ψ and takes an intermediate maximum, decreases monotonically until $\psi \approx 0.85$, peaks again and then quickly reduces to 0

as ψ approaches 1. This effect prevails as sample size increases (Fig. SI C_2) even though the intermediate maximum slightly shifts towards lower ψ. However, this intermediate maximum is effectively washed out by increasing ρ such that the second peak becomes the maximum. Furthermore, the shape of the peak becomes more pronounced as sample size increases. Thus, reproductive skew and exponential growth leave complex and distinct genomic footprints on the SFS. While in theory population growth and reproductive skew should be identifiable, this in practice strongly depends on sample size (Spence *et al.* 2016). In the next section we will assess the accuracy of our joint estimation framework and perform extensive validation (eq. 20) on large-scale simulated data.

Simulated coalescent and demographic models

To test our inference algorithm, we followed two different simulation approaches, each corresponding to two biological limiting cases. In both, data was simulated for the Cartesian product set over $\psi = \{0, 0.15, 0.3, 0.45, 0.6, 0.75, 0.9\}$, $\rho = \{0, 1, 10, 100\}$, $k = \{20, 50, 100, 200\}$ and $s = \{100, 1\,000, 10\,000\}$ per locus over 10,000 replicates each. Note that we follow Eldon *et al.* (2015) and employ the fixed-s method (see also Depaulis *et al.* 1998; Ramos-Onsins and Rozas 2000), where – instead of treating the number of segregating sites as an observation of a random variable that depends on the population scaled mutation rate θ – the number of segregating sites is fixed, and s mutations are placed uniformly at random along the branches of the coalescent tree. For an in-depth discussion of the fixed-s approach and its implications when constructing statistical tests, we refer the reader to Eldon *et al.* (2015) (and references therein). The main motivation for using the fixed-s approach in this study was to make results comparable across different values of ψ and ρ given an observed number of segregating sites. Conversely, keeping s constant implies that θ increases with ψ and ρ. We will discuss the latter point in more detail in light of the results below. Data was simulated for the following two underlying genetic architectures:

- Case 1 (linkage equilibrium): When each site in a given locus is unlinked the underlying coalescent tree at each site is independent (Sawyer and Hartl 1992; Bhaskar *et al.* 2015). Thus, by averaging over independent realizations of the (shared) underlying coalescent process the SFS can be obtained by randomly drawing from a multinomial distribution such that $\boldsymbol{\eta} \sim \text{Multinomial}\left(s, \boldsymbol{\varphi}\right)$.

- Case 2 (complete linkage): In this scenario we consider a genome of 1,000 unlinked loci, where sites within each locus are completely linked (i.e., linkage disequilibrium). Thus, for each locus we draw a random genealogy according to equations 16 and 47, superimpose s random mutations onto the ancestral tree by multinomial sampling, and aggregate the individual locus SFS into a single genome-wide SFS.

Finally, data sets where $s = \eta_1$ (i.e., where all segregating sites were singletons) were discarded and simulated again since these do not allow the underlying coalescent parameter and demographic history to be identified.

Accuracy of joint estimation framework

Next we evaluated the accuracy of the joint estimation framework by means of the mean absolute deviation (MAD) $MAD = \frac{1}{n}\sum_{i=1}^{n}|x_i - \hat{x}_i|$, the mean deviation (MD) $MD =$

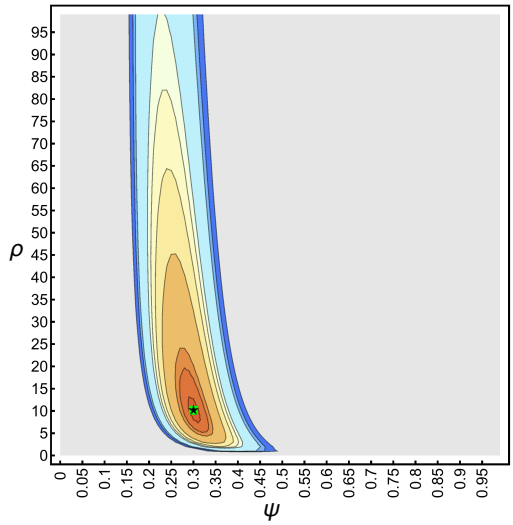

Figure 3 – Likelihood surface (eq. 20) of the idealized SFS with $k = 100$, $\psi = 0.3$, $\rho = 10$ and $s = 10{,}000$. Contours show the $0.95, 0.9675, 0.975, 0.99$, $0.99225, 0.9945$, $0.99675, 0.999, 0.99945$ and 0.9999 quantiles. Likelihoods below the 0.95 quantile are uniformly colored in gray. The green square shows the true ψ and ρ. The black star shows the maximum likelihood estimates $\hat{\psi}$ and $\hat{\rho}$.

$\frac{1}{n}\sum_{i=1}^{n} x_i - \hat{x}_i$, the mean squared error (MSE) $MSE = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x}_i)^2$, and the median deviation (MDD) where x and $\hat{x}$ denote the true and the estimated parameter, respectively. If not stated otherwise, results in the main text are shown for the default parameters $k = 100$ and $s = 10{,}000$. More results are given in Supporting Information (Supporting Figures) and (Supporting Tables).

Inference assuming linkage equilibrium:. First, for a consistency check we applied our grid-search algorithm to estimate ψ and ρ from an idealized SFS (i.e., where the SFS accurately reflects the expected branch length under the generating coalescent and demographic model $\boldsymbol{\varphi}$ except for distortions due to rounding). An exemplary likelihood surface (eq. 20) for such an idealized SFS is depicted in Figure 3, which shows that the likelihood surface – up to the resolution of the grid point – is smooth and generally unimodal and that the true parameters can be estimated accurately. Furthermore, Figure 3 shows that there is generally a negative correlation between ψ and ρ, and that the likelihood surface tends to be steeper and more concentrated along the ψ direction, which suggests that growth rate estimates might show a larger variance and could in general be more difficult to estimate. The steepness of the likelihood surface along the ψ axis tends to increase with ψ and sample size k, suggesting that the accuracy for estimating ψ should increase as well, while it should become more difficult to estimate ρ accurately.

An exemplary distribution of the jointly inferred maximum likelihood estimates $(\hat{\psi}, \hat{\rho})$ assuming linkage equilibrium is shown in Figure 4. The shape of this distribution resembles that of the likelihood surface (Fig. 3), indicating that there is some variance – in particular along the ρ-axis – in the maximum likelihood estimates. However, the median and the mean of the distribution match the true underlying coalescent and growth rate parameters (i.e., ψ and ρ) very well, implying that under linkage equilibrium (i.e., complete independence between sites) $\hat{\psi}$ and $\hat{\rho}$ are unbiased estimators.

Generally, as expected from the shape of the likelihood surface, ψ is estimated with high

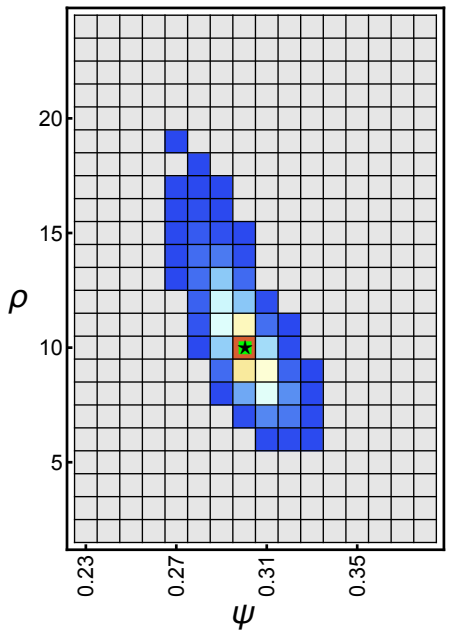

Figure 4 – Heatplot of the frequency of the maximum likelihood estimates for 10,000 data sets assuming linkage equilibrium with $k = 100$, $\psi = 0.3$, $\rho = 10$ and $s = 10{,}000$. Counts increase from blue to red with grey squares showing zero counts. The green square shows the true ψ and ρ. The black star shows the median (and mean) of the maximum likelihood estimates $\hat{\psi}$ and $\hat{\rho}$.

accuracy and precision, even for large sample sizes ($k = 200$) with only a few segregating sites ($s = 100$) and (nearly) independent of ρ (Fig. 5A, SI C_3A; Table SI D_1). However, as expected, growth rate estimates $\hat{\rho}$ show a larger variance and for some parameters – namely large k and small s – might be slightly upwardly biased when both the coalescent parameter and the growth rate are large (Fig. 5B, SI C_3B; Table SI D_2). Though, as the number of segregating sites increases, this bias vanishes and the variance decreases (Fig. SI C_4), highlighting that the joint estimation procedure gives asymptotically unbiased estimators.

For constant s, increasing sample size k increases the signal-to-noise ratio and thus the error in both $\hat{\psi}$ and $\hat{\rho}$ (Tables SI D_1, SI D_2, SI D_3, SI D_3) which is most noticeable

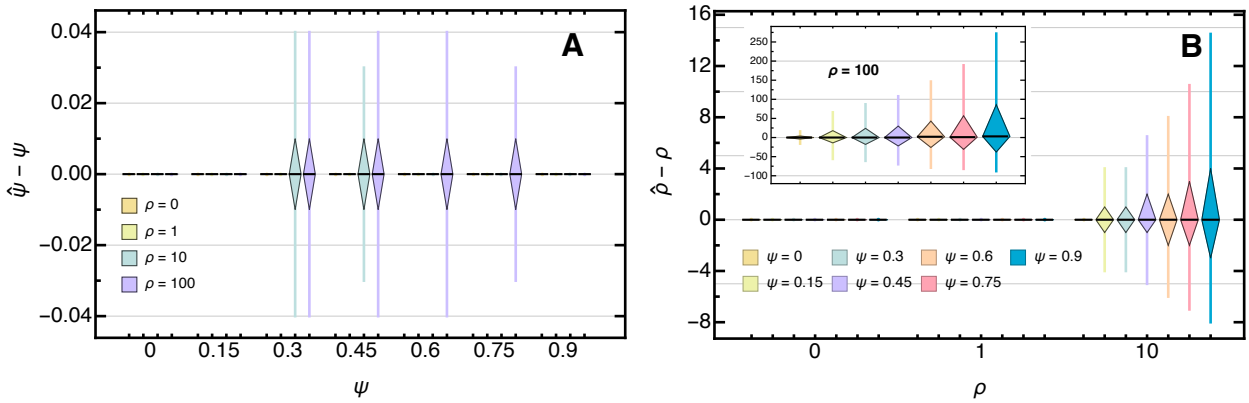

Figure 5 – Boxplot of the deviation of the maximum likelihood estimate from the true (**A**) ψ and (**B**) ρ for 10,000 data sets assuming linkage equilibrium with $k = 100$ and $s = 10{,}000$. Boxes represent the interquartile range (i.e., the 50% C.I.) and whiskers extend to the highest/lowest data point within the box ± 1.5 times the interquartile range.

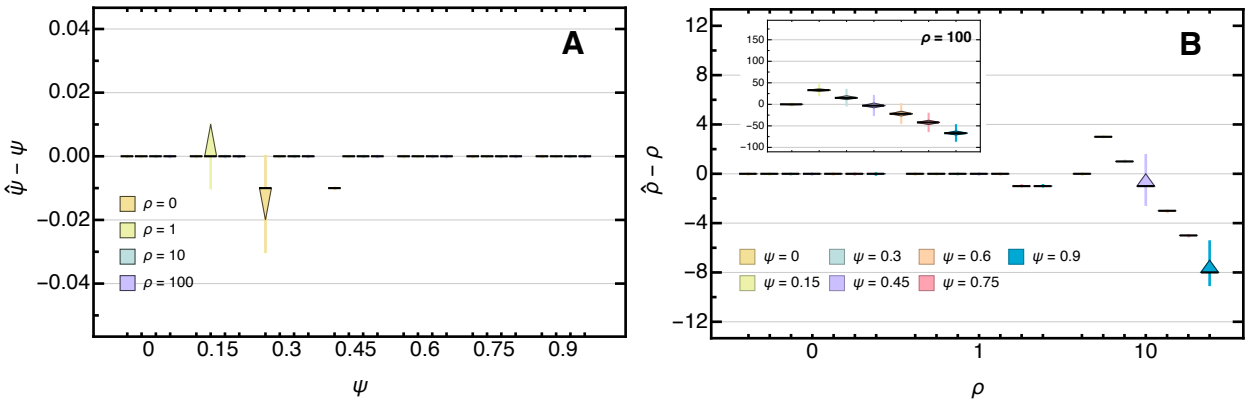

Figure 6 – Boxplot of the deviation of the maximum likelihood estimate from the true (**A**) ψ and (**B**) ρ for 10,000 data sets assuming complete linkage with $k = 100$ and $s = 10{,}000$. Boxes represent the interquartile range (i.e., the 50% C.I.) and whiskers extend to the highest/lowest data point within the box ± 1.5 times the interquartile range.

in growth rate estimates, in particular when ρ is large (Fig. SI C_5). This increase in estimation error can (partially) be compensated by increasing the number of segregating sites s (Fig. SI C_6, Table SI D_5). Specifically, if the true underlying ψ is large (i.e., if the offspring distribution is heavily skewed) an increasing number of segregating sites is needed to accurately infer ρ. However, the total tree length $\bar{T}_{\text{Total}}$ – and thus the number of segregating sites s – is expected to decrease sharply with ψ (Eldon and Wakeley 2006), implying that trees tend to become shorter under heavily skewed offspring distributions. This effect could (again, partially) be overcome by increasing sample size since $\bar{T}_{\text{Total}}$ – unlike the Kingman coalescent – scales linearly with k as ψ approaches 1 (Eldon and Wakeley 2006). However, population growth will reduce $\bar{T}_{\text{Total}}$ and the number of segregating sites even further.

Fixing the number of segregating sites for the assessment of the accuracy of the estimation method (i.e., the fixed-s approach) evades this problem to some extent. Strictly speaking, θ in our simulations increases with ψ and ρ. However, our results suggest that even more segregating sites (i.e., an even larger θ) would be needed to accurately infer population growth. Thus, unless (effective) population sizes and/or genome-wide mutation rates are large it might be very difficult to infer population growth if the offspring distribution is heavily skewed. On the other hand, the few studies that have estimated ψ generally found it to be small (Eldon and Wakeley 2006; Birkner *et al.* 2013; Árnason and Halldórsdóttir 2015) leaving it unresolved whether this problem is of any practical importance when studying natural populations.

Inference assuming complete linkage:. Figure 6 indicates the performance of the joint estimation method when considering a genome of 1,000 unlinked loci, where sites within each locus are completely linked and thus in linkage disequilibrium. Again ψ can be estimated accurately, except for $\psi = 0.3$ and $\psi = 0.45$ with $\rho = 0$ which both seem to be systematically underestimated. When growth rates are large (e.g., $\rho \geq 10$) and the underlying coalescent is not the Kingman coalescent (i.e., if $\psi \neq 0$), ρ is constantly either over- or underestimated, depending on the exact value of ψ. When fitting a linear model to the mean deviation against $\psi \neq 0$ (i.e., when excluding the Kingman coalescent) for $\rho = 10$ and $\rho = 100$ we find a clear (negative) linear relation ($R^2 = 0.997$ and 0.996, respectively) indicative of a systematic bias

introduced when assuming complete linkage within each locus. This effect persists when increasing the number of segregating sites and/or the number of loci, which both only decrease the variance in the estimates, and/or when using a finer grid (to exclude biases due to poor resolution), highlighting that this bias does not arise from a lack of data (results not shown).

The underlying cause is rather that the (Poisson random field) approximation in the likelihood function (eq. 20) breaks down since the underlying coalescent trees at each site are no longer independent. A similar observation – i.e., that linkage disequilibrium causes the likelihood function to become increasingly inaccurate – has recently been made by Bhaskar *et al.* (2015). We find that this effect is amplified as the skewness in the offspring distribution (i.e., ψ) increases, consistently biasing growth rate estimates downwards. Note that we expect the same phenomenon to occur with smaller growth rates as well, but on a smaller absolute scale that is unobservable due to the resolution set by the coarseness of the underlying grid. The systematic deviation for $\psi = 0.3$ and 0.45 with $\rho = 0$ suggests that there might be similar effects when estimating ψ. However, as argued above these might simply be unobservable within our resolution.

Distance-based inference and the effect of lumping:. As an alternative to the likelihood-based method, Eldon *et al.* (2015) proposed an ABC approach based on a minimum-distance statistic (eq. 21). In this section we assess the accuracy of $\hat{\psi}_d$ and $\hat{\rho}_d$ for d_1 and d_2 (i.e., l_1 and l_2 distance). We find that for the l_1 and the l_2 distance results are comparable to those of the likelihood-based estimates, but generally display a larger variance (Fig. SI C_7, SI C_8). Likelihood-based estimates $\hat{\psi}_{ML}$ tend to be more accurate across the entire parameter space though differences between the two are marginal. Over the majority of the parameter space the same holds true for $\hat{\rho}_{ML}$. Particularly for small to intermediate ψ the likelihood-based approach outperforms both distance-based approaches considerably (Tables SI D_6, SI D_7). Interestingly though, for large ψ and ρ (i.e., in the part of the parameter space where estimating ρ is generally difficult) the l_1 distance approach give more accurate estimates. When increasing the number of segregating sites, though, the likelihood approach becomes more accurate again, suggesting that the l_1 distance-based approach only outperforms the likelihood-based approach when there is insufficient data (not shown). Note that under complete linkage, $\hat{\rho}_d$ displays the same systematic bias as $\hat{\rho}_{ML}$, consistently biasing growth rate estimates downwards as ψ increases (not shown), due to its dependence on the expected normalized SFS $\boldsymbol{\varphi}^{(k)}$. However, given the asymptotic properties, the l_1 and the l_2 distance should perform reasonably well when used in a rejection-based ABC analysis (Fig. SI C_9).

Finally, we investigated the effect of lumping (i.e., aggregating the higher-frequency classes of the SFS into a single entry after a given threshold i) on the performance of our estimator. In contrast to Eldon *et al.* (2015), who found that lumping can improve the power to distinguish between multiple-merger coalescent models and models of population growth, we find that estimates based on the lumped SFS (using $i = 5$ and $i = 15$) show considerably more error (Tables SI D_8, SI D_9). While ψ can again be reasonably well estimated, $\hat{\rho}$ – in particular when ψ and/or ρ are large – is orders of magnitude more inaccurate when higher frequency classes are lumped. The reason is that when trying to differentiate between different coalescent or growth models, lumping can reduce the noise associated with the individual higher frequency classes and thus increases the power, provided that the different candidate models show different mean behaviors in the lumped classes (Eldon *et al.* 2015). While this

seems to hold true when considering "pure" coalescent or growth models, the joint footprints of skewed offspring distributions and (exponential) population growth are more subtle. In particular, since growth induces a systematic left shift in the SFS towards lower frequency classes, most of the information to distinguish between a psi-coalescent with or without growth is lost when aggregated.

Mis-inference of coalescent parameters when neglecting demography

As argued above, both reproductive skew and population growth result in an excess of singletons (i.e., low-frequency mutations) in the SFS. However, topological differences between the two generating processes in the right tail of the SFS allows distinguishing between the two. In particular, fitting an exponential growth model and not accounting for reproductive skewness results in a vastly (and often unrealistically) overestimated growth rate (Eldon *et al.* 2015). Here, we investigate how coalescent parameter estimates (i.e., $\hat{\psi}$) are affected when not accounting for (exponential) population growth (i.e., assuming $\rho = 0$) when both processes act simultaneously.

As expected we find that $\hat{\psi}$ is consistently overestimated (Fig. 7) and that the estimation error – independent of ψ – increases with larger (unaccounted for) growth rates. This is because, unless the underlying genealogy is star-shaped (e.g., when $\psi = 1$), growth will always left-shift the SFS and hence increase the singleton class. Thus, when assuming $\rho = 0$, increasing ψ compensates for the "missing" singletons. Interestingly though, the estimation error changes non-monotonically with ψ, and for large ρ can be as great as twice the value of the true underlying coalescent parameter. Furthermore, for low to intermediate ψ, even small growth rates can result in a relative error of up to 23%. Overall, not accounting for demography can lead to serious biases in ψ with broad ecological implications when trying to understand the variation in reproductive success.

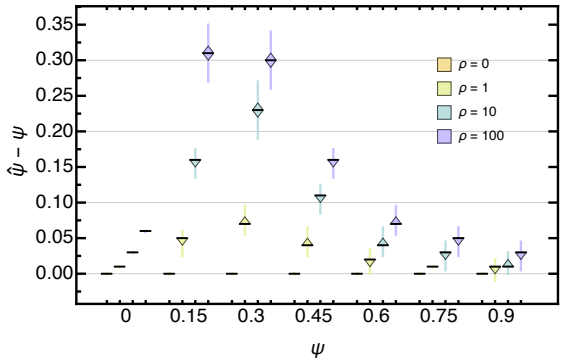

Figure 7 – Boxplot of the deviation of the maximum likelihood estimate from the true ψ for 10,000 data sets assuming linkage equilibrium with $k = 100$ and $s = 10{,}000$ when not accounting for population growth. Boxes represent the interquartile range (i.e., the 50% C.I.), whiskers extend to the highest/lowest data point within the box ± 1.5 times the interquartile range.

Application to sardine data

Finally, we applied our joint inference framework to a derived SFS for the control region of mtDNA in Japanese sardine (*Sardinops melanostictus*; File E1). Niwa *et al.* (2016) recently analyzed this data to test whether the observed excess in singletons was more likely caused by a recent population expansion or by sweepstake reproductive events and found that the latter is the more likely explanation. However, there is of course no *a priori* reason to believe that both reproductive skew and population growth could not have acted simultaneously.

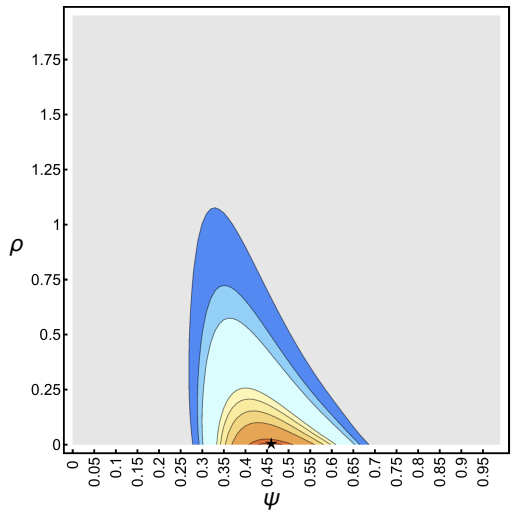

Figure 8 – Likelihood surface (eq. 20) of the unfolded SFS (given the ML rooted tree) of the sardine mtDNA sequences with $k = 106$ and $s = 78$. Contours show the $0.95, 0.9675, 0.975, 0.99,\ 0.99225, 0.9945, 0.99675,\ 0.999, 0.99945$ and 0.9999 quantiles. Likelihoods below the 0.95 quantile are uniformly colored in gray. The black star shows the maximum likelihood estimates $\hat{\psi} = 0.28$ and $\hat{\rho} = 0$.

When estimated jointly, the maximum likelihood estimate is $(\hat{\psi}, \hat{\rho}) = (0.28, 0)$ (Fig. 8), which implies considerable reproductive skew, but no (exponential) population growth. While our analysis confirms their results at first glance there are two points that warrant caution with this interpretation. First, as indicated by the contour lines in the plot there is some probability that the Japanese sardine population underwent a recent population expansion, though if it did, it only grew at a very low rate. Second, mtDNA does not recombine implying that all sites are in complete linkage. Thus, given our simulation results assuming complete linkage (see above; Fig. 6), ρ could be (slightly) underestimated.

Concluding remarks

This study marks the first multiple-merger coalescent with time-varying population sizes derived from a discrete time random mating model, and provides the first in-depth analyses of the joint inference of coalescent and demographic parameters. Since the Kingman coalescent represents a special case of the general class of multiple-merger coalescents (Irwin *et al.* 2016; Spence *et al.* 2016), it is interesting and encouraging to see that our analytical results – i.e., the time-change function (eq. 47) and the first expected coalescence times (eq. 51) – are generalizations of results derived for the Kingman coalescent (Griffiths and Tavaré 1998; Polanski and Kimmel 2003). In fact, when growth rates are measured within the corresponding coalescent framework (e.g., as $\tilde{\rho} = \rho\gamma$ for the psi-coalescent) these formulas should extend to other, more general multiple-merger coalescents. This also holds true for the challenges arising when calculating the expected normalized SFS (eq. 19) which is central to estimating coalescent parameters and growth rates: Because of catastrophic cancellation errors – mainly due to numerical representations of the exponential integral $\mathrm{Ei}(x)$ – computations have to be carried out using multi-precision libraries (Spence *et al.* 2016).

While both ψ and ρ can generally be estimated precisely, accurate estimation of the latter requires sufficient information (i.e., a large number of segregating sites) especially when offspring distributions are heavily skewed (i.e., if ψ is large). However, since strong recurrent sweepstake reproductive events – analogous to recurrent selective sweeps – constantly erase

genetic variation (i.e., reduce the number of segregating sites), there might be little power to accurately infer ρ in natural populations in these cases. In accordance with previous findings derived for the Kingman coalescent (Terhorst and Song 2015), increasing sample size does not improve the accuracy of demographic inference (i.e., estimating ρ) for a fixed number of segregating sites s. However, unlike in the Kingman coalescent where s increases logarithmically with sample size, genetic variation in ψ increases linearly for large ψ which could offset – or at least the hamper – this effect.

What is more crucial though for inferring growth rates is the effect of linkage. Under linkage equilibrium, maximum likelihood estimates were asymptotically unbiased; linkage disequilibrium however introduced a systematic bias that resulted in an overestimation of ρ for small ψ. As ψ increased, growth rate estimates gradually decreased resulting in a systematic underestimation. Similar observations have been made when trying to estimate the duration and the rate of exponential growth under the Kingman coalescent (Bhaskar *et al.* 2015). Likewise, Schrider *et al.* (2016) recently found that linked positive selection can severely bias demographic estimates. While their analyses assumed a Kingman framework, positive selection and recurrent selective sweeps typically result in multiple merger events (Durrett and Schweinsberg 2004, 2005; Neher and Hallatschek 2013). Thus, if neutral regions are tightly linked they will – at least partially – share the genealogical relationship with the selected region and skew inference. Despite the fact that our model here considers organisms with skewed offspring distributions under neutrality owing to the specifics of their reproductive biology, increasing ψ is tantamount to increasing the strength of positive selection under a non-neutral model, which is thus relevant to a very broad class of organisms indeed.

Furthermore, our analyses only consider the two extreme cases – assuming that loci are either completely linked or completely unlinked. While these are adequate for some biological systems, both extremes clearly mark the endpoint of a continuum of recombination rates. However, due to the lack of explicit coalescent simulators that allow for multiple-mergers, non-constant population sizes, and varying recombination rates, the role of recombination on the joint estimation of coalescent and demographic parameters remains open for future research. Importantly, linkage (dis)equilibrium depends not only on the rate of recombination but also on the reproduction parameters, and is thus expected to be highly variable – potentially being elevated despite frequent recombination, or largely absent despite infrequent recombination (Eldon and Wakeley 2008; Birkner *et al.* 2012). This, combined with our analyses, emphasizes the importance of accounting for demography and illuminates the serious biases that can arise in the inferred coalescent model if ignored. Such bias can have broad implications on inferred patterns of genetic variation (Eldon and Wakeley 2006; Tellier and Lemaire 2014; Niwa *et al.* 2016), including misguiding conservation efforts (Montano 2016), and obscuring the extent of reproductive skew.

Finally, most of the current analytical and computational tools have been derived and developed under the Kingman coalescent. In order to achieve the overall aim of generalizing the Kingman coalescent model (Wakeley 2013), these tools, though often computationally challenging, need to be extended. Great efforts have recently been undertaken towards developing a statistical inference framework allowing for model selection (Birkner and Blath 2008; Eldon 2011; Birkner *et al.* 2011, 2013, 2012; Steinrücken *et al.* 2013; Eldon *et al.* 2015; Spence *et al.* 2016). By setting a discrete-time random mating model and deriving the ancestral process, along with providing the analytical tools necessary to enable the joint inference of offspring

distribution and demography, this study makes an important contribution towards this goal.

Acknowledgements

This project was funded by grants from the Swiss National Science Foundation (FNS) and a European Research Council (ERC) Starting Grant to JDJ. We thank Kristen Irwin, Adamandia Kapopoulou, Martin Möehle, Sylvain Mousset and Meike Wittmann for helpful discussion and comments on earlier versions of this manuscript, and Hiro-Sato Niwa for providing the sardine data.

Literature Cited

Achaz, G., 2009 Frequency spectrum neutrality tests: One for all and all for one. Genetics **183**: 249–258.

Árnason, E. and K. Halldórsdóttir, 2015 Nucleotide variation and balancing selection at the *Ckma* gene in atlantic cod: analysis with multiple merger coalescent models. PeerJ **3**: e786.

Bhaskar, A., A. G. Clark, and Y. S. Song, 2014 Distortion of genealogical properties when the sample is very large. Proc. Natl. Acad. Sci. USA **111**: 2385–2390.

Bhaskar, A., Y. R. Wang, and Y. S. Song, 2015 Efficient inference of population size histories and locus-specific mutation rates from large-sample genomic variation data. Genome Res. **25**: 268–279.

Birkner, M. and J. Blath, 2008 Computing likelihoods for coalescents with multiple collisions in the infinitely many sites model. J. Math. Biol. **57**: 435–465.

Birkner, M., J. Blath, and B. Eldon, 2012 An ancestral recombination graph for diploid populations with skewed offspring distribution. Genetics **193**: 255–290.

Birkner, M., J. Blath, and B. Eldon, 2013 Statistical properties of the site-frequency spectrum associated with λ-coalescents. Genetics **195**: 1037–1053.

Birkner, M., J. Blath, and M. Steinrücken, 2011 Importance sampling for lambda-coalescents in the infinitely many sites model. Theor. Pop. Biol. **79**: 155 – 173.

Bolthausen, E. and A.-S. Sznitman, 1998 On ruelle's probability cascades and an abstract cavity method. Commun. Math. Phys. **197**: 247–276.

Cannings, C., 1974 The latent roots of certain markov chains arising in genetics: A new approach, i. haploid models. Adv. Appl. Probab. **6**: 260–290.

Depaulis, F., M. Veuille, *et al.*, 1998 Neutrality tests based on the distribution of haplotypes under an infinite-site model. Mol. Biol. Evol. **15**: 1788–1790.

Donnelly, P. and T. G. Kurtz, 1999 Particle representations for measure-valued population models. Ann. Probab. **27**: 166–205.

Durrett, R. and J. Schweinsberg, 2004 Approximating selective sweeps. Theor. Pop. Biol. **66**: 129–138.

Durrett, R. and J. Schweinsberg, 2005 A coalescent model for the effect of advantageous mutations on the genealogy of a population. Stoch. Proc. Appl. **115**: 1628 – 1657.

Eldon, B., 2011 Estimation of parameters in large offspring number models and ratios of coalescence times. Theor. Pop. Biol. **80**: 16 – 28.

Eldon, B., M. Birkner, J. Blath, and F. Freund, 2015 Can the site-frequency spectrum distinguish exponential population growth from multiple-merger coalescents? Genetics **199**: 841–856.

Eldon, B. and J. Wakeley, 2006 Coalescent processes when the distribution of offspring number among individuals is highly skewed. Genetics **172**: 2621–2633.

Eldon, B. and J. Wakeley, 2008 Linkage disequilibrium under skewed offspring distribution among individuals in a population. Genetics **178**: 1517–1532.

Fay, J. C. and C.-I. Wu, 2000 Hitchhiking under positive darwinian selection. Genetics **155**: 1405–1413.

Ferretti, L., M. Pérez-Enciso, and S. E. Ramos-Onsins, 2010 Optimal neutrality tests based on frequency spectrum. Genetics .

Fisher, R., 1930 *The genetical theory of natural selection*. Clarendon Press, Oxford.

Grant, W. S., E. Árnason, and B. Eldon, 2016 New dna coalescent models and old population genetics software. ICES J. Mar. Sci. **73**: 2178–2180.

Griffiths, R. and S. Tavaré, 1998 The age of a mutation in a general coalescent tree. Commun. Statist. Stoch. Model. **14**: 273–295.

Hedgecock, D., 1994 Does variance in reproductive success limit effective population sizes of marine organisms? In *Genetics and evolution of aquatic organisms*, edited by A. Beaumont 1222–1344, Chapman and Hall, London, UK.

Hedgecock, D. and A. I. Pudovkin, 2011 Sweepstakes reproductive success in highly fecund marine fish and shellfish: a review and commentary. Bull. Mar. Sci. **87**: 971–1002.

Huillet, T. and M. Möhle, 2013 On the extended moran model and its relation to coalescents with multiple collisions. Theor. Pop. Biol. **87**: 5 – 14.

Irwin, K. K., S. Laurent, S. Matuszewski, S. Vuilleumier, L. Ormond, H. Shim, C. Bank, and J. D. Jensen, 2016 On the importance of skewed offspring distributions and background selection in virus population genetics. Heredity **117**: 393–399.

Kingman, J., 1982a The coalescent. Stoch. Proc. Appl. **13**: 235 – 248.

Kingman, J. F. C., 1982b On the Genealogy of Large Populations. J. Appl. Probab. **19**: 27–43.

Kingman, J. F. C., 1982c Exchangeability and the Evolution of Large Populations. In *Exchangeability in Probability and Statistics*, edited by G. Koch and F. Spizzichino 97–112, North-Holland, Amsterdam.

Kingman, J. F. C., 2000 Origins of the coalescent: 1974-1982. Genetics **156**: 1461–1463.

Möhle, M., 1998 Robustness results for the coalescent. J. Appl. Probab. **35**: 438–447.

Möhle, M., 1999 Weak convergence to the coalescent in neutral population models. J. Appl. Probab. **36**: 446–460.

Möhle, M., 2002 The coalescent in population models with time-inhomogeneous environment. Stoch. Proc. Appl. **97**: 199–227.

Möhle, M. and S. Sagitov, 2001 A classification of coalescent processes for haploid exchangeable population models. Ann. Probab. **29**: 1547–1562.

Montano, V., 2016 Coalescent inferences in conservation genetics: should the exception become the rule? Biol. Lett. **12**.

Moran, P. A. P., 1958 Random processes in genetics. Proc. Cambridge Philos. Soc. **54**: 60.

Moran, P. A. P., 1962 *The statistical processes of evolutionary theory.*. Clarendon Press; Oxford University Press.

Neher, R. A. and O. Hallatschek, 2013 Genealogies of rapidly adapting populations. Proc. Natl. Acad. Sci. USA **110**: 437–442.

Neuhauser, C. and S. M. Krone, 1997 The genealogy of samples in models with selection.

Genetics **145**: 519–534.

Niwa, H.-S., K. Nashida, and T. Yanagimoto, 2016 Reproductive skew in japanese sardine inferred from dna sequences. ICES J. Mar. Sci. **73**: 2181–2189.

Nordborg, M., 1997 Structured coalescent processes on different time scales. Genetics **146**: 1501–1514.

Pitman, J., 1999 Coalescents with multiple collisions. Ann. Probab. **27**: 1870–1902.

Polanski, A., A. Bobrowski, and M. Kimmel, 2003 A note on distributions of times to coalescence, under time-dependent population size. Theor. Pop. Biol. **63**: 33 – 40.

Polanski, A. and M. Kimmel, 2003 New explicit expressions for relative frequencies of single-nucleotide polymorphisms with application to statistical inference on population growth. Genetics **165**: 427–436.

Ramos-Onsins, S. E. and J. Rozas, 2000 Statistical properties of new neutrality tests against population growth. Mol. Biol. Evol. **19**: 2092.

Sagitov, S., 1999 The general coalescent with asynchronous mergers of ancestral lines. J. Appl. Probab. **36**: 1116–1125.

Sawyer, S. A. and D. L. Hartl, 1992 Population genetics of polymorphism and divergence. Genetics **132**: 1161–1176.

Schrider, D. R., A. G. Shanku, and A. D. Kern, 2016 Effects of linked selective sweeps on demographic inference and model selection. Genetics .

Schweinsberg, J., 2000 Coalescents with simultaneous multiple collisions. Electron. J. Probab. **5**: 1–50.

Spence, J. P., J. A. Kamm, and Y. S. Song, 2016 The site frequency spectrum for general coalescents. Genetics **202**: 1549–1561.

Steinrücken, M., M. Birkner, and J. Blath, 2013 Analysis of dna sequence variation within marine species using beta-coalescents. Theor. Pop. Biol. **87**: 15–24.

Tajima, F., 1989 Statistical method for testing the neutral mutation hypothesis by dna polymorphism. Genetics **123**: 585–595.

Tellier, A. and C. Lemaire, 2014 Coalescence 2.0: a multiple branching of recent theoretical developments and their applications. Mol. Ecol. **23**: 2637–2652.

Terhorst, J. and Y. S. Song, 2015 Fundamental limits on the accuracy of demographic inference based on the sample frequency spectrum. Proc. Natl. Acad. Sci. USA **112**: 7677–7682.

Wakeley, J., 2009 *Coalescent Theory: An Introduction*. Coalescent theory: An introduction., Roberts & Company Publishers, Greenwood Village.

Wakeley, J., 2013 Coalescent theory has many new branches. Theor. Pop. Biol. **87**: 1 – 4.

Wakeley, J. and T. Takahashi, 2003 Gene genealogies when the sample size exceeds the effective size of the population. Mol. Biol. Evol. **20**: 208–213.

Watterson, G., 1975 On the number of segregating sites in genetical models without recombination. Theor. Pop. Biol. **7**: 256 – 276.

Wilkinson-Herbots, H. M., 1998 Genealogy and subpopulation differentiation under various models of population structure. J. Math. Biol. **37**: 535–585.

Wright, S., 1931 Evolution in mendelian populations. Genetics **16**: 97–159.

Supporting Information

Detailed derivation of results

In this Supporting Information we will give a detailed derivation of the results which in the main text have been abbreviated to keep formulas concise. First, Equation 29 showing that $\Delta_N^{(n)}$ is bounded for all $n \in \mathbb{N}$ is derived as

$$\begin{aligned}
\lim_{N\to\infty} \Delta_N^{(n)} &= \lim_{N\to\infty} (N_n - N_{n-1}) \\
&\leq \lim_{N\to\infty} \left(N \left(1 - \rho \frac{\psi^2}{N^\gamma}\right)^n - N \left(1 - \rho \frac{\psi^2}{N^\gamma}\right)^{n-1} \right) = \lim_{N\to\infty} N \exp\left[-\rho \frac{\psi^2}{N^\gamma} n\right] \left(\exp\left[-\rho \frac{\psi^2}{N^\gamma}\right] - 1\right) \\
&\leq \lim_{N\to\infty} N \left(\exp\left[-\rho \frac{\psi^2}{N^\gamma}\right] - 1\right) = \lim_{N\to\infty} N\rho \frac{\psi^2}{N^\gamma} \leq \lim_{N\to\infty} N\rho \frac{\psi^2}{N^\gamma} \\
&= 0.
\end{aligned} \tag{S1}$$

Second, the step function $F_N(s)$ (eq. 37) for the ancestral process is derived as

$$\begin{aligned}
F_N(s) &= \sum_{n=1}^{s} c_N(n) \sim \sum_{n=1}^{s} \frac{\psi^2}{N_n^\gamma} = \frac{\psi^2}{N^\gamma} \sum_{n=1}^{s} \left(\frac{1}{\left(1 - \rho \frac{\psi^2}{N^\gamma}\right)^\gamma} \right)^n \\
&= \frac{\psi^2}{N^\gamma} \left(1 - \rho \frac{\psi^2}{N^\gamma}\right)^{-\gamma} \frac{\left(\left(1 - \rho \frac{\psi^2}{N^\gamma}\right)^{-\gamma s} - 1\right)}{\left(1 - \rho \frac{\psi^2}{N^\gamma}\right)^{-\gamma} - 1}.
\end{aligned} \tag{S2}$$

Then, by solving for s the time-change function $\mathcal{G}_N^{-1}(t)$ (eq. 38) for the ancestral process can be derived as

$$
\begin{aligned}
\mathcal{G}_N^{-1}(t) &= \inf\left\{s > 0 : f_N\lfloor s\rfloor) > t\right\} - 1 \\
&\sim \inf\left\{ s > 0 : [s] > \frac{\log\left[1 - \frac{tN^\gamma}{\psi^2}\frac{\left(1-\frac{\rho\psi^2}{N^\gamma}\right)^{-\gamma}-1}{\left(1-\frac{\rho\psi^2}{N^\gamma}\right)^{-\gamma}}\right]}{\log\left[\left(1-\frac{\rho\psi^2}{N^\gamma}\right)^{-\gamma}\right]} \right\} - 1 \\
&= \left[\frac{\log\left[1 - \frac{tN^\gamma}{\psi^2}\frac{\left(1-\frac{\rho\psi^2}{N^\gamma}\right)^{-\gamma}-1}{\left(1-\frac{\rho\psi^2}{N^\gamma}\right)^{-\gamma}}\right]}{\log\left[\left(1-\frac{\rho\psi^2}{N^\gamma}\right)^{-\gamma}\right]}\right] \\
&= \left[\frac{\log\left[1 - \frac{tN^\gamma}{\psi^2}\left(1-\left(1-\frac{\rho\psi^2}{N^\gamma}\right)^{\gamma}\right)\right]}{-\gamma\log\left[\left(1-\frac{\rho\psi^2}{N^\gamma}\right)\right]}\right] \\
&\sim \left[\frac{\log\left[1 - \frac{tN^\gamma}{\psi^2}\left(1-\exp\left[\frac{\rho\gamma\psi^2}{N^\gamma}\right]\right)\right]}{-\gamma\log\left[\left(1-\frac{\rho\psi^2}{N^\gamma}\right)\right]}\right] \\
&\sim \left[\frac{\log\left[1 + \frac{tN^\gamma}{\psi^2}\left(1-(1-\frac{\rho\gamma\psi^2}{N^\gamma}\right)\right]}{-\gamma\log\left[\left(1-\frac{\rho\psi^2}{N^\gamma}\right)\right]}\right] \\
&= \left[\frac{\log\left[1+\rho\gamma t\right]}{-\gamma\log\left[\left(1-\frac{\rho\psi^2}{N^\gamma}\right)\right]}\right] \\
&\sim \left[\frac{\log\left[1+\rho\gamma t\right]}{\rho\gamma}\frac{N^\gamma}{\psi^2}\right], \qquad \text{(S3)}
\end{aligned}
$$

where we have used that $\log\left[1-\frac{\rho\psi^2}{N^\gamma}\right] \sim -\frac{\rho\psi^2}{N^\gamma}$ for sufficiently large N. Finally, the derivation showing that the ancestral process of the Moran model converges to a (time-inhomogeneous) psi-coalescent (eq. 45) is given by

$$
\begin{aligned}
\lim_{N\to\infty}\sum_{n=1}^{\mathcal{G}_N^{-1}(t)}\Phi_1^{(N)}(r;b) &= \lim_{N\to\infty}\sum_{n=1}^{\mathcal{G}_N^{-1}(t)} G_{b,b} \\
&= \lim_{N\to\infty}\sum_{n=1}^{\mathcal{G}_N^{-1}(t)}\binom{b}{b}\sum_{u=2}^{N}P_U(u)\frac{(u)_b(N_n-u)_{b-b}}{(N_n)_b} \\
&= \lim_{N\to\infty}\sum_{n=1}^{\mathcal{G}_N^{-1}(t)}\frac{(1-N_n^{\gamma})(2)_b+N_n^{-\gamma}(N_n\psi)_b}{(N_n)_b} \\
&= \lim_{N\to\infty}\sum_{n=1}^{\mathcal{G}_N^{-1}(t)}\frac{N_n^{-\gamma}(N_n\psi)^b)}{N_n^b} \\
&= \lim_{N\to\infty}\sum_{n=1}^{\mathcal{G}_N^{-1}(t)}\frac{\psi^b}{N_n^{\gamma}} \\
&= \lim_{N\to\infty}\frac{\psi^b}{N^{\gamma}}\sum_{n=1}^{\mathcal{G}_N^{-1}(t)}\left(\left(1-\frac{\rho\psi^2}{N^{\gamma}}\right)^{-\gamma}\right)^n \\
&= \lim_{N\to\infty}\frac{\psi^b}{N^{\gamma}}\frac{\left(1-\frac{\rho\psi^2}{N^{\gamma}}\right)^{-\gamma\mathcal{G}_N^{-1}(t)}-1}{1-\left(1-\frac{\rho\psi^2}{N^{\gamma}}\right)^{\gamma}} \\
&= \lim_{N\to\infty}\frac{\psi^b}{N^{\gamma}}\frac{\left(1-\frac{\rho\psi^2}{N^{\gamma}}\right)^{-\gamma\mathcal{G}_N^{-1}(t)}-1}{\frac{\rho\gamma\psi^2}{N^{\gamma}}} \\
&= \lim_{N\to\infty}\psi^{b-2}\frac{\left(1-\frac{\rho\psi^2}{N^{\gamma}}\right)^{-\gamma\mathcal{G}_N^{-1}(t)}-1}{\rho\gamma} \\
&= \lim_{N\to\infty}\psi^{b-2}\frac{\exp\left(\frac{\rho\gamma\psi^2}{N^{\gamma}}\mathcal{G}_N^{-1}(t)\right)-1}{\rho\gamma} \\
&= \lim_{N\to\infty}\psi^{b-2}\frac{\exp\left(\frac{\rho\gamma\psi^2}{N^{\gamma}}\frac{N^{\gamma}}{\rho\gamma\psi^2}\log\left(1+\rho\gamma t\right)\right)-1}{\rho\gamma} \\
&= \psi^{b-2}\frac{\exp\left(\log\left(1+\rho\gamma t\right)\right)-1}{\rho\gamma} \\
&= \psi^{b-2}t. \qquad \text{(S4)}
\end{aligned}
$$

Extended Moran model simulations

In this Supporting Information we compare a set of tree statistics obtained from the ancestral process $\left(\mathcal{A}^{\psi,\rho}_{\lfloor t/c_N(0)\rfloor,k}\right)_{t\in\mathbb{R}_{\geq 0}}$ to those obtained from the underlying extended Moran model $\left(\Pi^{\psi,\rho}_{t,k}\right)_{t\geq 0}$ for different values of $\rho \in \mathbb{R}_{\geq 0}$, $\psi \in (0,1)$ with $\gamma = 1.5$ (i.e., in the regime where sweepstake reproductive events dominate and the corresponding ancestral process is a psi-coalescent). Coalescent simulations follow the algorithm outlined in the main text.

The rational behind the extended Moran model simulations is outlined in Figure SI B_1.

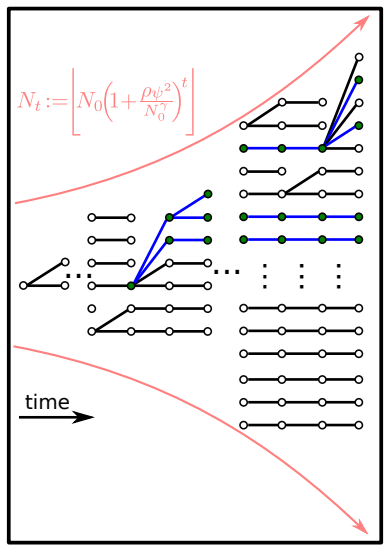

Figure SI B_1 – Illustration of genealogical relation in the extended Moran model with exponential growth. Circles denote individuals; time increases from left to right with the rightmost individuals constituting the present population. Green circles denote sampled individuals. Connecting lines between circles show the genealogical relationship between individuals; the genealogy of the sampled individuals is highlighted with blue. Pink lines illustrate that population size increases (forward in time) exponentially with rate ρ.

Starting from the present (i.e., $n = 0$) and choosing k random samples, these were followed successively backwards in time thus creating a coalescent tree from the leaves to the root. At each time step the population size was adjusted (shrinking backward in time; eq. 6) and the type of the reproductive event (eq. 5) along with the corresponding number of offspring $U_N(n)$ (eq. 4) were (randomly) determined . Finally, to determine whether a merger event has occurred – i.e., if one of the k active lineages had found its parent in the previous time step – $U_N(n)$ Bernoulli random numbers $\mathbf{Y} = (Y_i)_{i=1,2,\ldots,U_N(n)}$ with parameter

$$\frac{k - \sum_{j=1}^{i} \mathbb{1}_{y_i}}{N_n - i}$$

are drawn. Note that the denominator changes with the number of individuals that have not been assigned to a parent in the previous time step, while the nominator changes with the number of active lineages that have found their parent in the previous time step. In particular, $\sum_{j=1}^{U_N(n)} \mathbb{1}_{y_i}$ denotes the total number of active lineages that found their parent in the previous time step. Thus, if $\sum_{j=1}^{U_N(n)} \mathbb{1}_{y_i} > 1$ a coalescent event has occurred, and the number of active lineages in the next time step becomes $k - \sum_{j=1}^{U_N(n)} \mathbb{1}_{y_i}$. This process is repeated until $k = 0$, that is until all samples have found their most recent common ancestor.

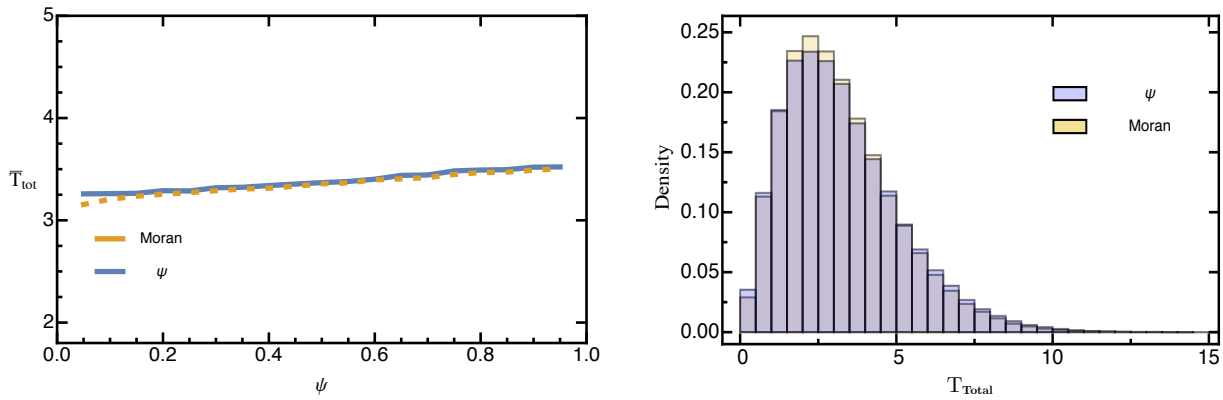

Figure SI B_2 – The left figure shows a comparison of $\overline{T}_{tot}$ based on simulated trajectories of $\left(\mathcal{A}^{\psi,\rho}_{\lfloor t/c_N(0)\rfloor,4}\right)_{t\geq 0}$ and $\left(\Pi^{\psi,\rho}_{t,4}\right)_{t\geq 0}$ for $\rho = 0.1$, $\gamma = 1.5$, and different values of ψ. The right figure depicts the corresponding empirical distributions of $T^{(i)}_{tot}$ for $\psi = 0.1$. Averages were taken over 100,000 replicates. Other parameters: $N = 100{,}000$.

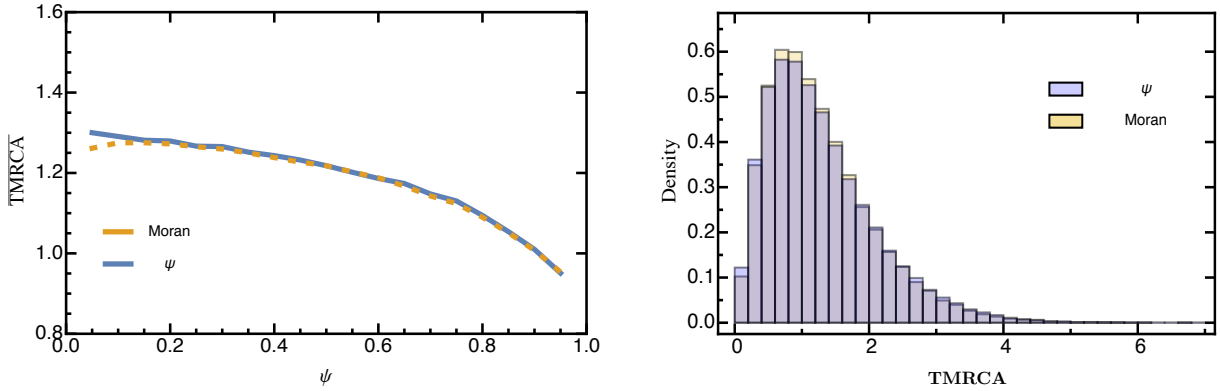

Figure SI B_3 – The left figure shows a comparison of $\overline{T}_{MRCA}$ based on simulated trajectories of $\left(\mathcal{A}^{\psi,\rho}_{\lfloor t/c_N(0)\rfloor,4}\right)_{t\geq 0}$ and $\left(\Pi^{\psi,\rho}_{t,4}\right)_{t\geq 0}$ for $\rho = 0.1$, $\gamma = 1.5$, and different values of ψ. The right figure depicts the corresponding empirical distributions of $T^{(i)}_{MRCA}$ for $\psi = 0.1$. Averages were taken over 100,000 replicates. Other parameters: $N = 100{,}000$.

For both processes the total tree length T_{total}, the time to the most recent common ancestor $TMRCA$, and the ratio of the sum of the length of all branches with i descendants over the total tree length T_i/T_{total} were recorded. While the number of samples was limited to $k = 4$ for computational reasons, the match between coalescent and Moran model simulations was almost perfect across the entire range of ψ (Fig. SI B_2-SI B_4). Only for very small ψwere slight deviations observed.

These results highlight two important points. First and reassuringly these results indicate that for large initial population sizes N the extended Moran model can be described by its limiting ancestral process (i.e., the psi-coalescent). Second, simulating the population dynamics forward in time is computationally intensive and prohibitively slow. However, since the ancestral process accurately captures the forward-in-time population dynamics, the coalescent process can be used to simulate the process quickly and efficiently over a large parameter space and for large sample sizes.

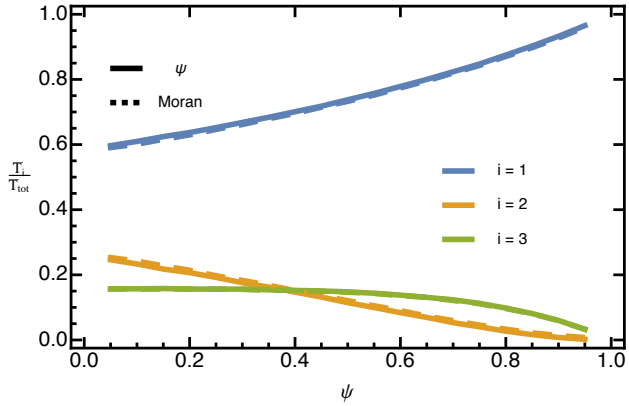

Figure SI B_4 – A comparison of $(T_i/\overline{T}_{tot})_{i=1,2,3}$ based on simulated trajectories of $\left(\mathcal{A}^{\psi,\rho}_{\lfloor t/c_N(0)\rfloor,4}\right)_{t\geq 0}$ and $\left(\Pi^{\psi,\rho}_{t,4}\right)_{t\geq 0}$ for $\rho = 0.1$, $\gamma = 1.5$, and different values of ψ. Averages were taken over 100,000 replicates. Other parameters: $N = 100{,}000$.

Supporting Figures

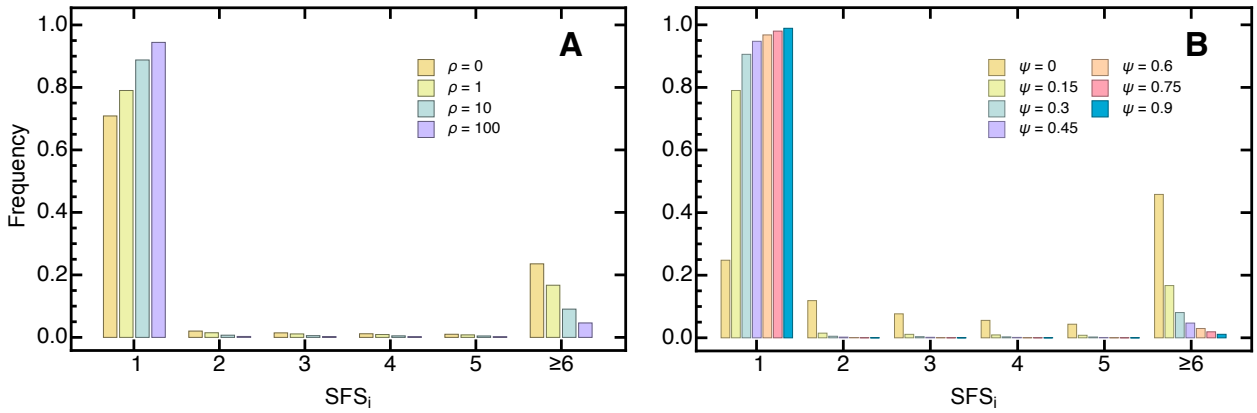

Figure SI C_1 – The expected normalized (lumped) SFS for the psi-coalescent for an exponentially growing population (eq. 24) with sample size $k = 100$ (**A**) for different values of ψ and fixed $\rho = 1$ and (**B**) for different values of ρ and fixed $\psi = 0.15$. The sixth entry in the SFS contains the aggregate of the higher frequency classes.

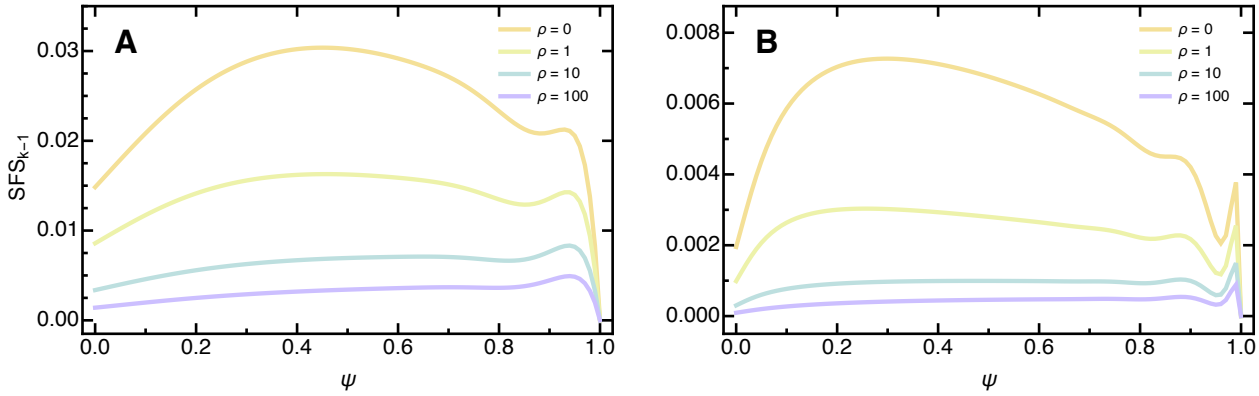

Figure SI C_2 – The expected normalized last entry of the SFS (eq. 24) as a function of ψ for various values of ρ with sample size $k = 100$ (**A**) for $k = 20$ and (**B**) for $k = 100$.

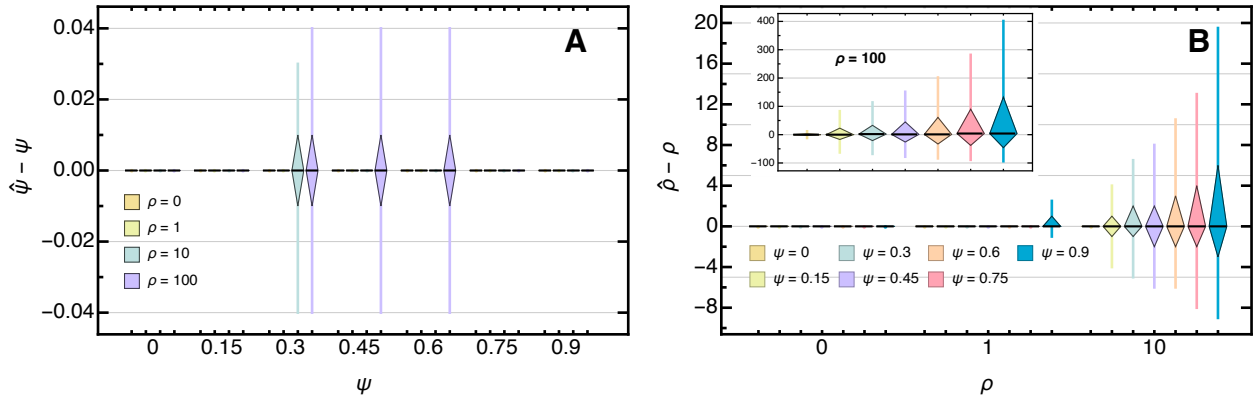

Figure SI C_3 – Boxplot of the deviation of the maximum likelihood estimate from the true (**A**) ψ and (**B**) ρ for 10,000 data sets assuming linkage equilibrium with $k = 200$ and $s = 10{,}000$. Boxes represent the interquartile range (i.e., the 50% C.I.) and whiskers extend to the highest/lowest data point within the box ± 1.5 times the interquartile range.

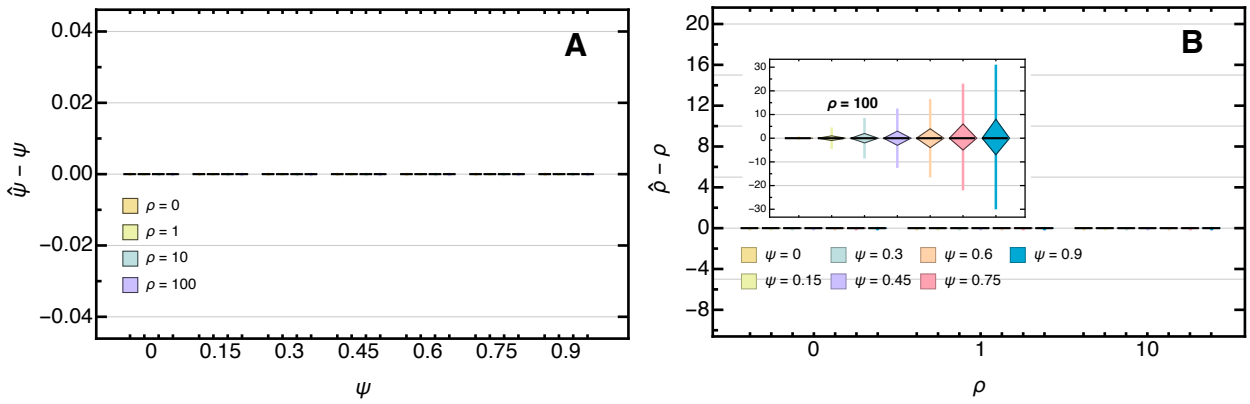

Figure SI C_4 – Boxplot of the deviation of the maximum likelihood estimate from the true (**A**) ψ and (**B**) ρ for 10,000 data sets assuming linkage equilibrium with $k = 200$ and $s = 1{,}000{,}000$. Boxes represent the interquartile range (i.e., the 50% C.I.) and whiskers extend to the highest/lowest data point within the box ± 1.5 times the interquartile range. As the number of segregating sites s increases the variance of the estimator decreases and approaches its true underlying value.

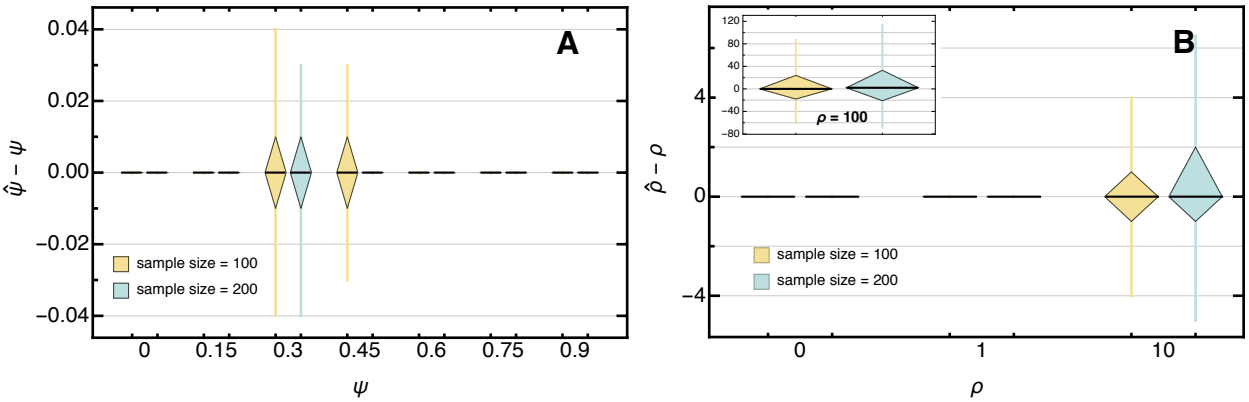

Figure SI C_5 – Boxplot of the deviation of the maximum likelihood estimate from the true (**A**) ψ and fixed $\rho = 10$, and (**B**) ρ and fixed $\psi = 0.3$ for 10,000 data sets assuming linkage equilibrium for different sample sizes k with $s = 10{,}000$. Boxes represent the interquartile range (i.e., the 50% C.I.) and whiskers extend to the highest/lowest data point within the box ± 1.5 times the interquartile range.

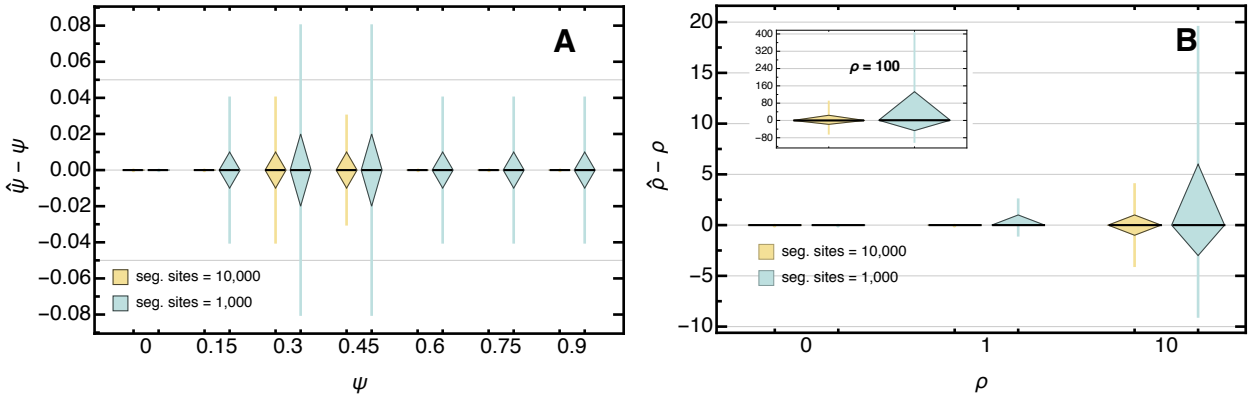

Figure SI C_6 – Boxplot of the deviation of the maximum likelihood estimate from the true (**A**) ψ and fixed $\rho = 10$, and (**B**) ρ and fixed $\psi = 0.3$ for 10,000 data sets assuming linkage equilibrium for different segregating sites s with $k = 100$. Boxes represent the interquartile range (i.e., the 50% C.I.) and whiskers extend to the highest/lowest data point within the box ± 1.5 times the interquartile range.

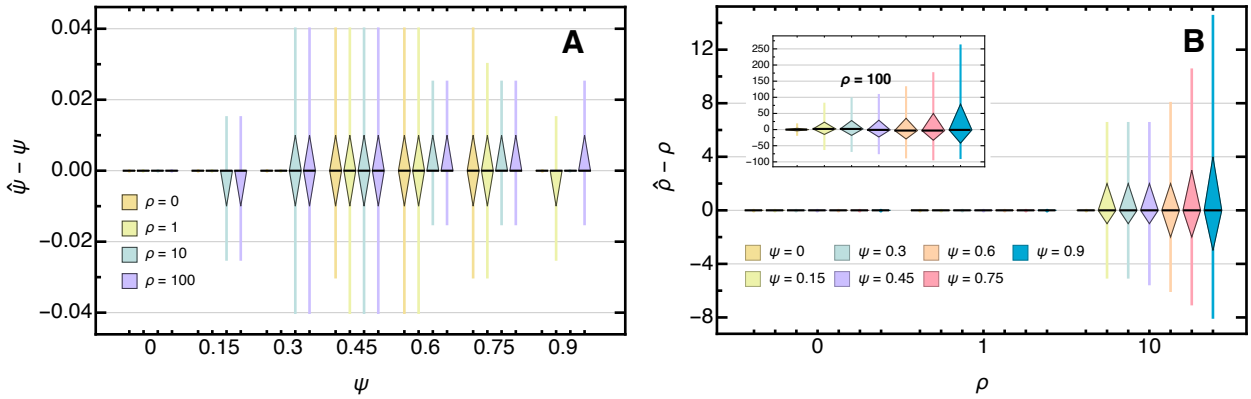

Figure SI C_7 – Boxplot of the deviation of the L1-distance estimate from the true (**A**) ψ and fixed $\rho = 10$, and (**B**) ρ and fixed $\psi = 0.3$ for 10,000 data sets assuming linkage equilibrium for different segregating sites s with $k = 100$. Boxes represent the interquartile range (i.e., the 50% C.I.) and whiskers extend to the highest/lowest data point within the box ± 1.5 times the interquartile range.

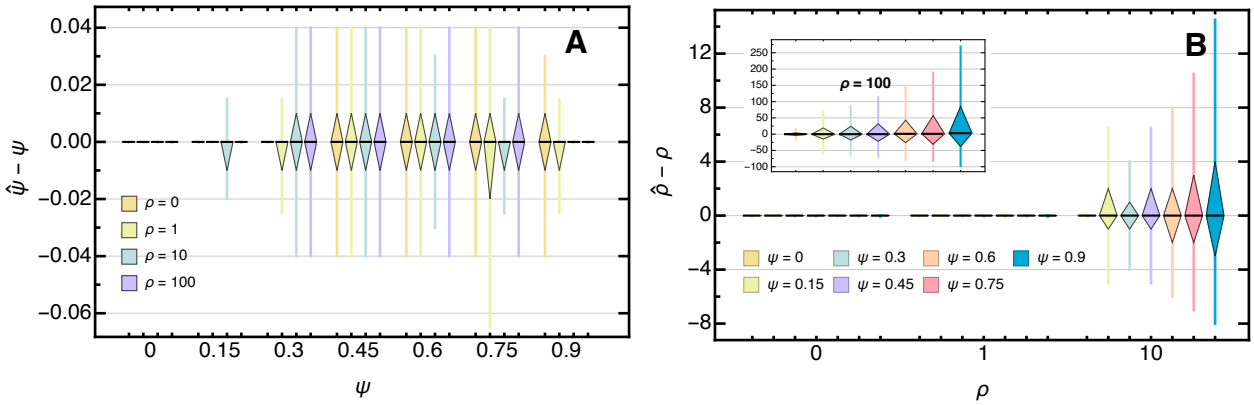

Figure SI C_8 – Boxplot of the deviation of the L2-distance estimate from the true (**A**) ψ and fixed $\rho = 10$, and (**B**) ρ and fixed $\psi = 0.3$ for 10,000 data sets assuming linkage equilibrium for different segregating sites s with $k = 100$. Boxes represent the interquartile range (i.e., the 50% C.I.) and whiskers extend to the highest/lowest data point within the box ± 1.5 times the interquartile range.

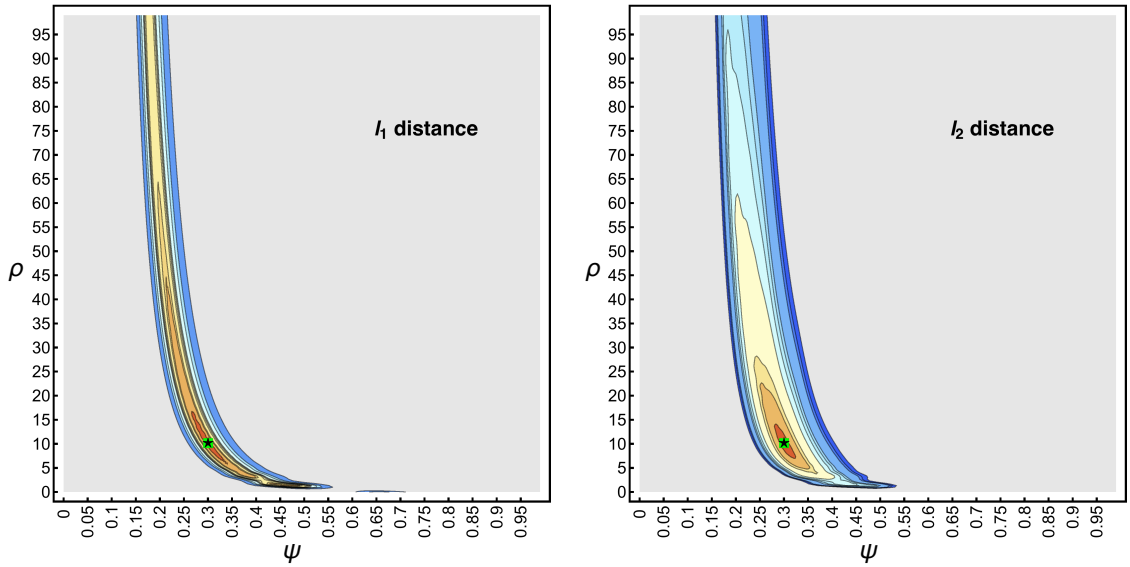

Figure SI C_9 – Surface of the l_1 (left) and the l_2 (right) distance of the idealized SFS with $k = 100$, $\psi = 0.3$, $\rho = 10$ and $s = 10{,}000$. Contours show the $0.95, 0.9675, 0.975, 0.99, 0.99225, 0.9945, 0.99675, 0.999, 0.99945$ and 0.9999 quantiles. Distances below the 0.95 quantile are uniformly colored in gray. The green square shows the true ψ and ρ. The black star shows the minimum distance estimates $\hat{\psi}$ and $\hat{\rho}$.

Supporting Tables

Table SI D_1 – Overview of the (marginal) accuracy $\hat{\psi}$. Each cell shows the mean difference (first row), the mean squared error (second row) and the median difference (third row) of $\hat{\psi}$ calculated over 10,000 data sets assuming linkage equilibrium. Colors within each sub-table range from light yellow to dark red and scale between the minimal and the maximal absolute value to aid interpretation.

$s = 100$

$k = 20$

Mean difference ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	1.707×10^{-2}	4.756×10^{-3}	1.886×10^{-2}	2.448×10^{-2}	2.346×10^{-2}	1.807×10^{-2}	6.254×10^{-3}
10	1.452×10^{-2}	8.58×10^{-4}	5.352×10^{-3}	2.549×10^{-3}	8.22×10^{-4}	8.557×10^{-3}	2.802×10^{-3}
1	1.414×10^{-2}	-2.042×10^{-3}	5.9×10^{-4}	-2.028×10^{-3}	-5.754×10^{-3}	6.19×10^{-4}	-1.172×10^{-3}
0	1.14×10^{-2}	-8.644×10^{-3}	-1.174×10^{-2}	-1.872×10^{-2}	-1.707×10^{-2}	-7.016×10^{-3}	-6.517×10^{-3}

Mean squared error ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	9.705×10^{-4}	3.823×10^{-3}	1.06×10^{-2}	1.066×10^{-2}	1.904×10^{-2}	1.346×10^{-2}	4.801×10^{-3}
10	7.076×10^{-4}	2.695×10^{-3}	6.746×10^{-3}	1.15×10^{-2}	1.417×10^{-2}	1.086×10^{-2}	4.115×10^{-3}
1	7.14×10^{-4}	2.149×10^{-3}	4.649×10^{-3}	7.723×10^{-3}	9.018×10^{-3}	7.653×10^{-3}	3.127×10^{-3}
0	4.85×10^{-4}	1.763×10^{-3}	2.94×10^{-3}	4.996×10^{-3}	6.336×10^{-3}	4.506×10^{-3}	2.323×10^{-3}

Median difference ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	0.	$-1. \times 10^{-2}$	$-1. \times 10^{-2}$	$-2. \times 10^{-2}$	$-2. \times 10^{-2}$	$-2. \times 10^{-2}$
10	0.	0.	0.	0.	$-1. \times 10^{-2}$	$-1. \times 10^{-2}$	$-1. \times 10^{-2}$
1	0.	$1. \times 10^{-2}$	0.	0.	0.	$-1. \times 10^{-2}$	0.
0	0.	$1. \times 10^{-2}$	$1. \times 10^{-2}$	$2. \times 10^{-2}$	$1. \times 10^{-2}$	0.	0.

$s = 1{,}000$

$k = 20$

Mean difference ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	5.261×10^{-3}	-9.4×10^{-5}	7.13×10^{-4}	-2.43×10^{-4}	1.11×10^{-4}	2.404×10^{-3}	1.836×10^{-3}
10	4.57×10^{-3}	3.6×10^{-5}	2.27×10^{-4}	-2.55×10^{-4}	-3.63×10^{-4}	4.3×10^{-4}	4.41×10^{-4}
1	3.508×10^{-3}	-2.262×10^{-3}	-3.959×10^{-3}	-3.377×10^{-3}	-3.705×10^{-3}	-1.63×10^{-3}	-3.83×10^{-4}
0	4.021×10^{-3}	9.5×10^{-3}	-2.86×10^{-4}	-4.27×10^{-4}	-3.62×10^{-4}	-6.72×10^{-4}	-1.512×10^{-3}

Mean squared error ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	9.595×10^{-6}	3.662×10^{-4}	8.847×10^{-4}	1.26×10^{-3}	1.436×10^{-3}	1.093×10^{-3}	6.51×10^{-4}
10	7.644×10^{-6}	2.439×10^{-4}	5.724×10^{-4}	9.246×10^{-4}	8.694×10^{-4}	6.61×10^{-4}	4.091×10^{-4}
1	5.004×10^{-6}	1.868×10^{-4}	3.882×10^{-4}	6.416×10^{-4}	5.8×10^{-4}	4.546×10^{-4}	2.666×10^{-4}
0	5.911×10^{-6}	1.505×10^{-4}	2.376×10^{-4}	3.875×10^{-4}	3.718×10^{-4}	2.866×10^{-4}	1.722×10^{-4}

Median difference ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	0.	0.	0.	0.	0.	0.
10	0.	0.	0.	0.	0.	0.	0.
1	0.	0.	0.	$1. \times 10^{-2}$	0.	0.	0.
0	0.	0.	0.	0.	0.	0.	0.

$s = 10{,}000$

$k = 20$

Mean difference ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	1.222×10^{-2}	4.4×10^{-5}	-3.2×10^{-5}	-1.44×10^{-4}	-2.5×10^{-5}	1.77×10^{-4}	6.6×10^{-5}
10	9.17×10^{-3}	-7.3×10^{-5}	-1.26×10^{-4}	-2.48×10^{-4}	-1.68×10^{-4}	-1.01×10^{-4}	-7.8×10^{-5}
1	5.37×10^{-3}	$8. \times 10^{-5}$	8.4×10^{-5}	-1.24×10^{-4}	-1.11×10^{-4}	-3.51×10^{-4}	-6.37×10^{-4}
0	6.75×10^{-3}	2.6×10^{-5}	9.8×10^{-5}	6.9×10^{-5}	4.2×10^{-5}	1.46×10^{-4}	-1.12×10^{-4}

Mean squared error ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	1.228×10^{-5}	4.378×10^{-5}	8.878×10^{-5}	1.483×10^{-3}	1.351×10^{-3}	1.094×10^{-3}	7.062×10^{-5}
10	9.17×10^{-6}	3.053×10^{-5}	6.426×10^{-5}	9.764×10^{-5}	8.942×10^{-5}	7.361×10^{-5}	4.734×10^{-5}
1	5.37×10^{-6}	1.45×10^{-5}	2.804×10^{-5}	4.706×10^{-5}	5.017×10^{-5}	4.349×10^{-5}	2.757×10^{-5}
0	6.75×10^{-6}	1.704×10^{-5}	2.83×10^{-5}	3.979×10^{-5}	3.816×10^{-5}	3.132×10^{-5}	1.644×10^{-5}

Median difference ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	0.	0.	0.	0.	0.	0.
10	0.	0.	0.	0.	0.	0.	0.
1	0.	0.	0.	0.	0.	0.	0.
0	0.	0.	0.	0.	0.	0.	0.

$k = 50$

Mean difference ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	6.262×10^{-3}	1.82×10^{-2}	5.508×10^{-2}	4.433×10^{-2}	3.244×10^{-2}	2.079×10^{-2}	8.99×10^{-4}
10	5.686×10^{-3}	4.676×10^{-3}	2.178×10^{-2}	1.952×10^{-2}	1.491×10^{-2}	1.251×10^{-2}	5.47×10^{-4}
1	5.767×10^{-3}	1.448×10^{-3}	2.754×10^{-3}	-5.62×10^{-4}	-9.61×10^{-4}	5.744×10^{-3}	3.88×10^{-4}
0	4.876×10^{-3}	-5.741×10^{-3}	-1.223×10^{-2}	-1.547×10^{-2}	-1.342×10^{-2}	-4.506×10^{-3}	-3.615×10^{-3}

Mean squared error ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	1.415×10^{-4}	6.202×10^{-3}	2.716×10^{-2}	3.329×10^{-2}	3.396×10^{-2}	2.6×10^{-2}	1.051×10^{-2}
10	1.186×10^{-4}	2.599×10^{-3}	1.511×10^{-2}	2.402×10^{-2}	2.796×10^{-2}	2.279×10^{-2}	9.653×10^{-3}
1	1.294×10^{-4}	1.56×10^{-3}	7.384×10^{-3}	1.378×10^{-2}	1.82×10^{-2}	1.644×10^{-2}	7.305×10^{-3}
0	9.468×10^{-5}	9.742×10^{-4}	3.831×10^{-3}	8.054×10^{-3}	1.165×10^{-2}	1.1×10^{-2}	5.905×10^{-3}

Median difference ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	0.	$-2. \times 10^{-2}$	$-1. \times 10^{-2}$	$-2. \times 10^{-2}$	$-2. \times 10^{-2}$	$-1. \times 10^{-2}$
10	0.	0.	0.	0.	$-1. \times 10^{-2}$	$-2. \times 10^{-2}$	$-1. \times 10^{-2}$
1	0.	0.	$1. \times 10^{-2}$	0.	0.	$-1. \times 10^{-2}$	$-1. \times 10^{-2}$
0	0.	$1. \times 10^{-2}$	$1. \times 10^{-2}$	$1. \times 10^{-2}$	0.	0.	0.

$s = 1{,}000$

$k = 50$

Mean difference ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	1.523×10^{-3}	7.39×10^{-4}	1.887×10^{-3}	-5.71×10^{-4}	1.219×10^{-3}	2.66×10^{-3}	2.1×10^{-3}
10	1.427×10^{-3}	4.7×10^{-4}	1.65×10^{-4}	-6.7×10^{-4}	9.3×10^{-5}	7.92×10^{-4}	2.74×10^{-4}
1	1.075×10^{-3}	-2.178×10^{-3}	-4.067×10^{-3}	-4.499×10^{-3}	-1.831×10^{-3}	-7.68×10^{-4}	-1.02×10^{-4}
0	1.365×10^{-3}	1.3×10^{-5}	-2.94×10^{-4}	-6.92×10^{-4}	-3.05×10^{-4}	-4.8×10^{-5}	-7.14×10^{-4}

Mean squared error ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	1.551×10^{-5}	2.984×10^{-4}	1.385×10^{-3}	1.893×10^{-3}	1.481×10^{-3}	1.143×10^{-3}	6.986×10^{-4}
10	1.461×10^{-5}	1.953×10^{-4}	8.177×10^{-4}	1.12×10^{-3}	8.126×10^{-4}	5.771×10^{-4}	3.903×10^{-4}
1	1.081×10^{-5}	1.269×10^{-4}	5.005×10^{-4}	7.426×10^{-4}	4.568×10^{-4}	3.416×10^{-4}	2.198×10^{-4}
0	1.381×10^{-5}	8.499×10^{-5}	2.807×10^{-4}	4.933×10^{-4}	3.233×10^{-4}	2.311×10^{-4}	1.461×10^{-4}

Median difference ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	0.	0.	0.	0.	0.	0.
10	0.	0.	0.	0.	0.	0.	0.
1	0.	0.	0.	0.	0.	0.	0.
0	0.	0.	0.	0.	0.	0.	0.

$s = 10{,}000$

$k = 50$

Mean difference ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	$6. \times 10^{-6}$	1.55×10^{-4}	2.93×10^{-4}	1.08×10^{-4}	1.17×10^{-4}	$-1. \times 10^{-6}$	-1.9×10^{-5}
10	$2. \times 10^{-6}$	6.2×10^{-5}	8.8×10^{-5}	$-9. \times 10^{-6}$	-1.6×10^{-5}	9.8×10^{-6}	-1.5×10^{-4}
1	$2. \times 10^{-6}$	9.2×10^{-5}	-8.7×10^{-5}	-1.14×10^{-4}	-9.4×10^{-5}	-3.15×10^{-4}	-3.29×10^{-4}
0	$1. \times 10^{-6}$	9.4×10^{-5}	1.52×10^{-4}	$6. \times 10^{-5}$	2.2×10^{-5}	1.35×10^{-4}	-3.3×10^{-5}

Mean squared error ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	$6. \times 10^{-8}$	3.705×10^{-6}	1.15×10^{-4}	1.344×10^{-4}	1.021×10^{-4}	8.581×10^{-5}	5.823×10^{-5}
10	$2. \times 10^{-8}$	2.416×10^{-6}	7.918×10^{-5}	8.801×10^{-5}	6.498×10^{-5}	5.396×10^{-5}	3.762×10^{-5}
1	$2. \times 10^{-8}$	5.56×10^{-6}	3.423×10^{-5}	5.516×10^{-5}	4.29×10^{-5}	3.467×10^{-5}	2.265×10^{-5}
0	$1. \times 10^{-8}$	6.24×10^{-6}	2.982×10^{-5}	4.404×10^{-5}	3.298×10^{-5}	2.541×10^{-5}	1.471×10^{-5}

Median difference ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	0.	0.	0.	0.	0.	0.
10	0.	0.	0.	0.	0.	0.	0.
1	0.	0.	0.	0.	0.	0.	0.
0	0.	0.	0.	0.	0.	0.	0.

$k = 100$

Mean difference ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	2.841×10^{-3}	4.217×10^{-2}	7.603×10^{-2}	5.058×10^{-2}	2.976×10^{-2}	1.129×10^{-2}	-5.523×10^{-3}
10	2.69×10^{-3}	1.346×10^{-2}	4.7×10^{-2}	3.552×10^{-2}	2.37×10^{-2}	1.23×10^{-2}	-7.894×10^{-3}
1	2.693×10^{-3}	2.985×10^{-3}	1.238×10^{-2}	1.204×10^{-2}	1.43×10^{-2}	9.416×10^{-3}	-7.867×10^{-3}
0	2.477×10^{-3}	-6.209×10^{-3}	-1.084×10^{-2}	-1.306×10^{-2}	-6.375×10^{-3}	7.879×10^{-3}	-8.884×10^{-3}

Mean squared error ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	3.581×10^{-5}	1.683×10^{-2}	4.414×10^{-2}	4.929×10^{-2}	5.07×10^{-2}	4.061×10^{-2}	1.766×10^{-2}
10	3.312×10^{-5}	5.591×10^{-3}	2.986×10^{-2}	4.118×10^{-2}	4.557×10^{-2}	3.778×10^{-2}	1.829×10^{-2}
1	3.465×10^{-5}	2.106×10^{-3}	1.45×10^{-2}	2.651×10^{-2}	3.428×10^{-2}	3.154×10^{-2}	1.782×10^{-2}
0	2.991×10^{-5}	1.056×10^{-3}	6.487×10^{-3}	1.528×10^{-2}	2.382×10^{-2}	2.091×10^{-2}	1.637×10^{-2}

Median difference ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	$-1. \times 10^{-2}$	$-2. \times 10^{-2}$	$-1. \times 10^{-2}$	$-2. \times 10^{-2}$	$-2. \times 10^{-2}$	$-1. \times 10^{-2}$
10	0.	0.	$-1. \times 10^{-2}$	$-1. \times 10^{-2}$	$-2. \times 10^{-2}$	$-2. \times 10^{-2}$	$-1. \times 10^{-2}$
1	0.	0.	0.	0.	$-1. \times 10^{-2}$	$-1. \times 10^{-2}$	$-1. \times 10^{-2}$
0	0.	$1. \times 10^{-2}$	$-2. \times 10^{-2}$	$1. \times 10^{-2}$	0.	$-1. \times 10^{-2}$	0.

$s = 1{,}000$

$k = 100$

Mean difference ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	2.27×10^{-4}	9.9×10^{-4}	4.484×10^{-3}	-1.01×10^{-3}	1.003×10^{-3}	6.766×10^{-3}	4.336×10^{-3}
10	1.91×10^{-4}	6.76×10^{-4}	7.96×10^{-4}	-3.137×10^{-3}	-1.724×10^{-3}	5.65×10^{-4}	9.86×10^{-4}
1	1.06×10^{-4}	-1.931×10^{-3}	-3.749×10^{-3}	-3.694×10^{-3}	-1.228×10^{-3}	-2.77×10^{-4}	-1.36×10^{-4}
0	2.14×10^{-4}	2.17×10^{-4}	-6.92×10^{-4}	-8.42×10^{-4}	-6.5×10^{-5}	-4.54×10^{-4}	-9.15×10^{-4}

Mean squared error ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	2.27×10^{-6}	4.479×10^{-4}	2.694×10^{-3}	2.649×10^{-3}	3.055×10^{-3}	2.289×10^{-3}	9.102×10^{-4}
10	1.91×10^{-6}	2.725×10^{-4}	1.377×10^{-3}	1.598×10^{-3}	1.381×10^{-3}	7.137×10^{-4}	4.912×10^{-4}
1	1.06×10^{-6}	1.651×10^{-4}	7.966×10^{-4}	8.498×10^{-4}	5.571×10^{-4}	3.061×10^{-4}	2.001×10^{-4}
0	2.14×10^{-6}	9.177×10^{-5}	3.956×10^{-4}	5.847×10^{-4}	2.782×10^{-4}	1.676×10^{-4}	1.156×10^{-4}

Median difference ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	0.	0.	0.	0.	0.	0.
10	0.	0.	0.	0.	0.	0.	0.
1	0.	0.	0.	0.	0.	0.	0.
0	0.	0.	0.	0.	0.	0.	0.

$s = 10{,}000$

$k = 100$

Mean difference ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	1.29×10^{-4}	2.3×10^{-5}	-4.5×10^{-5}	$3. \times 10^{-5}$	-7.7×10^{-5}	-4.3×10^{-5}
10	0.	1.9×10^{-5}	-1.07×10^{-4}	-8.8×10^{-5}	$8. \times 10^{-5}$	$-6. \times 10^{-5}$	-1.99×10^{-4}
1	0.	3.6×10^{-5}	-2.38×10^{-4}	-1.47×10^{-4}	-6.1×10^{-5}	-2.57×10^{-4}	-3.2×10^{-4}
0	0.	$8. \times 10^{-6}$	-2.2×10^{-5}	2.7×10^{-5}	-2.8×10^{-5}	$3. \times 10^{-6}$	$-2. \times 10^{-6}$

Mean squared error ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	4.489×10^{-5}	1.18×10^{-4}	9.935×10^{-5}	8.168×10^{-5}	6.719×10^{-5}	4.705×10^{-5}
10	0.	3.183×10^{-5}	8.423×10^{-5}	6.778×10^{-5}	5.154×10^{-5}	4.198×10^{-5}	2.985×10^{-5}
1	0.	6.62×10^{-6}	4.694×10^{-5}	4.933×10^{-5}	3.447×10^{-5}	2.735×10^{-5}	1.672×10^{-5}
0	0.	$6. \times 10^{-6}$	3.558×10^{-5}	4.251×10^{-5}	2.834×10^{-5}	1.827×10^{-5}	8.92×10^{-6}

Median difference ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	0.	0.	0.	0.	0.	0.
10	0.	0.	0.	0.	0.	0.	0.
1	0.	0.	0.	0.	0.	0.	0.
0	0.	0.	0.	0.	0.	0.	0.

$k = 200$

Mean difference ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	9.88×10^{-4}	6.758×10^{-2}	6.479×10^{-2}	3.734×10^{-2}	2.086×10^{-2}	5.614×10^{-3}	-9.349×10^{-3}
10	9.9×10^{-4}	2.996×10^{-2}	6.008×10^{-2}	4.057×10^{-2}	2.411×10^{-2}	5.231×10^{-3}	-9.197×10^{-3}
1	1.099×10^{-3}	5.279×10^{-3}	3.08×10^{-2}	3.465×10^{-2}	1.593×10^{-2}	2.928×10^{-3}	-1.005×10^{-3}
0	9.45×10^{-4}	-7.512×10^{-3}	-2.289×10^{-3}	3.408×10^{-3}	4.032×10^{-3}	2.14×10^{-3}	-1.379×10^{-3}

Mean squared error ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	9.94×10^{-4}	3.142×10^{-2}	5.312×10^{-2}	6.503×10^{-2}	6.963×10^{-2}	5.729×10^{-2}	2.906×10^{-2}
10	9.98×10^{-4}	1.377×10^{-2}	4.602×10^{-2}	6.071×10^{-2}	6.588×10^{-2}	5.466×10^{-2}	2.905×10^{-2}
1	1.121×10^{-3}	4.072×10^{-3}	2.806×10^{-2}	4.828×10^{-2}	6.014×10^{-2}	5.303×10^{-2}	2.958×10^{-2}
0	9.49×10^{-4}	1.647×10^{-3}	1.403×10^{-2}	3.028×10^{-2}	4.367×10^{-2}	4.347×10^{-2}	2.947×10^{-2}

Median difference ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	$-1. \times 10^{-2}$	0.	0.	$-2. \times 10^{-2}$	$-2. \times 10^{-2}$	$-1. \times 10^{-2}$
10	0.	0.	$-1. \times 10^{-2}$	$-1. \times 10^{-2}$	$-2. \times 10^{-2}$	$-2. \times 10^{-2}$	$-1. \times 10^{-2}$
1	0.	$1. \times 10^{-2}$	0.	$-1. \times 10^{-2}$	$-1. \times 10^{-2}$	$-2. \times 10^{-2}$	$-1. \times 10^{-2}$
0	0.	$1. \times 10^{-2}$	$1. \times 10^{-2}$	0.	$-1. \times 10^{-2}$	$-1. \times 10^{-2}$	$-1. \times 10^{-2}$

$s = 1{,}000$

$k = 200$

Mean difference ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	3.816×10^{-3}	1.128×10^{-2}	6.613×10^{-3}	1.541×10^{-2}	2.375×10^{-2}	1.105×10^{-2}
10	0.	8.86×10^{-4}	5.59×10^{-4}	-5.564×10^{-3}	-2.634×10^{-3}	8.205×10^{-3}	4.794×10^{-3}
1	0.	-1.969×10^{-3}	-3.33×10^{-3}	-4.909×10^{-3}	-2.511×10^{-3}	9.65×10^{-4}	5.6×10^{-4}
0	$1. \times 10^{-4}$	-4.27×10^{-4}	-2.36×10^{-3}	-1.538×10^{-3}	-1.282×10^{-3}	-1.75×10^{-4}	-4.47×10^{-4}

Mean squared error ψ

$p \mid \psi$	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	1.048×10^{-3}	5.545×10^{-3}	7.062×10^{-3}	9.731×10^{-3}	6.746×10^{-3}	1.81×10^{-3}
10	0.	4.718×10^{-4}	2.939×10^{-3}	3.16×10^{-3}	4.425×10^{-3}	3.076×10^{-3}	1.084×10^{-3}
1	0.	2.868×10^{-4}	1.474×10^{-3}	1.489×10^{-3}	1.095×10^{-3}	7.221×10^{-4}	5.448×10^{-4}
0	$1. \times 10^{-8}$	1.307×10^{-4}	6.774×10^{-4}	8.781×10^{-4}	5.212×10^{-4}	2.164×10^{-4}	1.753×10^{-4}

Median difference ψ

| $p

Table SI D_2 – Overview of the (marginal) accuracy $\hat{\rho}$. Each cell shows the mean difference (first row), the mean squared error (second row) and the median difference (third row) of $\hat{\rho}$ (right column) calculated over 10,000 data sets assuming linkage equilibrium. Colors within each sub-table range from light yellow to dark red and scale between the minimal and the maximal absolute value to aid interpretation.

$s = 100$

$k = 20$

Mean difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	1.245×10^{1}	1.585×10^{2}	2.235×10^{2}	2.548×10^{2}	2.524×10^{2}	2.165×10^{2}	9.386×10^{1}
10	2.586×10^{-1}	1.038×10^{1}	3.102×10^{1}	6.858×10^{1}	1.117×10^{2}	1.164×10^{2}	1.059×10^{2}
1	1.48×10^{-1}	6.505×10^{-1}	1.177	3.246	1.185×10^{1}	2.658×10^{1}	4.085×10^{1}
0	9.09×10^{-2}	2.062×10^{-1}	2.808×10^{-1}	5.153×10^{-1}	1.264	4.267	1.169×10^{1}

Mean squared error ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	1.345×10^{4}	1.322×10^{5}	2.019×10^{5}	2.387×10^{5}	2.448×10^{5}	2.221×10^{5}	1.262×10^{5}
10	4.408×10^{1}	2.411×10^{3}	1.66×10^{4}	4.904×10^{4}	9.1×10^{4}	9.718×10^{4}	8.951×10^{4}
1	7.644×10^{-1}	3.699	1.836×10^{1}	7.302×10^{2}	7.272×10^{3}	2.076×10^{4}	3.189×10^{4}
0	9.39×10^{-2}	2.704×10^{-1}	5.364×10^{-1}	7.736×10^{1}	4.372×10^{2}	3.073×10^{3}	8.885×10^{3}

Median difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	-2.1×10^{1}	1.	9.	8.	-9.	-3.7×10^{1}	$-9. \times 10^{1}$
10	-1.	0.	0.	1.	1.	-1.	-7.
1	0.	0.	0.	0.	0.	0.	0.
0	0.	0.	0.	0.	0.	0.	0.

$k = 50$

Mean difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	5.106	1.975×10^{2}	2.533×10^{2}	2.493×10^{2}	1.993×10^{2}	4.816×10^{1}	-5.662×10^{1}
10	1.743×10^{-1}	2.258×10^{1}	7.956×10^{1}	1.353×10^{2}	1.552×10^{2}	9.787×10^{1}	2.753×10^{1}
1	1.259×10^{-1}	8.85×10^{-1}	5.567	1.955×10^{1}	4.136×10^{1}	4.992×10^{1}	2.37×10^{1}
0	8.25×10^{-2}	2.401×10^{-1}	4.802×10^{-1}	1.576	7.867	1.805×10^{1}	1.622×10^{1}

Mean squared error ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	6.899×10^{3}	1.767×10^{5}	2.442×10^{5}	2.457×10^{5}	2.171×10^{5}	8.605×10^{4}	3.484×10^{4}
10	3.448×10^{1}	9.44×10^{3}	6.016×10^{4}	1.144×10^{5}	1.404×10^{5}	8.105×10^{4}	2.954×10^{4}
1	6.159×10^{-1}	1.011×10^{1}	2.479×10^{2}	1.377×10^{4}	3.369×10^{4}	4.191×10^{4}	2.046×10^{4}
0	8.45×10^{-2}	3.823×10^{-1}	8.497	5.556×10^{2}	5.979×10^{3}	1.568×10^{4}	1.433×10^{4}

Median difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	-1.7×10^{1}	1.	-6.	-2.1×10^{1}	-6.2×10^{1}	-8.7×10^{1}	-9.5×10^{1}
10	-1.	1.	0.	1.	-1.	-5.	-8.
1	0.	0.	0.	0.	0.	0.	0.
0	0.	0.	0.	0.	0.	0.	0.

$k = 100$

Mean difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	5.204	2.292×10^{2}	2.342×10^{2}	2.079×10^{2}	9.533×10^{1}	1.199×10^{1}	-6.232×10^{1}
10	2.34×10^{-1}	5.372×10^{1}	1.334×10^{2}	1.737×10^{2}	1.379×10^{2}	8.905×10^{1}	2.825×10^{1}
1	1.695×10^{-1}	2.009	1.96×10^{1}	5.467×10^{1}	6.883×10^{1}	7.3×10^{1}	3.522×10^{1}
0	8.09×10^{-2}	3.203×10^{-1}	2.421	1.148×10^{1}	2.686×10^{1}	3.35×10^{1}	3.103×10^{1}

Mean squared error ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	5.624×10^{3}	2.159×10^{5}	2.355×10^{5}	2.189×10^{5}	1.407×10^{5}	9.281×10^{4}	3.907×10^{4}
10	3.102×10^{1}	3.511×10^{4}	1.135×10^{5}	1.554×10^{5}	1.269×10^{5}	9.224×10^{4}	3.678×10^{4}
1	6.511×10^{-1}	2.741×10^{2}	1.286×10^{4}	4.608×10^{4}	6.077×10^{4}	7.098×10^{4}	3.585×10^{4}
0	8.29×10^{-2}	7.583×10^{-1}	1.341×10^{3}	9.040×10^{3}	2.344×10^{4}	3.171×10^{4}	3.105×10^{4}

Median difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	-1.4×10^{1}	-5.	-2.9×10^{1}	-6.3×10^{1}	-8.8×10^{1}	-9.6×10^{1}	-9.8×10^{1}
10	-1.	1.	0.	0.	-5.	-8.	-9.
1	0.	0.	0.	0.	0.	0.	-1.
0	0.	0.	0.	0.	0.	0.	0.

$k = 200$

Mean difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	8.32	2.483×10^{2}	2.486×10^{2}	1.621×10^{2}	$9. \times 10^{1}$	2.713×10^{1}	-5.286×10^{1}
10	7.004 $\times 10^{-1}$	1.044×10^{2}	1.888×10^{2}	1.875×10^{2}	1.521×10^{2}	1.081×10^{2}	3.682×10^{1}
1	2.047×10^{-1}	1.037×10^{1}	6.002×10^{1}	1.015×10^{2}	1.237×10^{2}	1.054×10^{2}	4.512×10^{1}
0	8.77×10^{-2}	1.047	1.61×10^{1}	3.682×10^{1}	6.2×10^{1}	7.368×10^{1}	4.441×10^{1}

Mean squared error ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	4.34×10^{3}	2.406×10^{5}	2.503×10^{5}	1.984×10^{5}	1.524×10^{5}	1.12×10^{5}	4.828×10^{4}
10	3.215×10^{1}	8.272×10^{4}	1.692×10^{5}	1.768×10^{5}	1.506×10^{5}	1.159×10^{5}	4.664×10^{4}
1	7.077×10^{-1}	5.755×10^{3}	5.003×10^{4}	9.252×10^{4}	1.182×10^{5}	1.061×10^{5}	4.678×10^{4}
0	9.05×10^{-2}	2.893×10^{2}	1.302×10^{4}	3.226×10^{4}	5.827×10^{4}	7.274×10^{4}	4.508×10^{4}

Median difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	-5.	-1.45×10^{1}	-4.9×10^{1}	-7.4×10^{1}	-9.1×10^{1}	-9.7×10^{1}	-9.9×10^{1}
10	0.	1.	1.	-3.	-8.	-9.	$-1. \times 10^{1}$
1	0.	0.	0.	0.	0.	-1.	-1.
0	0.	0.	0.	0.	0.	0.	0.

$s = 1{,}000$

$k = 20$

Mean difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	-4.7	1.634×10^{1}	3.311×10^{1}	5.378×10^{1}	7.642×10^{1}	9.887×10^{1}	1.543×10^{2}
10	-3.441×10^{-1}	5.722×10^{-1}	1.013	1.75	2.412	3.546	1.151×10^{1}
1	2.8×10^{-2}	1.271×10^{-1}	1.833×10^{-1}	2.234×10^{-1}	2.58×10^{-1}	3.268×10^{-1}	5.512×10^{-1}
0	0.	2.5×10^{-3}	6.7×10^{-3}	1.19×10^{-2}	1.52×10^{-2}	3.3×10^{-2}	1.161×10^{-1}

Mean squared error ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	6.231×10^{2}	4.446×10^{3}	1.412×10^{4}	2.912×10^{4}	4.903×10^{4}	7.484×10^{4}	1.321×10^{5}
10	3.699	1.169×10^{1}	2.397×10^{1}	4.722×10^{1}	8.49×10^{1}	1.908×10^{2}	2.892×10^{3}
1	2.96×10^{-2}	1.351×10^{-1}	2.199×10^{-1}	3.04×10^{-1}	3.826×10^{-1}	6.798×10^{-1}	2.812
0	0.	2.5×10^{-3}	6.7×10^{-3}	1.19×10^{-2}	1.52×10^{-2}	3.3×10^{-2}	1.231×10^{-1}

Median difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	-7.	1.	-1.	3.	3.	-4.	-1.
10	0.	0.	0.	0.	0.	0.	0.
1	0.	0.	0.	0.	0.	0.	0.
0	0.	0.	0.	0.	0.	0.	0.

$k = 50$

Mean difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	-3.066	2.532×10^{1}	6.375×10^{1}	1.051×10^{2}	1.348×10^{2}	1.666×10^{2}	2.286×10^{2}
10	-1.704×10^{-1}	8.411×10^{-1}	2.017	3.386	5.279	1.059×10^{1}	3.405×10^{1}
1	2.8×10^{-2}	1.548×10^{-1}	2.312×10^{-1}	2.714×10^{-1}	3.077×10^{-1}	4.699×10^{-1}	1.118
0	$1. \times 10^{-4}$	3.4×10^{-3}	1.19×10^{-2}	1.9×10^{-2}	2.23×10^{-2}	5.22×10^{-2}	1.595×10^{-1}

Mean squared error ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	4.799×10^{2}	9.139×10^{3}	3.915×10^{4}	7.543×10^{4}	1.045×10^{5}	1.419×10^{5}	2.079×10^{5}
10	3.202	1.944×10^{1}	5.89×10^{1}	3.117×10^{2}	8.246×10^{2}	2.777×10^{3}	$2. \times 10^{4}$
1	2.9×10^{-2}	1.75×10^{-1}	3.3×10^{-1}	4.334×10^{-1}	5.613×10^{-1}	1.523	3.975×10^{1}
0	$1. \times 10^{-4}$	3.4×10^{-3}	1.19×10^{-2}	1.9×10^{-2}	2.23×10^{-2}	5.26×10^{-2}	1.969×10^{-1}

Median difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	-4.	1.	0.	4.	6.	1.	8.
10	0.	0.	0.	0.	0.	0.	0.
1	0.	0.	0.	0.	0.	0.	0.
0	0.	0.	0.	0.	0.	0.	0.

$k = 100$

Mean difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	7.21×10^{-1}	5.038×10^{1}	1.118×10^{2}	1.597×10^{2}	1.959×10^{2}	2.248×10^{2}	2.488×10^{2}
10	6.33×10^{-2}	1.379	4.063	1.074×10^{1}	1.72×10^{1}	2.937×10^{1}	8.004×10^{1}
1	3.29×10^{-2}	1.911×10^{-1}	2.928×10^{-1}	3.788×10^{-1}	5.623×10^{-1}	8.344×10^{-1}	3.284
0	0.	9.5×10^{-3}	2.43×10^{-2}	2.92×10^{-2}	3.75×10^{-2}	8.34×10^{-2}	2.88×10^{-1}

Mean squared error ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	3.516×10^{2}	2.534×10^{4}	8.351×10^{4}	1.279×10^{5}	1.688×10^{5}	2.009×10^{5}	2.439×10^{5}
10	2.632	3.597×10^{1}	3.492×10^{2}	4.319×10^{3}	8.492×10^{3}	1.597×10^{4}	6.132×10^{4}
1	3.31×10^{-2}	2.443×10^{-1}	5.748×10^{-1}	2.195	3.155×10^{1}	1.127×10^{2}	9.346×10^{2}
0	0.	9.5×10^{-3}	2.43×10^{-2}	3.04×10^{-2}	3.95×10^{-2}	8.56×10^{-2}	1.101×10^{1}

Median difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	2.	1.	1.1×10^{1}	$1. \times 10^{1}$	8.	-6.
10	0.	0.	0.	0.	0.	0.	1.
1	0.	0.	0.	0.	0.	0.	0.
0	0.	0.	0.	0.	0.	0.	0.

$k = 200$

Mean difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	1.595	8.491×10^{1}	1.629×10^{2}	2.094×10^{2}	2.319×10^{2}	2.248×10^{2}	1.838×10^{2}
10	1.273×10^{-1}	2.728	1.501×10^{1}	2.954×10^{1}	5.211×10^{1}	7.246×10^{1}	9.777×10^{1}
1	3.03×10^{-2}	2.564×10^{-1}	4.647×10^{-1}	9.021×10^{-1}	1.534	3.409	1.402×10^{1}
0	0.	2.53×10^{-2}	5.32×10^{-2}	5.89×10^{-2}	8.3×10^{-2}	1.86×10^{-1}	6.24×10^{-1}

Mean squared error ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	2.625×10^{2}	5.72×10^{4}	1.352×10^{5}	1.83×10^{5}	2.145×10^{5}	2.229×10^{5}	2.019×10^{5}
10	2.362	1.236×10^{2}	6.429×10^{3}	1.645×10^{4}	3.711×10^{4}	5.74×10^{4}	8.2×10^{4}
1	3.03×10^{-2}	4.592×10^{-1}	2.377	1.16×10^{2}	4.881×10^{2}	1.695×10^{3}	9.622×10^{3}
0	0.	2.53×10^{-2}	5.46×10^{-2}	7.47×10^{-2}	1.492×10^{-1}	1.884×10^{1}	3.067×10^{1}

Median difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	2.	6.	1.2×10^{1}	2.	-7.	-1.6×10^{1}
10	0.	0.	1.	1.	1.	0.	0.
1	0.	0.	0.	0.	0.	0.	0.
0	0.	0.	0.	0.	0.	0.	0.

$s = 10{,}000$

$k = 20$

Mean difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	-1.574	1.664	2.727	4.297	5.176	7.516	1.693×10^{1}
10	-6.07×10^{-2}	8.51×10^{-2}	1.333×10^{-1}	1.769×10^{-1}	2.391×10^{-1}	3.182×10^{-1}	5.957×10^{-1}
1	0.	$2. \times 10^{-4}$	$4. \times 10^{-4}$	5.6×10^{-3}	$1. \times 10^{-2}$	2.61×10^{-2}	1.1×10^{-1}
0	0.	0.	0.	0.	0.	0.	0.

Mean squared error ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	7.551×10^{1}	2.615×10^{2}	4.617×10^{2}	7.366×10^{2}	9.65×10^{2}	1.696×10^{3}	5.284×10^{3}
10	4.857×10^{-1}	1.166	1.887	2.552	3.397	5.354	1.291×10^{1}
1	0.	$2. \times 10^{-4}$	$4. \times 10^{-4}$	5.6×10^{-3}	$1. \times 10^{-2}$	2.61×10^{-2}	1.118×10^{-1}
0	0.	0.	0.	0.	0.	0.	0.

Median difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	-1.	0.	0.	1.	0.	-1.	0.
10	0.	0.	0.	0.	0.	0.	0.
1	0.	0.	0.	0.	0.	0.	0.
0	0.	0.	0.	0.	0.	0.	0.

$k = 50$

Mean difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	2.126×10^{-1}	2.379	4.309	6.546	1.04×10^{1}	1.665×10^{1}	3.377×10^{1}
10	2.17×10^{-2}	9.78×10^{-2}	1.64×10^{-1}	2.067×10^{-1}	2.617×10^{-1}	5.249×10^{-1}	1.067
1	0.	$7. \times 10^{-4}$	4.6×10^{-3}	$1. \times 10^{-2}$	1.77×10^{-2}	5.75×10^{-2}	1.476×10^{-1}
0	0.	0.	0.	0.	0.	0.	$1. \times 10^{-4}$

Mean squared error ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	3.586×10^{1}	4.501×10^{2}	8.538×10^{2}	1.273×10^{3}	2.207×10^{3}	4.64×10^{3}	1.433×10^{4}
10	3.197×10^{-1}	1.716	2.794	3.362	4.789	9.538	2.354×10^{1}
1	0.	$7. \times 10^{-4}$	4.6×10^{-3}	$1. \times 10^{-2}$	1.77×10^{-2}	5.75×10^{-2}	1.616×10^{-1}
0	0.	0.	0.	0.	0.	0.	$1. \times 10^{-4}$

Median difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	0.	0.	0.	1.	0.	2.
10	0.	0.	0.	0.	0.	0.	0.
1	0.	0.	0.	0.	0.	0.	0.
0	0.	0.	0.	0.	0.	0.	0.

$k = 100$

Mean difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	1.288×10^{-1}	3.931	6.772	1.075×10^{1}	1.988×10^{1}	3.211×10^{1}	6.551×10^{1}
10	2.42×10^{-2}	1.578×10^{-1}	2.59×10^{-1}	3.666×10^{-1}	4.911×10^{-1}	8.883×10^{-1}	1.904
1	0.	$3. \times 10^{-3}$	1.35×10^{-2}	1.81×10^{-2}	3.7×10^{-2}	9.79×10^{-2}	2.088×10^{-1}
0	0.	0.	0.	0.	0.	0.	2.3×10^{-3}

Mean squared error ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	2.937×10^{1}	7.107×10^{2}	1.344×10^{3}	2.454×10^{3}	5.933×10^{3}	1.264×10^{4}	3.989×10^{4}
10	2.9×10^{-1}	2.583	3.8	5.188	8.617	1.748×10^{1}	5.212×10^{1}
1	0.	$3. \times 10^{-3}$	1.35×10^{-2}	1.81×10^{-2}	3.7×10^{-2}	9.87×10^{-2}	2.478×10^{-1}
0	0.	0.	0.	0.	0.	0.	2.3×10^{-3}

Median difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	0.	0.	0.	2.	1.	3.
10	0.	0.	0.	0.	0.	0.	0.
1	0.	0.	0.	0.	0.	0.	0.
0	0.	0.	0.	0.	0.	0.	0.

$k = 200$

Mean difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	1.442×10^{-1}	5.914	1.301×10^{1}	2.189×10^{1}	3.834×10^{1}	6.96×10^{1}	1.132×10^{2}
10	2.45×10^{-2}	2.47×10^{-1}	3.618×10^{-1}	5.416×10^{-1}	9.503×10^{-1}	1.672	3.814
1	0.	1.36×10^{-2}	2.42×10^{-2}	4.01×10^{-2}	8.13×10^{-2}	1.592×10^{-1}	2.999×10^{-1}
0	0.	0.	0.	0.	0.	0.	8.3×10^{-3}

Mean squared error ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	2.513×10^{1}	1.105×10^{3}	2.934×10^{3}	6.824×10^{3}	1.761×10^{4}	4.165×10^{4}	8.943×10^{4}
10	2.855×10^{-1}	3.744	5.213	8.901	1.868×10^{1}	4.334×10^{1}	2.198×10^{2}
1	0.	1.36×10^{-2}	2.42×10^{-2}	4.01×10^{-2}	8.17×10^{-2}	1.672×10^{-1}	4.585×10^{-1}
0	0.	0.	0.	0.	0.	0.	8.3×10^{-3}

Median difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	0.	2.	1.	1.	4.	4.
10	0.	0.	0.	0.	0.	0.	0.
1	0.	0.	0.	0.	0.	0.	0.
0	0.	0.	0.	0.	0.	0.	0.

Table SI D_3 – Comparison of the (marginal) accuracy of $\hat{\psi}$ (left column) and $\hat{\rho}$ (right column) for $k = 100$ (i.e., the reference) and $k = 200$. Each cell shows the difference of the absolute mean difference ($|MD_{k=100}| - |MD_{k=200}|$; first row), the difference of the mean absolute difference ($MAD_{k=100} - MAD_{k=200}$; third row) and the difference of the mean squared error ($MSE_{k=100} - MSE_{k=200}$; third row) each calculated over 10,000 data sets assuming linkage equilibrium and $s = 10{,}000$. Colors within each sub-table range from light yellow to dark red and scale between the minimal and the maximal absolute value to aid interpretation.

Mean difference ψ

100	0.	2.4×10^{-5}	-2.27×10^{-4}	-1.18×10^{-4}	-2.3×10^{-5}	$5. \times 10^{-6}$	-1.31×10^{-4}
10	0.	$7. \times 10^{-6}$	-6.9×10^{-5}	$-8. \times 10^{-5}$	6.8×10^{-5}	1.8×10^{-5}	6.6×10^{-5}
1	0.	-2.3×10^{-5}	-1.06×10^{-4}	-4.6×10^{-5}	-1.85×10^{-4}	2.7×10^{-5}	7.1×10^{-5}
0	0.	$-7. \times 10^{-6}$	$-5. \times 10^{-6}$	-6.6×10^{-5}	-9.5×10^{-5}	-7.7×10^{-5}	-4.1×10^{-5}
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Mean difference ρ

100	-1.54×10^{-2}	-1.982	-6.235	-1.113×10^{1}	-1.847×10^{1}	-3.749×10^{1}	-4.769×10^{1}
10	$-3. \times 10^{-4}$	-8.92×10^{-2}	-1.028×10^{-1}	-1.75×10^{-1}	-4.592×10^{-1}	-7.832×10^{-1}	-1.91
1	0.	-1.06×10^{-2}	-1.07×10^{-2}	-2.2×10^{-2}	-4.43×10^{-2}	-6.13×10^{-2}	-9.11×10^{-2}
0	0.	0.	0.	0.	0.	0.	$-6. \times 10^{-3}$
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Mean absolute difference ψ

100	0.	-6.48×10^{-4}	5.21×10^{-4}	7.42×10^{-4}	4.69×10^{-4}	2.89×10^{-4}	3.59×10^{-4}
10	0.	-8.47×10^{-4}	5.33×10^{-4}	7.42×10^{-4}	4.44×10^{-4}	3.24×10^{-4}	4.1×10^{-4}
1	0.	-6.13×10^{-4}	-5.08×10^{-4}	5.4×10^{-4}	4.79×10^{-4}	3.73×10^{-4}	3.67×10^{-4}
0	0.	-3.23×10^{-4}	-6.59×10^{-4}	3.56×10^{-4}	4.69×10^{-4}	2.91×10^{-4}	1.85×10^{-4}
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Mean absolute difference ρ

100	3.118×10^{-1}	-4.5	-8.458	-1.537×10^{1}	-2.412×10^{1}	-4.279×10^{1}	-5.423×10^{1}
10	4.9×10^{-3}	-2.87×10^{-1}	-2.364×10^{-1}	-4.912×10^{-1}	-8.356×10^{-1}	-1.258	-2.543
1	0.	-1.06×10^{-2}	-1.07×10^{-2}	-2.2×10^{-2}	-4.43×10^{-2}	-6.13×10^{-2}	-1.061×10^{-1}
0	0.	0.	0.	0.	0.	0.	$-6. \times 10^{-3}$
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Mean squared error ψ

100	0.	-9.7×10^{-6}	7.17×10^{-6}	1.476×10^{-5}	7.97×10^{-6}	3.61×10^{-6}	4.11×10^{-6}
10	0.	-1.033×10^{-5}	8.53×10^{-6}	1.214×10^{-5}	6.8×10^{-6}	3.84×10^{-6}	4.22×10^{-6}
1	0.	-7.51×10^{-6}	-7.88×10^{-6}	7.64×10^{-6}	5.23×10^{-6}	3.97×10^{-6}	3.69×10^{-6}
0	0.	-3.23×10^{-6}	-8.51×10^{-6}	4.38×10^{-6}	4.89×10^{-6}	2.91×10^{-6}	1.85×10^{-6}
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Mean squared error ρ

100	4.239	-3.945×10^{2}	-1.59×10^{3}	-4.371×10^{3}	-1.168×10^{4}	-2.902×10^{4}	-4.455×10^{4}
10	4.5×10^{-3}	-1.181	-1.414	-3.713	-1.006×10^{1}	-2.586×10^{1}	-1.677×10^{2}
1	0.	-1.06×10^{-2}	-1.07×10^{-2}	-2.2×10^{-2}	-4.47×10^{-2}	-6.85×10^{-2}	-2.107×10^{-1}
0	0.	0.	0.	0.	0.	0.	$-6. \times 10^{-3}$
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Table SI D_4 – Comparison of the (marginal) accuracy of $\hat{\psi}$ (left column) and $\hat{\rho}$ (right column) for $k = 100$ (i.e., the reference) and $k = 50$. Each cell shows the difference of the absolute mean difference ($|MD_{k=100}| - |MD_{k=50}|$; first row), the difference of the mean absolute difference ($MAD_{k=100} - MAD_{k=50}$; third row) and the difference of the mean squared error ($MSE_{k=100} - MSE_{k=50}$; third row) each calculated over 10,000 data sets assuming linkage equilibrium and $s = 10{,}000$. Colors within each sub-table range from light yellow to dark red and scale between the minimal and the maximal absolute value to aid interpretation.

Mean difference ψ

100	$-6. \times 10^{-5}$	-2.6×10^{-5}	-2.7×10^{-4}	-6.3×10^{-5}	-8.7×10^{-5}	7.6×10^{-5}	2.4×10^{-5}
10	$-2. \times 10^{-5}$	-4.3×10^{-5}	1.9×10^{-5}	7.9×10^{-5}	6.4×10^{-5}	-3.8×10^{-5}	4.9×10^{-5}
1	$-2. \times 10^{-5}$	-5.6×10^{-5}	1.51×10^{-4}	3.3×10^{-5}	-3.3×10^{-5}	-5.8×10^{-5}	$-9. \times 10^{-6}$
0	$-1. \times 10^{-5}$	-1.4×10^{-5}	-1.3×10^{-4}	-3.3×10^{-5}	$6. \times 10^{-6}$	-1.32×10^{-4}	-3.1×10^{-5}
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Mean difference ρ

100	-8.38×10^{-2}	1.553	2.463	4.208	9.473	1.547×10^{1}	3.174×10^{1}
10	2.5×10^{-3}	$6. \times 10^{-2}$	9.5×10^{-2}	1.599×10^{-1}	2.294×10^{-1}	3.634×10^{-1}	8.374×10^{-1}
1	0.	2.3×10^{-3}	8.9×10^{-3}	8.1×10^{-3}	1.93×10^{-2}	4.04×10^{-2}	6.12×10^{-2}
0	0.	0.	0.	0.	0.	0.	2.2×10^{-3}
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Mean absolute difference ψ

100	$-6. \times 10^{-5}$	5.82×10^{-4}	5.4×10^{-5}	-1.521×10^{-3}	-9.71×10^{-4}	-9.06×10^{-4}	-7.64×10^{-4}
10	$-2. \times 10^{-5}$	7.37×10^{-4}	2.03×10^{-4}	-1.041×10^{-3}	-8.66×10^{-4}	-8.5×10^{-4}	-6.77×10^{-4}
1	$-2. \times 10^{-5}$	8.2×10^{-5}	8.61×10^{-4}	-2.97×10^{-4}	-6.75×10^{-4}	-6.58×10^{-4}	-5.85×10^{-4}
0	$-1. \times 10^{-5}$	-2.4×10^{-5}	5.14×10^{-4}	-1.13×10^{-4}	-4.16×10^{-4}	-6.94×10^{-4}	-5.79×10^{-4}
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Mean absolute difference ρ

100	-4.316×10^{-1}	3.899	4.463	7.756	1.448×10^{1}	2.078×10^{1}	3.749×10^{1}
10	-2.55×10^{-2}	2.53×10^{-1}	2.012×10^{-1}	3.145×10^{-1}	5.108×10^{-1}	6.874×10^{-1}	1.248
1	0.	2.3×10^{-3}	8.9×10^{-3}	8.1×10^{-3}	1.93×10^{-2}	4.04×10^{-2}	6.52×10^{-2}
0	0.	0.	0.	0.	0.	0.	2.2×10^{-3}
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Mean squared error ψ

100	$-6. \times 10^{-8}$	7.84×10^{-6}	2.96×10^{-6}	-3.501×10^{-5}	-2.039×10^{-5}	-1.862×10^{-5}	-1.118×10^{-5}
10	$-2. \times 10^{-8}$	7.67×10^{-6}	5.05×10^{-6}	-2.023×10^{-5}	-1.344×10^{-5}	-1.198×10^{-5}	-7.77×10^{-6}
1	$-2. \times 10^{-8}$	1.06×10^{-6}	1.271×10^{-5}	-5.83×10^{-6}	-8.43×10^{-6}	-7.32×10^{-6}	-5.93×10^{-6}
0	$-1. \times 10^{-8}$	-2.4×10^{-7}	5.76×10^{-6}	-1.53×10^{-6}	-4.64×10^{-6}	-7.14×10^{-6}	-5.79×10^{-6}
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Mean squared error ρ

100	-6.487	2.606×10^{2}	4.898×10^{2}	1.181×10^{3}	3.726×10^{3}	7.996×10^{3}	2.455×10^{4}
10	-2.97×10^{-2}	8.468×10^{-1}	1.005	1.826	3.828	7.944	2.858×10^{1}
1	0.	2.3×10^{-3}	8.9×10^{-3}	8.1×10^{-3}	1.93×10^{-2}	4.12×10^{-2}	8.62×10^{-2}
0	0.	0.	0.	0.	0.	0.	2.2×10^{-3}
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Table SI D_5 – Comparison of the (marginal) accuracy of $\hat{\psi}$ (left column) and $\hat{\rho}$ (right column) for $s = 10{,}000$ (i.e., the reference) and $s = 1{,}000$. Each cell shows the difference of the absolute mean difference ($|MD_{s=10{,}000}| - |MD_{s=1{,}000}|$; first row), the difference of the mean absolute difference ($MAD_{s=10{,}000} - MAD_{s=1{,}000}$; third row) and the difference of the mean squared error ($MSE_{s=10{,}000} - MSE_{s=1{,}000}$; third row) each calculated over 10,000 data sets assuming linkage equilibrium and $k = 100$. Colors within each sub-table range from light yellow to dark red and scale between the minimal and the maximal absolute value to aid interpretation.

Mean difference ψ

100	-2.27×10^{-4}	-8.61×10^{-4}	-4.461×10^{-3}	-9.65×10^{-4}	-9.73×10^{-4}	-6.689×10^{-3}	-4.293×10^{-3}
10	-1.91×10^{-4}	-6.57×10^{-4}	-6.89×10^{-4}	-3.049×10^{-3}	-1.644×10^{-3}	-5.05×10^{-4}	-7.87×10^{-4}
1	-1.06×10^{-4}	-1.895×10^{-3}	-3.511×10^{-3}	-3.547×10^{-3}	-1.167×10^{-3}	$-2. \times 10^{-5}$	1.84×10^{-4}
0	-2.14×10^{-4}	-1.37×10^{-4}	-6.7×10^{-4}	-8.15×10^{-4}	-3.7×10^{-5}	-4.51×10^{-4}	-9.13×10^{-4}
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Mean difference ρ

100	-5.922×10^{-1}	-4.645×10^{1}	-1.05×10^{2}	-1.49×10^{2}	-1.76×10^{2}	-1.927×10^{2}	-1.833×10^{2}
10	-3.91×10^{-2}	-1.222	-3.804	-1.038×10^{1}	-1.671×10^{1}	-2.848×10^{1}	-7.814×10^{1}
1	-3.29×10^{-2}	-1.881×10^{-1}	-2.793×10^{-1}	-3.605×10^{-1}	-5.253×10^{-1}	-7.365×10^{-1}	-3.075
0	0.	-9.5×10^{-3}	-2.43×10^{-2}	-2.92×10^{-2}	-3.75×10^{-2}	-8.34×10^{-2}	-2.857×10^{-1}
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Mean absolute difference ψ

100	-2.27×10^{-4}	-1.169×10^{-2}	-2.687×10^{-2}	-2.576×10^{-2}	-2.493×10^{-2}	-2.251×10^{-2}	-1.606×10^{-2}
10	-1.91×10^{-4}	-9.343×10^{-3}	-1.957×10^{-2}	-1.95×10^{-2}	-1.654×10^{-2}	-1.306×10^{-2}	-1.167×10^{-2}
1	-1.06×10^{-4}	-8.323×10^{-3}	-1.605×10^{-2}	-1.534×10^{-2}	-1.156×10^{-2}	-9.416×10^{-3}	-8.246×10^{-3}
0	-2.14×10^{-4}	-5.947×10^{-3}	-1.087×10^{-2}	-1.279×10^{-2}	-9.423×10^{-3}	-7.779×10^{-3}	-6.701×10^{-3}
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Mean absolute difference ρ

100	-1.012×10^{1}	-6.693×10^{1}	-1.317×10^{2}	-1.723×10^{2}	-2.008×10^{2}	-2.183×10^{2}	-2.142×10^{2}
10	-9.315×10^{-1}	-2.884	-5.7	-1.234×10^{1}	-1.881×10^{1}	-3.084×10^{1}	-8.071×10^{1}
1	-3.31×10^{-2}	-2.169×10^{-1}	-3.375×10^{-1}	-4.049×10^{-1}	-5.959×10^{-1}	-8.959×10^{-1}	-3.418
0	0.	-9.5×10^{-3}	-2.43×10^{-2}	-2.92×10^{-2}	-3.75×10^{-2}	-8.34×10^{-2}	-2.857×10^{-1}
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Mean squared error ψ

100	-2.27×10^{-6}	-4.03×10^{-4}	-2.576×10^{-3}	-2.55×10^{-3}	-2.973×10^{-3}	-2.221×10^{-3}	-8.632×10^{-4}
10	-1.91×10^{-6}	-2.406×10^{-4}	-1.293×10^{-3}	-1.53×10^{-3}	-1.33×10^{-3}	-6.717×10^{-4}	-4.614×10^{-4}
1	-1.06×10^{-6}	-1.584×10^{-4}	-7.497×10^{-4}	-8.005×10^{-4}	-5.226×10^{-4}	-2.787×10^{-4}	-1.834×10^{-4}
0	-2.14×10^{-6}	-8.577×10^{-5}	-3.6×10^{-4}	-5.422×10^{-4}	-2.499×10^{-4}	-1.494×10^{-4}	-1.067×10^{-4}
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Mean squared error ρ

100	-3.222×10^{2}	-2.463×10^{4}	-8.216×10^{4}	-1.254×10^{5}	-1.628×10^{5}	-1.879×10^{5}	-2.016×10^{5}
10	-2.342	-3.341×10^{1}	-3.454×10^{2}	-4.314×10^{3}	-8.484×10^{3}	-1.595×10^{4}	-6.127×10^{4}
1	-3.31×10^{-2}	-2.413×10^{-1}	-5.613×10^{-1}	-2.177	-3.152×10^{1}	-1.126×10^{2}	-9.343×10^{2}
0	0.	-9.5×10^{-3}	-2.43×10^{-2}	-3.04×10^{-2}	-3.95×10^{-2}	-8.56×10^{-2}	-1.101×10^{1}
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Table SI D_6 – Comparison of the (marginal) accuracy of $\hat{\psi}$ (left column) and $\hat{\rho}$ (right column) for the maximum likelihood-based estimate (i.e., the reference) and the L1-distance-based estimate. Each cell shows the difference of the absolute mean difference ($|MD_{MLE}| - |MD_{L1}|$; first row), the difference of the mean absolute difference ($MAD_{MLE} - MAD_{L1}$; third row) and the difference of the mean squared error ($MSE_{MLE} - MSE_{L1}$; third row) each calculated over 10,000 data sets assuming linkage equilibrium, $k = 100$, and $s = 10{,}000$. Colors within each sub-table range from light yellow to dark red and scale between the minimal and the maximal absolute value to aid interpretation.

Mean difference ψ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	-4.49×10^{-4}	-8.54×10^{-4}	-3.07×10^{-4}	-3.539×10^{-3}	-2.962×10^{-3}	-1.851×10^{-3}
10	0.	-1.499×10^{-3}	-1.879×10^{-3}	-1.78×10^{-4}	-1.574×10^{-3}	-1.369×10^{-3}	-3.14×10^{-4}
1	0.	-2.74×10^{-4}	-3.83×10^{-4}	-6.26×10^{-4}	-1.061×10^{-3}	-9.23×10^{-4}	-1.145×10^{-3}
0	0.	-8.5×10^{-5}	-1.46×10^{-4}	-5.76×10^{-4}	-3.92×10^{-4}	-8.25×10^{-4}	-5.29×10^{-4}

Mean absolute difference ψ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	-1.187×10^{-3}	-2.464×10^{-3}	-1.747×10^{-3}	-2.495×10^{-3}	-2.14×10^{-3}	-1.693×10^{-3}
10	0.	-2.339×10^{-3}	-2.649×10^{-3}	-2.568×10^{-3}	-2.448×10^{-3}	-1.779×10^{-3}	-1.214×10^{-3}
1	0.	-6.2×10^{-4}	-1.527×10^{-3}	-4.418×10^{-3}	-7.551×10^{-3}	-5.447×10^{-3}	-2.625×10^{-3}
0	0.	-1.25×10^{-4}	-7.96×10^{-4}	-3.192×10^{-3}	-6.172×10^{-3}	-5.719×10^{-3}	-3.119×10^{-3}

Mean squared error ψ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	-1.775×10^{-5}	-7.334×10^{-5}	-4.629×10^{-5}	-6.731×10^{-5}	-5.28×10^{-5}	-3.119×10^{-5}
10	0.	-3.229×10^{-5}	-7.087×10^{-5}	-5.848×10^{-5}	-4.472×10^{-5}	-2.861×10^{-5}	-1.47×10^{-5}
1	0.	-1.094×10^{-5}	-3.303×10^{-5}	-9.844×10^{-5}	-1.618×10^{-4}	-8.273×10^{-5}	-2.851×10^{-5}
0	0.	-1.25×10^{-5}	-9.68×10^{-6}	-5.656×10^{-5}	-1.161×10^{-4}	-8.205×10^{-5}	-3.279×10^{-5}

Mean difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	-5.55×10^{-2}	-3.308	-3.131	3.784×10^{-1}	5.127	5.522	6.389
10	-2.18×10^{-2}	-5.28×10^{-1}	-3.566×10^{-1}	-8.07×10^{-2}	9.53×10^{-2}	1.103×10^{-1}	9.19×10^{-2}
1	0.	-2.25×10^{-2}	-1.18×10^{-2}	-7.1×10^{-3}	-4.3×10^{-3}	-9.4×10^{-3}	-9.3×10^{-3}
0	0.	0.	0.	0.	0.	0.	-1.8×10^{-3}

Mean absolute difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	-5.995×10^{-1}	-3.491	-3.436	-4.202×10^{-1}	1.98	3.216	4.165
10	-1.41×10^{-1}	-5.474×10^{-1}	-2.44×10^{-1}	-2.071×10^{-1}	-8.23×10^{-2}	-3.79×10^{-2}	1.99×10^{-2}
1	0.	-2.25×10^{-2}	-1.18×10^{-2}	-7.1×10^{-3}	-4.3×10^{-3}	-9.4×10^{-3}	-9.7×10^{-3}
0	0.	0.	0.	0.	0.	0.	-1.8×10^{-3}

Mean squared error ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	-8.662	-3.304×10^{2}	-5.77×10^{2}	-8.993×10^{1}	8.06×10^{2}	1.306×10^{3}	2.989×10^{3}
10	-1.71×10^{-1}	-2.73	-2.295	-1.171	-5.215×10^{-1}	4.01×10^{-2}	2.937×10^{-1}
1	0.	-2.31×10^{-2}	-1.2×10^{-2}	-7.1×10^{-3}	-4.3×10^{-3}	-9.6×10^{-3}	-1.51×10^{-2}
0	0.	0.	0.	0.	0.	0.	-1.8×10^{-3}

Table SI D_7 – Comparison of the (marginal) accuracy of $\hat{\psi}$ (left column) and $\hat{\rho}$ (right column) for the maximum likelihood-based estimate (i.e., the reference) and the L2-distance-based estimate. Each cell shows the difference of the absolute mean difference ($|MD_{MLE}| - |MD_{L2}|$; first row), the difference of the mean absolute difference ($MAD_{MLE} - MAD_{L2}$; third row) and the difference of the mean squared error ($MSE_{MLE} - MSE_{L2}$; third row) each calculated over 10,000 data sets assuming linkage equilibrium, $k = 100$, and $s = 10{,}000$. Colors within each sub-table range from light yellow to dark red and scale between the minimal and the maximal absolute value to aid interpretation.

Mean difference ψ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	$9.\times10^{-6}$	-1.22×10^{-4}	-3.9×10^{-5}	-2.3×10^{-5}	-4.4×10^{-5}	$2.\times10^{-6}$
10	0.	-5.92×10^{-4}	-1.496×10^{-3}	-6.25×10^{-4}	-1.21×10^{-4}	-1.67×10^{-4}	-1.39×10^{-4}
1	0.	-3.528×10^{-3}	-6.88×10^{-4}	-1.221×10^{-3}	-2.446×10^{-3}	-4.373×10^{-3}	-3.471×10^{-3}
0	0.	-1.1×10^{-5}	1.2×10^{-5}	-3.22×10^{-4}	-2.5×10^{-5}	-1.47×10^{-4}	-2.47×10^{-4}

Mean absolute difference ψ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	-3.51×10^{-4}	-5.5×10^{-4}	-8.43×10^{-4}	-1.367×10^{-3}	-1.558×10^{-3}	-1.136×10^{-3}
10	0.	-2.312×10^{-3}	-1.65×10^{-3}	-1.911×10^{-3}	-1.533×10^{-3}	-1.287×10^{-3}	-8.91×10^{-4}
1	0.	-3.534×10^{-3}	-1.762×10^{-3}	-5.133×10^{-3}	-9.1×10^{-3}	-9.775×10^{-3}	-4.943×10^{-3}
0	0.	-1.63×10^{-4}	-9.02×10^{-4}	-3.512×10^{-3}	-7.175×10^{-3}	-8.609×10^{-3}	-5.393×10^{-3}

Mean squared error ψ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	-4.77×10^{-5}	-1.5×10^{-5}	-2.063×10^{-5}	-3.135×10^{-5}	-3.174×10^{-5}	-2.028×10^{-5}
10	0.	-3.204×10^{-5}	-3.986×10^{-5}	-4.203×10^{-5}	-2.649×10^{-5}	-1.875×10^{-5}	-1.063×10^{-5}
1	0.	-8.106×10^{-5}	-4.354×10^{-5}	-1.234×10^{-4}	-2.206×10^{-4}	-2.279×10^{-4}	-7.141×10^{-5}
0	0.	-1.63×10^{-5}	-1.122×10^{-5}	-6.528×10^{-5}	-1.452×10^{-4}	-1.709×10^{-4}	-7.129×10^{-5}

Mean difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	-6.06×10^{-2}	-3.865×10^{-1}	-5.669×10^{-1}	5.49×10^{-2}	-4.865×10^{-1}	-3.713×10^{-1}	-5.043×10^{-1}
10	-2.9×10^{-3}	-2.719×10^{-1}	-2.394×10^{-1}	-9.55×10^{-2}	-4.25×10^{-2}	-3.17×10^{-2}	-5.84×10^{-2}
1	0.	-1.891×10^{-1}	-1.59×10^{-2}	-1.17×10^{-2}	-9.7×10^{-3}	-8.7×10^{-3}	-1.27×10^{-2}
0	0.	0.	0.	$-1.\times10^{-4}$	0.	0.	-1.2×10^{-3}

Mean absolute difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	-7.066×10^{-1}	-1.183	-9.345×10^{-1}	-4.367×10^{-1}	-7.791×10^{-1}	-5.373×10^{-1}	-7.103×10^{-1}
10	-1.435×10^{-1}	-4.847×10^{-1}	-6.34×10^{-2}	-1.753×10^{-1}	-8.75×10^{-2}	-7.07×10^{-2}	-9.34×10^{-2}
1	0.	-1.897×10^{-1}	-1.59×10^{-2}	-1.17×10^{-2}	-9.7×10^{-3}	-8.7×10^{-3}	-1.75×10^{-2}
0	0.	0.	0.	$-1.\times10^{-4}$	0.	0.	-1.2×10^{-3}

Mean squared error ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	-1.027×10^{1}	-8.835×10^{1}	-1.442×10^{2}	-8.158×10^{1}	-2.318×10^{2}	-3.466×10^{2}	-6.137×10^{2}
10	-1.739×10^{-1}	-2.154	-9.372×10^{-1}	-8.903×10^{-1}	-7.511×10^{-1}	-7.469×10^{-1}	-2.401
1	0.	-2.543×10^{-1}	-2.33×10^{-2}	-1.17×10^{-2}	-9.7×10^{-3}	-9.1×10^{-3}	-2.55×10^{-2}
0	0.	0.	0.	$-1.\times10^{-4}$	0.	0.	-1.2×10^{-3}

Table SI D_8 – Comparison of the (marginal) accuracy of $\hat{\psi}$ (left column) and $\hat{\rho}$ (right column) for the maximum likelihood estimate based on the full SFS (i.e., the reference) and the lumped SFS where the $i = 5$ entry in the SFS contains the aggregate of the higher frequency classes. Each cell shows the difference of the absolute mean difference ($|MD_{i=0}| - |MD_{i=5}|$; first row), the difference of the mean absolute difference ($MAD_{i=0} - MAD_{i=5}$; third row) and the difference of the mean squared error ($MSE_{i=0} - MSE_{i=5}$; third row) each calculated over 10,000 data sets assuming linkage equilibrium, $k = 100$, and $s = 10{,}000$. Colors within each sub-table range from light yellow to dark red and scale between the minimal and the maximal absolute value to aid interpretation.

Mean difference ψ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	-7.051×10^{-3}	-4.542×10^{-2}	-5.894×10^{-2}	-5.722×10^{-2}	-4.11×10^{-2}	-5.664×10^{-2}
10	0.	-4.245×10^{-3}	-4.103×10^{-2}	-3.082×10^{-2}	-1.789×10^{-2}	-1.74×10^{-3}	-1.894×10^{-2}
1	0.	-1.838×10^{-3}	-1.561×10^{-2}	-1.23×10^{-1}	-1.48×10^{-2}	-5.409×10^{-2}	-2.431×10^{-3}
0	0.	3.4×10^{-5}	-2.054×10^{-2}	-6.181×10^{-2}	-8.636×10^{-2}	-8.615×10^{-2}	-1.462×10^{-2}

Mean absolute difference ψ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	-1.793×10^{-3}	-9.782×10^{-2}	-1.439×10^{-1}	-1.565×10^{-1}	-1.314×10^{-1}	-7.858×10^{-2}
10	0.	-1.444×10^{-2}	-9.935×10^{-2}	-1.425×10^{-1}	-1.537×10^{-1}	-1.377×10^{-1}	-5.074×10^{-2}
1	0.	-9.936×10^{-3}	-7.792×10^{-2}	-1.235×10^{-1}	-1.377×10^{-1}	-1.04×10^{-1}	-2.877×10^{-2}
0	0.	-2.72×10^{-4}	-2.063×10^{-2}	-6.289×10^{-2}	-9.012×10^{-2}	-9.192×10^{-2}	-2.087×10^{-2}

Mean squared error ψ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	0.	-1.073×10^{-3}	-2.168×10^{-2}	-3.693×10^{-2}	-3.529×10^{-2}	-2.339×10^{-2}	-7.372×10^{-3}
10	0.	-6.26×10^{-4}	-2.037×10^{-2}	-3.271×10^{-2}	-3.274×10^{-2}	-2.948×10^{-2}	-5.431×10^{-3}
1	0.	-4.721×10^{-4}	-8.994×10^{-3}	-2.12×10^{-2}	-2.777×10^{-2}	-2.282×10^{-2}	-2.336×10^{-3}
0	0.	-6.44×10^{-5}	-2.017×10^{-3}	-1.162×10^{-2}	-2.236×10^{-2}	-2.459×10^{-2}	-2.296×10^{-3}

Mean difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	-1.6×10^{-3}	-1.437×10^{1}	-1.527×10^{2}	-2.103×10^{2}	-2.057×10^{2}	-1.657×10^{2}	6.263×10^{1}
10	-1.4×10^{-3}	-6.09×10^{-1}	-1.177×10^{1}	-4.799×10^{1}	-7.014×10^{1}	-8.445×10^{1}	-1.433×10^{1}
1	0.	-1.139×10^{-1}	-1.163	-4.029	-7.027	-9.382	-5.184×10^{-1}
0	0.	-2.9×10^{-3}	-2.584×10^{-1}	-1.008	-1.784	-2.024	-1.873×10^{-1}

Mean absolute difference ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	-3.64×10^{-2}	-4.183×10^{1}	-2.056×10^{2}	-2.656×10^{2}	-2.666×10^{2}	-2.306×10^{2}	-4.649×10^{1}
10	-4.32×10^{-2}	-2.98	-1.743×10^{1}	-5.386×10^{1}	-7.652×10^{1}	-9.035×10^{1}	-2.012×10^{1}
1	0.	-2.955×10^{-1}	-1.93	-4.808	-7.785	-9.757	-8.96×10^{-1}
0	0.	-2.9×10^{-3}	-2.584×10^{-1}	-1.008	-1.784	-2.024	-1.873×10^{-1}

Mean squared error ρ

ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9
100	-4.7×10^{-1}	-7.194×10^{3}	-1.351×10^{5}	-2.055×10^{5}	-2.193×10^{5}	-1.997×10^{5}	-2.057×10^{4}
10	-4.74×10^{-2}	-2.637×10^{1}	-1.286×10^{3}	-2.251×10^{4}	-4.713×10^{4}	-6.338×10^{4}	-1.272×10^{4}
1	0.	-3.241×10^{-1}	-1.048×10^{1}	-1.27×10^{2}	-8.785×10^{2}	-2.588×10^{3}	-2.495×10^{1}
0	0.	-2.9×10^{-3}	-3.688×10^{-1}	-5.518	-2.616×10^{1}	-7.375×10^{1}	-5.199×10^{-1}

Table SI D_9 – Comparison of the (marginal) accuracy of $\hat{\psi}$ (left column) and $\hat{\rho}$ (right column) for the maximum likelihood estimate based on the full SFS (i.e., the reference) and the lumped SFS where the $i = 15$ entry in the SFS contains the aggregate of the higher frequency classes. Each cell shows the difference of the absolute mean difference ($|MD_{i=0}| - |MD_{i=15}|$; first row), the difference of the mean absolute difference ($MAD_{i=0} - MAD_{i=15}$; third row) and the difference of the mean squared error ($MSE_{i=0} - MSE_{i=15}$; third row) each calculated over 10,000 data sets assuming linkage equilibrium, $k = 100$, and $s = 10{,}000$. Colors within each sub-table range from light yellow to dark red and scale between the minimal and the maximal absolute value to aid interpretation.

Mean difference ψ

100	0.	-1.24×10^{-4}	-7.469×10^{-3}	-1.886×10^{-2}	-4.001×10^{-2}	-6.134×10^{-2}	-7.884×10^{-3}
10	0.	-8.2×10^{-5}	-4.218×10^{-3}	-7.432×10^{-3}	-9.452×10^{-3}	-3.341×10^{-2}	-1.875×10^{-2}
1	0.	$-1. \times 10^{-6}$	-2.331×10^{-3}	-3.698×10^{-3}	-2.08×10^{-2}	-7.634×10^{-3}	-2.643×10^{-2}
0	0.	-1.2×10^{-5}	$-3. \times 10^{-5}$	-5.269×10^{-3}	-1.478×10^{-2}	-2.01×10^{-3}	-2.145×10^{-2}
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Mean difference ρ

100	1.3×10^{-3}	1.823×10^{-1}	-1.608×10^{1}	-8.871×10^{1}	-9.488×10^{1}	-4.772	-3.127×10^{1}
10	3.7×10^{-3}	7.3×10^{-3}	-5.424×10^{-1}	-3.989	-6.303	-1.987	-3.338×10^{1}
1	0.	-4.9×10^{-3}	-1.28×10^{-1}	-3.258×10^{-1}	-5.282×10^{-1}	-2.082×10^{-1}	-2.11
0	0.	0.	$-9. \times 10^{-4}$	-4.41×10^{-2}	-9.8×10^{-2}	-1.68×10^{-2}	-3.897×10^{-1}
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Mean absolute difference ψ

100	0.	-1.98×10^{-4}	-2.494×10^{-2}	-7.218×10^{-2}	-1.015×10^{-1}	-9.211×10^{-2}	-4.606×10^{-2}
10	0.	-2.68×10^{-4}	-1.914×10^{-2}	-5.761×10^{-2}	-6.927×10^{-2}	-6.442×10^{-2}	-4.582×10^{-2}
1	0.	-1.21×10^{-4}	-9.563×10^{-3}	-3.663×10^{-2}	-3.424×10^{-2}	-2.121×10^{-2}	-3.588×10^{-2}
0	0.	-3.2×10^{-5}	-8.94×10^{-4}	-8.647×10^{-3}	-2.088×10^{-2}	-9.404×10^{-3}	-2.617×10^{-2}
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Mean absolute difference ρ

100	3.1×10^{-3}	-5.161×10^{-1}	-3.883×10^{1}	-1.242×10^{2}	-1.412×10^{2}	-6.453×10^{1}	-7.039×10^{1}
10	-4.3×10^{-3}	-7.35×10^{-2}	-2.399	-7.257	-9.38	-6.204	-3.522×10^{1}
1	0.	-4.9×10^{-3}	-1.878×10^{-1}	-5.328×10^{-1}	-5.812×10^{-1}	-2.594×10^{-1}	-2.208
0	0.	0.	$-9. \times 10^{-4}$	-4.41×10^{-2}	-9.8×10^{-2}	-1.68×10^{-2}	-3.897×10^{-1}
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Mean squared error ψ

100	0.	-2.76×10^{-5}	-1.79×10^{-3}	-1.084×10^{-2}	-1.909×10^{-2}	-1.586×10^{-2}	-4.594×10^{-3}
10	0.	-3.1×10^{-6}	-1.059×10^{-3}	-6.13×10^{-3}	-1.026×10^{-2}	-9.869×10^{-3}	-6.722×10^{-3}
1	0.	-2.07×10^{-6}	-6.558×10^{-4}	-3.901×10^{-3}	-3.299×10^{-3}	-1.531×10^{-3}	-5.648×10^{-3}
0	0.	-3.2×10^{-7}	-1.572×10^{-5}	-7.607×10^{-4}	-2.417×10^{-3}	-4.666×10^{-4}	-4.628×10^{-3}
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Mean squared error ρ

100	3.37×10^{-2}	-3.355×10^{1}	-8.107×10^{3}	-7.226×10^{4}	-1.032×10^{5}	-4.275×10^{4}	-5.61×10^{4}
10	-5.3×10^{-3}	-3.027×10^{-1}	-2.292×10^{1}	-2.217×10^{2}	-7.406×10^{2}	-1.039×10^{3}	-2.002×10^{4}
1	0.	-4.9×10^{-3}	-1.97×10^{-1}	-9.284×10^{-1}	-2.034	-1.025	-1.062×10^{2}
0	0.	0.	$-9. \times 10^{-4}$	-4.41×10^{-2}	-9.98×10^{-2}	-1.8×10^{-2}	-2.159
ρ / ψ	0	0.15	0.3	0.45	0.6	0.75	0.9

Supporting Files

File E1: SFS-Sardine-NiwaEtAl2016.txt