

EXACT PHYLODYNAMICS VIA STRUCTURED MARKOV GENEALOGY PROCESSES

AARON A. KING, QIANYING LIN, AND EDWARD L. IONIDES

ABSTRACT.

1. Introduction.

Problem of phylodynamics. Factorization of problem into two subproblems.

Relation to previous work. Existing methods (Volz et al., 2009; Stadler, 2010). Large-population, small sample-size approximations.

Extension of previous results (King et al., 2022). Broader class of state-spaces. Accommodating discrete structure.

Classes of Markov processes. Utility and flexibility of Markov assumptions.

Population process induces Markov history and genealogy processes. Using these, we derive equations for the likelihood of a genealogy conditional on the history. We then integrate out the history to obtain nonlinear filtering equations, the solution of which yields the likelihood. These readily lend themselves to a family of sequential Monte Carlo algorithms for computing the likelihood. We demonstrate with several examples.

2. Population processes.

Motivating examples: compartmental models. Wide variety of models. Linear chain trick. Migration, superspreading, competition between strains.

Figure: example compartmental model with multiple demes.

Mathematical formalism. Definition of state space. Jumps and jump rates. Demes.

Kolmogorov equations.

Date: July 3, 2023.

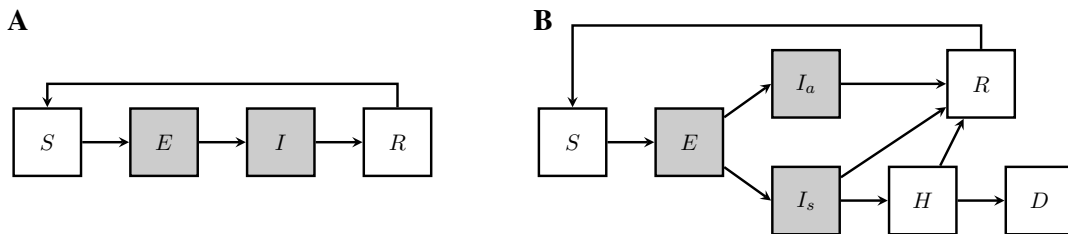


FIGURE 1. Examples of compartmental models. Demes are shaded.

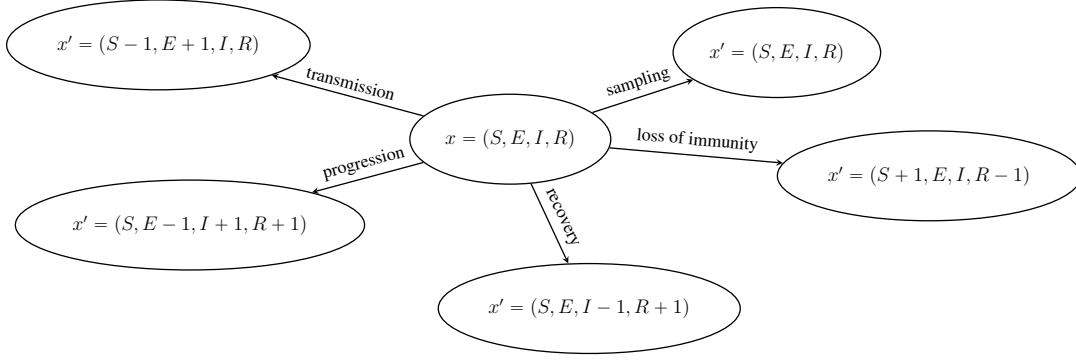


FIGURE 2. Markov state transition diagram for an SEIR model. The state is characterized by four numbers, S , E , I , and R . From a given state x , there are four possible jumps $x \mapsto x'$. From the point of view of the induced genealogy process, transmission is of birth type, progression is of migration type, and recovery is of death type, while loss of immunity is of neutral type. In this example, when a sampling event occurs, the state does not change.

2.1. Examples.

SEIRS model.

SIIR model.

Linear birth-death model.

Moran model and the Kingman coalescent.

3. History and genealogy processes.

3.1. Mathematical preliminaries.

State space. Definition of Ω as space of càdlàg functions of time. Define everything in terms of càdlàg functions rather than sequences. Then the topology comes directly from the Skorokhod metric. Explain we use a base measure coming from unit-rate Poisson processes.

Restrictions. Metric, Borel σ -algebra, base measure.

Probability density. Master process.

Markov jump processes. We will assume that our population process is a non-explosive Markov jump process \mathcal{X}_t parameterized by time $t \in \mathbb{R}_+$, and taking values in some space \mathbb{X} . In earlier work (King et al., 2022), we limited ourselves to the case $\mathbb{X} = \mathbb{Z}^d$, but we dispense with that restriction here. The population process is completely specified by its initial-state distribution and its generator. In particular, we suppose that

$$\mathbb{P}[\mathcal{X}_0 = x] = p_0(x), \quad (1)$$

for some choice of initial distribution, p_0 . We suppose there are countably many types of jumps and denote the set of jump types by \mathbb{U} . For each $u \in \mathbb{U}$, there is a corresponding jump rate or hazard α_u . We impose the further requirement that $\sum_{u, x'} \alpha_u(t, x, x') < \infty$ for all t, x . We will always use a version of \mathcal{X}_t in which the sample paths are right-continuous with left limits (càdlàg).

The above may be compactly summarized by means of the corresponding Kolmogorov equations, as follows. Given $f \in L^\infty(\mathbb{X})$, let $F(s, x) := \mathbb{E}[f(\mathcal{X}_t) \mid \mathcal{X}_s = x]$. It then follows that, for $s < t$ and $x \in \mathbb{X}$,

$$\frac{\partial F}{\partial s}(s, x) = - \int \alpha(s, x, x') [F(s, x') - F(s, x)] dx', \quad F(t, x) = f(x). \quad (2)$$

Eq. 2 is the so-called Kolmogorov backward equation. Its adjoint is the Kolmogorov forward equation (KFE, also called the *master equation* in this context), which states that for $t \geq 0$, $x \in \mathbb{X}$,

$$\frac{\partial w}{\partial t}(t, x) = \int w(t, x') \alpha(t, x', x) dx' - \int w(t, x) \alpha(t, x, x') dx', \quad w(0, x) = p_0(x). \quad (3)$$

If $w(t, x)$ satisfies Eqs. 3, then $w(t, x) = \mathbb{P}[\mathcal{X}_t = x]$.

Definitions. Our goal in this paper is to introduce a family of Markov processes induced by the jump process just described (??). While population processes of the kind described above can be constructed in a variety of time-honored ways (e.g., Andersen et al., 1993; Kallenberg, 1997), these classical approaches are not entirely sufficient for the tree-valued Markov processes we will erect on top of the population process. We therefore explicitly construct the probability space that underlies the stochastic processes we subsequently describe. Unavoidably, this leads to technicalities that are necessary for the firm establishment of the properties of these processes but that may distract from the overarching goals. Readers willing to stipulate these properties may skim the remainder of this section, in which we formally construct the probability space and make some definitions that will be needed in the sequel.

Let us define a *jump* to be an ordered quadruple $(t, u, x, n) \in \mathbb{R}_+ \times \mathbb{U} \times \mathbb{X} \times \mathbb{Z}_+$. We refer to t as the *time* of the jump; u is the *type* of the jump; x is the state achieved at the end of the jump; n is an *auxiliary number* whose use will be made clear below. A *jump sequence* is a countable sequence of jumps at increasing times. That is,

$$\omega = (t_k, u_k, x_k, n_k)_{k=0}^K$$

is a jump sequence if and only if $K \in \mathbb{Z}_+ \cup \{\infty\}$, $u_k \in \mathbb{U}$, $x_k \in \mathbb{X}$, $n_k \in \mathbb{Z}_+$ for all k , and $0 = t_0 < t_1 < t_2 < \dots$. We will take our sample space, Ω , to be the set of all jump sequences. For $\omega \in \Omega$ as above, we write

$$T_k(\omega) := t_k, \quad U_k(\omega) := u_k, \quad X_k(\omega) := x_k, \quad N_k(\omega) := n_k, \quad K(\omega) := K \quad (4)$$

and also

$$\begin{aligned} T(\omega) &:= (T_k(\omega))_{k=0}^{K(\omega)}, & U(\omega) &:= (U_k(\omega))_{k=0}^{K(\omega)}, \\ X(\omega) &:= (X_k(\omega))_{k=0}^{K(\omega)}, & N(\omega) &:= (N_k(\omega))_{k=0}^{K(\omega)}. \end{aligned} \quad (5)$$

We will denote by $\tilde{\Omega}$ the set of all finite jump sequences, i.e., $\tilde{\Omega} := \{\omega \in \Omega \mid K(\omega) < \infty\}$.

Note that every element of Ω corresponds to a unique sample path of \mathcal{X}_t . In particular,

$$\mathcal{X}_t(\omega) = \sum_{k=0}^{K(\omega)} X_k(\omega) \mathbb{1}_{[T_k(\omega), \infty)}(t). \quad (6)$$

Note that, for every $\omega \in \Omega$, the function $\mathcal{X}(\omega) : \mathbb{R}_+ \rightarrow \mathbb{X}$ is thus càdlàg. It will be useful to define $\tilde{\mathcal{X}}$ to be the left-limit function of \mathcal{X} :

$$\tilde{\mathcal{X}}(\omega, t) := \begin{cases} \lim_{s \uparrow t} \mathcal{X}(\omega, s), & t > 0, \\ \mathcal{X}(0), & t = 0. \end{cases}$$

The sample space Ω has a natural partial order. We write $\omega \leq \omega'$ if ω' is an extension of ω ; that is, if $K(\omega) \leq K(\omega')$, and $(T_k(\omega), U_k(\omega), X_k(\omega), N_k(\omega)) = (T_k(\omega'), U_k(\omega'), X_k(\omega'), N_k(\omega'))$ for

$k = 0, \dots, K(\omega)$. For $\omega \in \Omega$, the set $\{\omega' \in \Omega \mid \omega' \preceq \omega\}$ is totally ordered. Moreover, for each $\omega \in \mathring{\Omega}$, there is a unique *predecessor*, $\tilde{\omega}$, such that $\tilde{\omega} \preceq \omega$, $\tilde{\omega} \neq \omega$, and, for all ω' , $\tilde{\omega} \preceq \omega' \preceq \omega$ implies that either $\omega' = \tilde{\omega}$ or $\omega' = \omega$.

In order to define probabilistic *events*, it is necessary to define the σ -algebra of measurable subsets of Ω . We first impose the condition that \mathbb{X} be a complete, separable, measurable metric space. We then follow King et al. (2022), extending the Skorokhod metric (Kallenberg, 1997; Ethier & Kurtz, 2009) to make Ω a complete, separable metric space (*i.e.*, a Polish space) and take our event space, \mathcal{B} , to be its Borel σ -algebra.

We now turn to the probability measure, \mathbb{P} , on $\Omega \cong T(\Omega) \times U(\Omega) \times X(\Omega) \times N(\Omega)$. We will specify this by giving its density. In particular, for each $u \in \mathbb{U}$, $x, x' \in \mathbb{X}$, we suppose $\nu_{u,x,x'}$ is a given probability measure on \mathbb{Z}_+ ; these will take specific forms below. For $t \in \mathbb{R}_+$ and $\omega \in \Omega|_t$, define the probability density function

$$P_t(\omega) := p_0(\mathcal{X}_0) \prod_{k=1}^K \left(\alpha_{U_k}(T_k, \tilde{\mathcal{X}}_{T_k}, \mathcal{X}_{T_k}) \nu_{U_k, \tilde{\mathcal{X}}_{T_k}, \mathcal{X}_{T_k}}(N_k) \right) \exp \left(- \int_0^t \sum_{ux'} \alpha_u(s, \mathcal{X}_s, x') ds \right), \quad (7)$$

the latter sum being taken over $u \in \mathbb{U}$, $x' \in \mathbb{X}$. Here, for the sake of readability, we have suppressed the dependence of the random variables \mathcal{X}_t , K , T_k , U_k and N_k on ω .

History process. Next among our constellation of related processes (??) is the *history process*, which encapsulates the entire history of \mathcal{X} up to time t . Specifically, for $\omega \in \Omega$, we define

$$\mathcal{H}_t(\omega) := \left(t, (T_k(\omega), U_k(\omega), X_k(\omega))_{k=0}^{K(\omega|t)} \right),$$

where T_k , U_k , X_k , and K are as in Eq. 4. Thus \mathcal{H}_t contains exactly those elements of ω that are relevant to the history of the population process \mathcal{X}_t . It is trivial to verify that \mathcal{H}_t is Markov and to compute its probability density. In particular, given any history $h_t = \left(t, (t_k, u_k, x_k)_{j=0}^K \right)$, the marginal density at h_t is readily seen to be

$$\begin{aligned} P_{\mathcal{H}_t}(h_t) := & p_0(x_0) \prod_{k=1}^K \left\{ \alpha_{u_k}(t_k, x_{k-1}, x_k) \exp \left(- \int_{t_{k-1}}^{t_k} \sum_{ux'} \alpha_u(s, x_{k-1}, x') ds \right) \right\} \\ & \times \exp \left(- \int_{t_K}^t \sum_{ux'} \alpha_u(s, x_K, x') ds \right). \end{aligned}$$

Inventory process. Our main interest is in tracking the genealogical relationships among the lineages in our demes. Accordingly, we will assign each individual lineage a unique number. let $\mathcal{I}_t : \mathbb{R}_+ \rightarrow \mathbb{D}^{\mathbb{Z}_+}$ be such that $\mathcal{I}_t(n) = i$ if and only if lineage n is in deme i at time t . If t is before the birth or after the death of n , $\mathcal{I}_t(n) = \emptyset$. Note that $n_i(\mathcal{X}_t) = |\{n \mid \mathcal{I}_t(n) = i\}|$ for all $t \in \mathbb{R}_+$ and $i \in \mathbb{D}$.

Genealogy process.

Pruning.

Filter equations.

Definition 1. Suppose X_t is a Markov process with Kolmogorov forward equation (KFE)

$$\frac{\partial w}{\partial t} = \int w(t, x') \alpha(t, x', x) dx' - \int w(t, x) \alpha(t, x, x') dx', \quad w(0, x) = p_0(x). \quad (8)$$

Suppose that $B(x, x') > 0$ and $\lambda(t, x) > 0$ are given functions, and $t_1 < t_2 < \dots < t_K$ is a given sequence of times. We say that the equations

$$\frac{\partial w}{\partial t} = \int w(t, x') \alpha(t, x', x) B(x', x) dx' - \int w(t, x) \alpha(t, x, x') dx' - \lambda(t, x) w(t, x), \quad (9)$$

$$w(t_k, x) = \int \tilde{w}(t_k, x') \frac{\alpha(t, x', x)}{\mu} B(x', x) dx', \quad k = 1, \dots, K, \quad (10)$$

$$w(0, x) = p_0(x). \quad (11)$$

are the *filter equations* for X_t with *boost* B , *decay* λ , *initial density* p_0 , and *observation times* t_k .

The useful fact about filter equations is that they afford a means of computing the likelihood of a given sequence of events. This is facilitated by the following

Lemma 1. *Eq. 9 is satisfied by $w(t, x) = \int v U(t, x, v) dv$, where U satisfies the KFE*

$$\begin{aligned} \frac{\partial U}{\partial t} = & \int U(t, x', v') \alpha(t, x', x) \delta(v, B(x', x) v') dx' dv' \\ & - \int U(t, x, v) \alpha(t, x, x') \delta(v', B(x, x') v) dx' dv' + \frac{\partial}{\partial v} [\lambda(t, x) v U(t, x, v)]. \end{aligned} \quad (12)$$

Here, $\delta(v, v')$ is the familiar Dirac δ .

Proof.

$$\begin{aligned} \frac{\partial w}{\partial t} &= \int v \frac{\partial U}{\partial t}(t, x, v) dv \\ &= \int v U(t, x', v') \alpha(t, x', x) \delta(v, B(x', x) v') dv dx' dv' \\ &\quad - \int v U(t, x, v) \alpha(t, x, x') \delta(v', B(x, x') v) dv dx' dv' \\ &\quad + \int v \frac{\partial}{\partial v} [\lambda(t, x) v U(t, x, v)] dv. \end{aligned}$$

Evaluating the first integral with respect to v , the second with respect to v' , and the third by parts, we obtain

$$\begin{aligned} \frac{\partial w}{\partial t} &= \int v' U(t, x', v') \alpha(t, x', x) B(x', x) dv' dx' - \int v U(t, x, v) \alpha(t, x, x') dv dx' \\ &\quad - \lambda(t, x) \int v U(t, x, v) dv, \end{aligned}$$

which is simplified to obtain Eq. 9. □

We recognize in Eq. 12 the KFE of a certain process (X_t, V_t) . In particular, X_t is a jump process with KF Eq. 8. The V_t process has jumps wherever X_t does, such that when X_t jumps $x \mapsto x'$, V_t jumps by a factor of $B(x, x')$. Between jumps, V_t decays deterministically and exponentially at rate $\lambda(t, x)$. If we view V_t as a weight, then Lemma 1 says that the weighted average of X_t evolves according to Eq. 9.

Lineages, event-counter, deme-residence. Let $L = L(\mathcal{G})$ denote the finite set of all samples represented in genealogy \mathcal{G} . Let \ll be any ordering of L that is compatible with ancestry. That is, if $j, j' \in L$ are such that j is ancestral to j' , then we must have $j \ll j'$. Using this ordering, we can uniquely associate each point on a genealogical tree with the least of those lineages that descend from that point. In particular, any lineage $j \in L$, corresponding to a sample taken at time t_j^s , can be traced backward from

node to node until either it coalesces with some lesser lineage at some time $t_j^o > 0$ or a root is reached (in which case, we define $t_j^o = 0$). Each node encountered along the way represents a genealogical event from which j emerges. Moreover, at each time $t \in [t_j^o, t_j^s]$, lineage j is in precisely one of the demes D . However, for $t \notin [t_j^o, t_j^s]$, lineage j does not exist. To express this, we say that lineage j is in the *undeme*, which we denote using the symbol $\bar{\emptyset}$. We define $\bar{\mathbb{D}} := \mathbb{D} \cup \{\bar{\emptyset}\}$.

It will be useful to define a function that captures all the relevant features of a pruned genealogy. Accordingly, let $\mathbb{Y} = (\mathbb{Z}_+ \times \bar{\mathbb{D}} \times \mathbb{L})^{\mathbb{L}}$ and define $y : \mathbb{R}_+ \rightarrow \mathbb{Y}$ so that, for $j \in \mathbb{L}$ and $t \in \mathbb{R}_+$:

- (a) $\text{ct}(y_j(t)) \in \mathbb{Z}_+$ is a counting process which increases by 1 at each event along lineage j .
- (b) $\text{deme}(y_j(t)) \in \bar{\mathbb{D}}$ indicates in which deme lineage j lies at time t . In particular $\text{deme}(y_j(t)) = \bar{\emptyset}$ if $t \notin [t_j^o, t_j^s]$.
- (c) $\text{anc}(y_j(t)) \in \mathbb{L}$ indicates the lineage in which the ancestors of lineage j are found. In particular, $\text{anc}(y_j(t)) = j$ for $t \geq t_j^o$, but $\text{anc}(y_j(t))$ is well defined for all j and t .

One can verify that y is càdlàg. Define $\tilde{y}(t) = \lim_{s \uparrow t} y(s)$.

To visualize these functions, one can make a correspondence between demes and colors. Then a pruned genealogy is visualized as a tree with colored branches. Knowing the function $\text{deme}(y)$ is equivalent to knowing the coloring, while $\text{ct}(y)$ determines the locations of events in the genealogy and $\text{anc}(y)$ determines the topology. Note in particular that $y_j(t) \neq \tilde{y}_j(t)$ if and only if t is the time of an event from which lineage j emerges.

Lineage count, saturation. In the following, we will find that we need to count the deme-specific numbers of lineages present at a given time. Accordingly, for any $\mathbb{L}' \subseteq \mathbb{L}$, $\eta \in (\mathbb{Z}_+ \times \bar{\mathbb{D}} \times \mathbb{L}')^{\mathbb{L}'}$, and $i \in \mathbb{D}$, let us define

$$\ell_i(\eta) := |\{j \in \mathbb{L}' \mid \text{deme}(\eta_j) = i\}| \in \mathbb{Z}_+, \quad \ell(\eta) := (\ell_i(\eta))_{i \in \mathbb{D}} \in \mathbb{Z}_+^{\mathbb{D}}.$$

Note that lineages j for which $\eta_j = \bar{\emptyset}$ are not counted. With this definition, it follows that $\ell_i(y(t))$ is the number of lineages in deme i at time t .

We will also have occasion to refer to the deme-specific number of lineages emerging from a given event. Accordingly, for $\mathbb{L}' \subseteq \mathbb{L}$, $\eta, \eta' \in (\mathbb{Z}_+ \times \bar{\mathbb{D}} \times \mathbb{L}')^{\mathbb{L}'}$, and $i \in \mathbb{D}$, let us define

$$s_i(\eta, \eta') := |\{j \in \mathbb{L}' \mid \text{deme}(\eta_j) = i \ \& \ \text{ct}(\eta'_j) = \text{ct}(\eta_j) + 1\}|, \quad s(\eta, \eta') := (s_i(\eta, \eta'))_{i \in \mathbb{D}} \in \mathbb{Z}_+^{\mathbb{D}}.$$

With this definition, $s_i(\tilde{y}(t), y(t))$ is the number of lineages in deme i that emerge from an event at time t and that, if t is not an event time, then $s(\tilde{y}(t), y(t)) = 0$.

3.2. Obscuration.

4. Results.

Definition 2. For $n, r, \ell, n \in \mathbb{Z}_+^{\mathbb{D}}$, define the multivariable binomial coefficient by

$$\binom{n}{r} := \prod_{i \in \mathbb{D}} \binom{n_i}{r_i}.$$

Using this, we define the *binomial ratio*

$$\binom{n \quad \ell}{r \quad s} := \frac{\binom{n - \ell}{r - s}}{\binom{n}{r}} \in [0, 1].$$

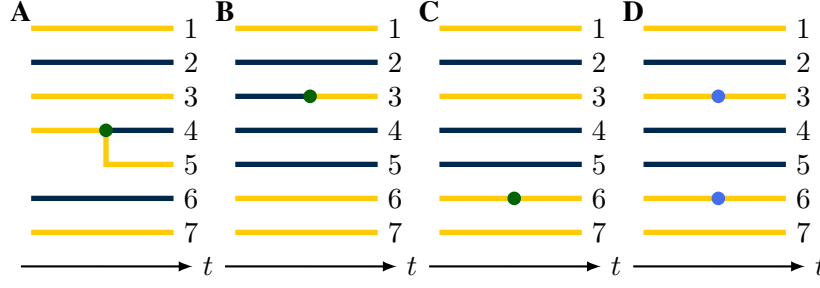


FIGURE 3.

In consequence of the Chu-Vandermonde identity, we have

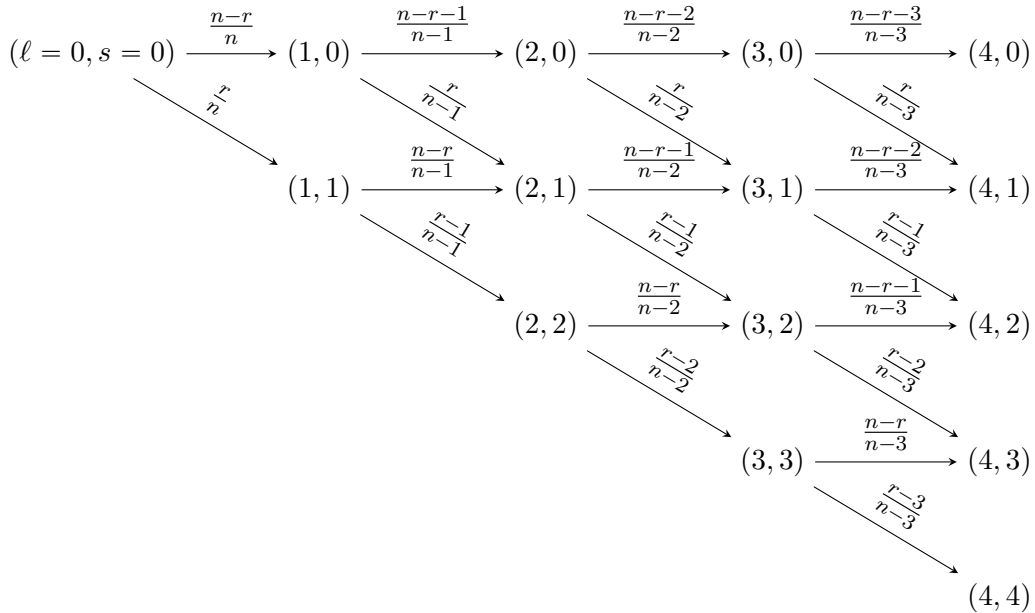
$$\sum_{s \in \mathbb{Z}_+^D} \binom{n}{r} \binom{\ell}{s} \binom{\ell}{s} = 1 \quad (13)$$

Likelihood of a pruned genealogy. Now certain population-process events at certain times are incompatible with any given pruned genealogy. Example: a sample event where no sample is seen, or a migration event involving a change of deme not observed in the pruned genealogy. Define the function $Q_u(\eta, \eta') = 1$ if a change $\eta \rightarrow \eta'$ at time t is compatible with an event of type u at that time, and $Q_u = 0$ otherwise. In other words, Q_u is the indicator function for the condition that there exist $\omega \in \Omega$, $t \in \mathbb{R}_+$, and $k \in \mathbb{Z}_+$ such that $t = T_k(\omega)$, $u = U_k(\omega)$, $\tilde{y}(t, \omega) = \eta$, and $y(t, \omega) = \eta'$. Using this, we define

$$\phi_u(\xi, \eta, \eta') := \binom{n(\xi)}{r_u} \binom{\ell(\eta')}{s(\eta, \eta')} Q_u(\eta, \eta').$$

Theorem 1.

$$\mathbb{P}[\mathcal{G}|\mathcal{H}] = \prod_{(u,t,x) \in \mathcal{H}} \phi_u(x, \tilde{y}(t), y(t)).$$



$$\frac{r!}{(r-s)!} \cdot \frac{(n-\ell)!}{n!} \cdot \frac{(n-r)!}{(n-r-\ell+s)!} = \frac{r!(n-r)!}{n!} \cdot \frac{(n-\ell)!}{(n-\ell-r+s)!(r-s)!} = \binom{n-\ell}{r-s}$$

Suppose we have a pruned genealogy \mathcal{P}^* , defined on the time-interval $[0, T]$, with event times $0 = t_0 < t_1 < \dots < t_n = T$. From the theorem, we have, for $t \notin \text{event}(\mathcal{P}^*)$,

$$\frac{\partial w}{\partial t} = \sum_u \int w(t, x') \alpha_u(t, x', x) \phi_u(x, y(t), y(t)) dx' - \sum_u \int w(t, x) \alpha_u(t, x, x') dx'. \quad (14)$$

At event times, $t \in \text{event}(\mathcal{P}^*)$, one has

$$w(t, x) = \sum_u \int \tilde{w}(t, x') \frac{\alpha_u(t_k, x', x)}{\mu} \phi_u(x, \tilde{y}(t), y(t)) dx', \quad (15)$$

In addition, there are the initial and boundary conditions

$$w(0, x) = p_0(x), \quad w(t, x) = 0 \quad \text{whenever} \quad n(x) < \ell(y(t)). \quad (16)$$

Likelihood of an obscured genealogy. Let $\kappa^\mathcal{V}$ be the indicator function for the condition that a pruned genealogy is compatible with a given obscured genealogy. In particular, given an obscured genealogy \mathcal{V} , set $\kappa_u^\mathcal{V}(t, x, x', \eta, \eta') = \mathbb{1}\{\exists \mathcal{P} \text{ s.t. } \kappa_u^\mathcal{P}(t, x, x', y, y') = 1 \text{ \& } \text{obs}(\mathcal{P}) = \mathcal{V}\}$.

5. Examples.

5.1. SEIRS.

Jumps: $\mathbb{U} = \{\text{Inf}, \text{Prog}, \text{Recov}, \text{Wane}, \text{Birth}, \text{Deaths}, \text{Death}_\text{E}, \text{Death}_\text{I}, \text{Death}_\text{R}, \text{Sample}\}$.

Demes: $\mathbb{D} = \{\text{E}, \text{I}\}$.

Jump rates:

- $\alpha_\text{Inf}(t, x, x') = \beta(t) \frac{x^\text{S} x^\text{I}}{N(t)} \mathbb{1}\{x' = x + (-1, 1, 0, 0)\}$
- $\alpha_\text{Prog}(x, x') = \rho x^\text{E} \mathbb{1}\{x' = x + (0, -1, 1, 0)\}$
- $\alpha_\text{Recov}(x, x') = \gamma x^\text{I} \mathbb{1}\{x' = x + (0, 0, -1, 1)\}$
- $\alpha_\text{Wane}(x, x') = \nu x^\text{R} \mathbb{1}\{x' = x + (1, 0, 0, -1)\}$
- $\alpha_\text{Sample}(t, x, x') = \psi x^\text{I} \mathbb{1}\{x' = x\}$
- $\alpha_\text{Birth}(t, x, x') = B(t) \mathbb{1}\{x' = x + (1, 0, 0, 0)\}$
- $\alpha_\text{Death}_k(x, x') = \mu x^k \mathbb{1}\{x'^j = x^j - \delta_{jk}\}, k \in \{\text{S}, \text{E}, \text{I}, \text{R}\}$

6. Discussion.

References.

- Andersen, P. K., Borgan, Ø., Gill, R. D., & Keiding, N. (1993) *Statistical models based on counting processes*. Springer Series in Statistics. New York: Springer-Verlag.
- Ethier, S. N. & Kurtz, T. G. (2009) *Markov processes: characterization and convergence*. John Wiley & Sons.
- Kallenberg, O. (1997) *Foundations of Modern Probability*. Springer, 2nd edn.
- King, A. A., Lin, Q., & Ionides, E. L. (2022) Markov genealogy processes. *Theoretical Population Biology* **143**:77–91.
- Stadler, T. (2010) Sampling-through-time in birth-death trees. *Journal of Theoretical Biology* **267**:396–404.

Volz, E. M., Kosakovsky Pond, S. L., Ward, M. J., Leigh Brown, A. J., & Frost, S. D. W. (2009) Phylodynamics of infectious disease epidemics. *Genetics* **183**:1421–1430.

A. A. KING, DEPARTMENT OF ECOLOGY & EVOLUTIONARY BIOLOGY, CENTER FOR THE STUDY OF COMPLEX SYSTEMS, AND DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN, ANN ARBOR, MI 48109 USA

Email address: kingaa@umich.edu

URL: <https://kinglab.eeb.lsa.umich.edu/>

Q.-Y. LIN, THEORETICAL BIOLOGY AND BIOPHYSICS, LOS ALAMOS NATIONAL LABORATORY, LOS ALAMOS, NM XXXXX USA

E. L. IONIDES, DEPARTMENT OF STATISTICS UNIVERSITY OF MICHIGAN, ANN ARBOR, MI 48109 USA