

DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS
Volume 36, Number 1, January 2016

doi:10.3934/dcds.2016.36.63

pp. 63–95

THE GENERAL RECOMBINATION EQUATION IN CONTINUOUS TIME AND ITS SOLUTION

ELLEN BAAKE

Technische Fakultät, Universität Bielefeld
Postfach 100131, 33501 Bielefeld, Germany

MICHAEL BAAKE

Fakultät für Mathematik, Universität Bielefeld
Postfach 100131, 33501 Bielefeld, Germany

MAJID SALAMAT

Technische Fakultät, Universität Bielefeld
Postfach 100131, 33501 Bielefeld, Germany

(Communicated by Rafael de la Llave)

ABSTRACT. The process of recombination in population genetics, in its deterministic limit, leads to a nonlinear ODE in the Banach space of finite measures on a locally compact product space. It has an embedding into a larger family of nonlinear ODEs that permits a systematic analysis with lattice-theoretic methods for general partitions of finite sets. We discuss this type of system, reduce it to an equivalent finite-dimensional nonlinear problem, and establish a connection with an ancestral partitioning process, backward in time. We solve the finite-dimensional problem recursively for generic sets of parameters and briefly discuss the singular cases, and how to extend the solution to this situation.

1. Introduction. This contribution is concerned with differential equation models for the dynamics of the genetic composition of populations that evolve under recombination. Here, recombination is the genetic mechanism in which two parent individuals are involved in creating the mixed type of their offspring during sexual reproduction. The essence of this process is illustrated in Fig. 1 and may be idealised and summarised as follows.

Genetic information is encoded in terms of finite sequences. Eggs and sperm (i.e., female and male germ cells or *gametes*) each carry one such sequence. They go through the following life cycle: At fertilisation, two gametes meet randomly and unite, thus starting the life of a new individual, which is equipped with both the maternal and the paternal sequence. At maturity, this individual generates its own germ cells. This process may include recombination, that is, the maternal and paternal sequences perform one or more *crossovers* and are cut and relinked accordingly, so that two ‘mixed’ sequences emerge. These are the new gametes and start the next round of fertilisation (by random mating within a large population).

Models of this process aim at describing the dynamics of the genetic composition of a population that goes through this life cycle repeatedly. These models come in

2010 *Mathematics Subject Classification.* 34G20, 06B23, 92D10.

Key words and phrases. Nonlinear recombination equation, population genetics, measure-valued ODE, lattice of partitions, recursive solution.

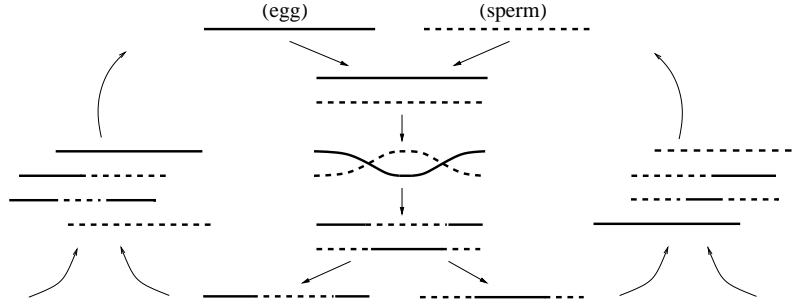


FIGURE 1. Life cycle of a population under sexual reproduction and recombination. Each line symbolises a sequence of sites that defines a gamete (such as the two at the top that start the cycle as ‘egg’ and ‘sperm’). The pool of gametes at the left and the right comes from a large population of recombining individuals. These sequences meet randomly to start the life of a new individual. Altogether, each of these sequences has been pieced together from two randomly chosen parental sequences.

various flavours: in discrete or continuous time; with various assumptions about the crossover pattern; and in a deterministic or a stochastic formulation, depending on whether or not the population is assumed to be so large that stochastic fluctuations may be neglected. We will employ the deterministic continuous-time approach here, but allow for very general crossover patterns. The biologically relevant cases will be mentioned throughout the paper, as will various connections to the existing body of literature.

From now on, we describe populations at the level of their gametes and thus identify gametes with individuals. Their genetic information is encoded in terms of a linear arrangement of sites, indexed by the set $S := \{1, \dots, n\}$. For each site $i \in S$, there is a set X_i of ‘letters’ that may possibly occur at that site. For the sake of concreteness, we use *finite* sets X_i for the moment; we generalise this to arbitrary locally compact spaces X_i in Section 6.

A *type* is thus defined as a sequence $x = (x_1, \dots, x_n) \in X_1 \times \dots \times X_n =: X$, where X is called the *type space*. By construction, x_i is the i -th component or coordinate of x , and we define $x_I := (x_i)_{i \in I}$ as the collection of ‘coordinates’ with indices in I , where I is a subset of S . A *population* is identified with (or described by) a probability vector $p = (p(x))_{x \in X}$ on X , where $p(x)$ denotes the proportion of individuals of type x in X . Note that we assume the sequences to have fixed length. Additional processes that may change this, such as copying blocks, are disregarded here; see [22] and references therein for possible extensions.

With Fig. 1 in mind, recombination may now be modelled as follows. A new (‘offspring’) sequence is formed as the ‘mixture’ of two randomly chosen parental sequences (say x and y) from the population: It copies the letters of x at some of its sites and those of y at all others. If, for example, a double crossover happens between sites i and $i + 1$ and between j and $j + 1$ ($i < j$), then the offspring sequence reads $(x_1, \dots, x_i, y_{i+1}, \dots, y_j, x_{j+1}, \dots, x_n)$. The offspring sequence replaces a randomly chosen sequence (possibly one of the parents, but this is negligibly rare in a large population). Viewed differently, the offspring sequence in our example

reads x whenever the parents are of the form $(x_1, \dots, x_i, *, \dots, *, x_{j+1}, \dots, x_n)$ and $(*, \dots, *, x_{i+1}, \dots, x_j, *, \dots, *)$. Here, a ‘*’ at site i stands for an arbitrary element of X_i , so means marginalisation. This will be helpful when formulating the differential equation.

The sites that come from the paternal and the maternal sequences, respectively, define a *partition* \mathcal{A} of S into two parts. Due to the random choice of the parents, we need not keep track of which sequence was ‘maternal’ and which was ‘paternal’. In principle, all partitions of S into two parts ($\mathcal{A} = \{A_1, A_2\}$) can be realised, via a suitable number of crossovers at suitable positions. If no crossover happens, then the partition is $\mathcal{A} = \{S\}$, and the offspring is an exact copy of the first parent. Reproduction with recombination according to a partition \mathcal{A} happens at rate $\varrho(\mathcal{A})$.

We shall introduce all notions with more care later. For now, we turn the verbal description into a differential equation system and obtain

$$\dot{p}_t(x) = \sum_{\mathcal{A} \in \mathbb{P}_2(S)} \varrho(\mathcal{A}) (p_t(x_{A_1}, *) p_t(*, x_{A_2}) - p_t(x)) \quad (1)$$

for all $x \in X$, where $\mathbb{P}_2(S)$ denotes the set of partitions of S into two parts. Eq. (1) may be understood as a ‘mass balance’ equation: For every $\mathcal{A} \in \mathbb{P}_2(S)$, sequences of type x are ‘produced’ from the corresponding parental sequences at overall rate $\varrho(\mathcal{A}) p_t(x_{A_1}, *) p_t(*, x_{A_2})$, where the product reflects the random combination; at the same time, sequences of type x are lost (i.e., replaced by new ones) at overall rate $\varrho(\mathcal{A}) p_t(x)$. Note that the case $\mathcal{A} = \{S\}$ provides no net contribution to $\dot{p}_t(x)$, since gain and loss are equal in this case.

The resulting ODE system appears difficult to handle, due to the large number of possible states and the nonlinearity of the right-hand side. In previous papers [6, 5, 4, 26, 3, 7, 8], we have concentrated on a special case, namely, the situation in which at most one crossover happens at any given time. That is, we restricted attention to *ordered partitions into two parts*, corresponding to the sites before and after a single-crossover point. We have analysed the resulting models in continuous time (both deterministic and stochastic), as well as in discrete time. For the deterministic continuous-time system, a simple explicit solution is available [6, 5]. This simplicity is due to some underlying linearity.

It is now time to tackle the case of general partitions (in continuous time). Therefore, in this contribution, we give up the single-crossover assumption – and even allow for an arbitrary number of parents in a given recombination event, which leads to partitions with more than two parts. Even though this is not a common biological feature, we will see that it requires little extra mathematical effort. Also, it is a very natural structure on the lattice of partitions of a (finite) set. The restriction to partitions form the biologically most relevant subset $\mathbb{P}_2(S)$ will always be possible by a suitable choice of the model parameters, which are the recombination rates $\varrho(\mathcal{A})$.

This contribution is motivated by the pioneering work of Geiringer [18] and Bennett [10], who worked on a similar system in *discrete* time (but restricted to a special type space); the later work of Lyubich [19, Chapter 6], who worked out much of the underlying structure and got close to a solution in 1992; and by more recent work of Dawson [12, 13], who presented a (recursive) solution in 2000 and 2002. It relies on a certain nonlinear transformation from (gamete or type) frequencies to suitable correlation functions, which decouple from each other and decay geometrically. If sequences of more than three sites are involved, this transformation

must be constructed via recursions that involve the parameters of the recombination process.

Dawson's construction testifies to remarkable insight into the problem. However, it is not easy to penetrate to the mathematical core of his arguments. We therefore start at the very beginning and formulate the model on a fairly general type space, in a measure-theoretic framework, and allowing for arbitrary partitions. More importantly, we put the problem into a systematic lattice-theoretic setting; this will become the key for the transparent construction of the solution. Furthermore, we establish a connection with a partitioning process backward in time, which describes how an individual in the present population has been pieced together from the genetic material of its ancestors. This provides a link to the *ancestral recombination graph* (ARG), which is the ancestral process commonly used in models of recombination in *finite* populations, see [14, Ch. 3.4].

The paper is organised as follows. After introducing the mathematical objects we need and some of their properties in Section 2, the general recombination equation is discussed in Section 3. As a first step, this is done in the setting of a measure-valued ordinary differential equation (ODE), which is then reduced to a finite-dimensional ODE system. Section 4 solves this system under a linearity assumption, which is motivated by previous work, but does not give the solution in sufficient generality.

As a further preparation for the general solution, we study the behaviour of the system under marginalisation in Section 5. Based on this, Section 6 establishes the connection with the partitioning process. This is followed by the derivation of the general solution in Section 7, which is recursive in nature and applies to the generic choice of the recombination rates. More detailed properties of the solution are investigated in Section 8, while Section 9 deals with various types of non-generic cases. The Appendix provides some material for the treatment of degenerate cases.

This paper builds on previous work, most importantly on [6, 5]. Some of the results from these papers will freely be used below, and not re-derived here (though we will always provide precise references).

2. Partitions, measures and recombinators. Let S be a finite set, and consider the lattice $\mathbb{P} = \mathbb{P}(S)$ of partitions of S ; see [1] for general background on lattice theory. When the cardinality of S is $|S| = n$, the set \mathbb{P} contains $B(n)$ elements, known as the *Bell number*; compare [23, A000110]. With $B(0) := 1$, these numbers are recursively computed as $B(n+1) = \sum_{k=0}^n \binom{n}{k} B(k)$ for $n \geq 0$, with generating function $F(z) := \sum_{n=0}^{\infty} \frac{B(n)}{n!} z^n = \exp(e^z - 1)$ and explicit formula $B(n) = \frac{1}{e} \sum_{k=0}^{\infty} \frac{k^n}{k!}$. A subset of relevance to us, for the biological applications, consists of all partitions of S into two parts, $\mathbb{P}_2(S) = \{\mathcal{A} \in \mathbb{P}(S) \mid |\mathcal{A}| = 2\}$, which contains $2^{n-1} - 1$ elements. Note that $\mathbb{P}_2(S)$ generates the lattice $\mathbb{P}(S)$ in an obvious way.

Here, we write a partition of S as $\mathcal{A} = \{A_1, \dots, A_r\}$, where $r = |\mathcal{A}|$ is the number of its parts (or blocks), and one has $A_i \cap A_j = \emptyset$ for all $i \neq j$ together with $A_1 \cup \dots \cup A_r = S$. The natural ordering relation is denoted by \preceq , where $\mathcal{A} \preceq \mathcal{B}$ means that \mathcal{A} is *finer* than \mathcal{B} , or that \mathcal{B} is *coarser* than \mathcal{A} . The conditions $\mathcal{A} \preceq \mathcal{B}$ and $\mathcal{B} \succeq \mathcal{A}$ are synonymous, while $\mathcal{A} \prec \mathcal{B}$ means $\mathcal{A} \preceq \mathcal{B}$ together with $\mathcal{A} \neq \mathcal{B}$.

The joint refinement of two partitions \mathcal{A} and \mathcal{B} is written as $\mathcal{A} \wedge \mathcal{B}$, and is the coarsest partition below \mathcal{A} and \mathcal{B} . The unique *minimal* partition within the lattice \mathbb{P} is denoted as $\mathbb{0} = \{\{x\} \mid x \in S\}$, while the unique *maximal* one is $\mathbb{1} = \{S\}$. When $\mathcal{A} \preceq \mathcal{B}$, we also employ the interval notation $[\mathcal{A}, \mathcal{B}] := \{\mathcal{C} \mid \mathcal{A} \preceq \mathcal{C} \preceq \mathcal{B}\}$. For a general subset \mathbb{G} of \mathbb{P} , we write the complement as $\mathbb{G}^c = \mathbb{P} \setminus \mathbb{G}$. Finally, when

$\mathbb{G} \subseteq \mathbb{P}$, the coarsest partition below all elements of \mathbb{G} is denoted by $\bigwedge \mathbb{G}$. Note that $\bigwedge \emptyset = \underline{1}$ by convention.

When U and V are disjoint (finite) sets, two partitions $\mathcal{A} \in \mathbb{P}(U)$ and $\mathcal{B} \in \mathbb{P}(V)$ can be joined to form an element of $\mathbb{P}(U \cup V)$. We denote such a *joining* by $\mathcal{A} \sqcup \mathcal{B}$, and similarly for multiple joinings. Conversely, if $U \subset S$, a partition $\mathcal{A} \in \mathbb{P}(S)$, with $\mathcal{A} = \{A_1, \dots, A_r\}$ say, defines a unique partition of U by restriction. The latter is denoted by $\mathcal{A}|_U$, and its parts are precisely all non-empty sets of the form $A_i \cap U$ with $1 \leq i \leq r$.

Let now $S = \{1, 2, \dots, n\}$ and define $X = X_1 \times \dots \times X_n$, where each X_i is a locally compact space. In many concrete applications, the X_i will be finite sets, but we do not make such a restriction as it is neither necessary nor desirable. In particular, there are situations in quantitative genetics [11] that will profit from the more general setting we employ here.

When S and X are given, we denote the natural projection to the i th component by π_i , so $\pi_i(X) = X_i$. Similarly, for an arbitrary non-empty subset $U \subseteq S$, we use the notation $\pi_U: X \rightarrow X_U := \times_{i \in U} X_i$ for the projection to the subspace X_U .

Let $\mathcal{M}(X)$ denote the space of finite, regular Borel measures on X , equipped with the usual total variation norm $\|\cdot\|$, which makes it into a Banach space. Also, we need the closed subset (or cone) $\mathcal{M}_+(X)$ of positive measures, which we mean to include the zero measure. Within $\mathcal{M}_+(X)$, we denote the closed subset of probability measures by $\mathcal{P}(X)$. Note that $\mathcal{M}_+(X)$ and $\mathcal{P}(X)$ are convex sets. The restriction of a measure $\mu \in \mathcal{M}(X)$ to a subspace X_U is written as $\pi_U \cdot \mu := \mu \circ \pi_U^{-1}$, which is consistent with marginalisation of measures. For any Borel set $A \subseteq X_U$, one thus has the relation $(\pi_U \cdot \mu)(A) = \mu(\pi_U^{-1}(A))$.

Given a measure $\nu \in \mathcal{M}(X)$ and a partition $\mathcal{A} = \{A_1, \dots, A_r\} \in \mathbb{P}$, we define the mapping $R_{\mathcal{A}}: \mathcal{M}(X) \rightarrow \mathcal{M}(X)$ by $\nu \mapsto R_{\mathcal{A}}(\nu)$ with $R_{\mathcal{A}}(0) := 0$ and, for $\nu \neq 0$,

$$R_{\mathcal{A}}(\nu) := \frac{1}{\|\nu\|^{r-1}} \bigotimes_{i=1}^r (\pi_{A_i} \cdot \nu) \quad (2)$$

where the product is (implicitly) ‘site ordered’, i.e. it matches the ordering of the sites as specified by the set S . We shall also use site ordering for cylinder or product sets. We call a mapping of type $R_{\mathcal{A}}$ a *recombinator*. Note that recombinators are nonlinear whenever $\mathcal{A} \neq \underline{1}$.

Proposition 1. *Let $S = \{1, 2, \dots, n\}$ and $X = X_1 \times \dots \times X_n$ as above. Now, let $\mathcal{A} \in \mathbb{P}(S)$ be arbitrary, and consider the corresponding recombinator as defined by Eq. (2). Then, the following assertions are true.*

1. $R_{\mathcal{A}}$ is positive homogeneous of degree 1, which means that $R_{\mathcal{A}}(a\nu) = aR_{\mathcal{A}}(\nu)$ holds for all $\nu \in \mathcal{M}(X)$ and $a \geq 0$.
2. $R_{\mathcal{A}}$ is globally Lipschitz on $\mathcal{M}(X)$, with Lipschitz constant $L \leq 2|\mathcal{A}| + 1$.
3. $\|R_{\mathcal{A}}(\nu)\| \leq \|\nu\|$ holds for all $\nu \in \mathcal{M}(X)$.
4. $R_{\mathcal{A}}$ maps $\mathcal{M}_+(X)$ into itself.
5. $R_{\mathcal{A}}$ preserves the norm of positive measures, and hence also maps $\mathcal{P}(X)$ into itself.

Proof. Claims (1) and (4) are elementary consequences of the definition. Claim (3) follows easily from the arguments used in [6, Sec. 3.1], while (5) is clear from the observation that $\|\nu\| = \nu(X)$ for $\nu \geq 0$, which then implies $\|R_{\mathcal{A}}(\nu)\| = \|\nu\|$ by standard arguments; compare [6, Fact 2].

It remains to prove (2). Let $\mathcal{A} = \{A_1, \dots, A_r\}$ be a partition with r parts, where $1 \leq r \leq n$. Then, one can show inductively in r that

$$\left\| \bigotimes_{i=1}^r (\pi_{A_i} \cdot \mu) - \bigotimes_{j=1}^r (\pi_{A_j} \cdot \nu) \right\| \leq \|\mu - \nu\| \sum_{k=0}^{r-1} \|\mu\|^k \|\nu\|^{r-1-k}, \quad (3)$$

for arbitrary $\mu, \nu \in \mathcal{M}(X)$. If one of them is the zero measure, the Lipschitz estimate is a consequence of claim (3).

Let now $\mu, \nu \in \mathcal{M}(X)$ both be non-zero, and assume that $\|\mu\| \leq \|\nu\|$. Using Eq. (3) together with positive homogeneity of $R_{\mathcal{A}}$, we have

$$\begin{aligned} & \|R_{\mathcal{A}}(\mu) - R_{\mathcal{A}}(\nu)\| \\ &= \left\| \|\mu\| \left(\bigotimes_{i=1}^r \pi_{A_i} \cdot \frac{\mu}{\|\mu\|} - \bigotimes_{j=1}^r \pi_{A_j} \cdot \frac{\nu}{\|\nu\|} \right) + (\|\mu\| - \|\nu\|) \bigotimes_{k=1}^r \pi_{A_k} \cdot \frac{\nu}{\|\nu\|} \right\| \\ &\leq r \min(\|\mu\|, \|\nu\|) \left\| \frac{\mu}{\|\mu\|} - \frac{\nu}{\|\nu\|} \right\| + \|\mu - \nu\|, \end{aligned} \quad (4)$$

where we used that $\mu/\|\mu\|$ and $\nu/\|\nu\|$ are measures of norm 1. Next, one has

$$\left\| \frac{\mu}{\|\mu\|} - \frac{\nu}{\|\nu\|} \right\| \leq \frac{2\|\mu - \nu\|}{\max(\|\mu\|, \|\nu\|)}$$

which, on inserting into Eq. (4), gives the inequality

$$\|R_{\mathcal{A}}(\mu) - R_{\mathcal{A}}(\nu)\| \leq (2r + 1) \|\mu - \nu\|$$

from which our claim follows, with $L \leq 2r + 1$. \square

Remark 1. Let us mention that there is an alternative way to see the Lipschitz property, at least for finite X . This is because the dynamics may then be reformulated in terms of a chemical reaction system. This is a large class of models, for which a substantial body of theory is available; see [24] for a review. In particular, the Lipschitz property applies under very general conditions, which are satisfied in our case. \diamond

The estimate used above is rather coarse, but suffices for our needs here. When using the invariance of $\mathcal{P}(X)$, one simply employs Eq. (3) from the last proof to establish the following consequence.

Corollary 1. *On $\mathcal{P}(X)$, with $\mathcal{A} \in \mathbb{P}(S)$, the recombinator $R_{\mathcal{A}}$ is Lipschitz with $L \leq |\mathcal{A}|$.* \square

Before we embark on the recombination equation and its solution, we need to establish one technical result on the relation between recombinators and projectors to subsystems, as defined by non-empty sets $U \subset S$. Here, the system on which a recombinator acts is marked by an upper index S (for the full system) or U (for the subsystem). Later, we will drop this index whenever the meaning is unambiguous.

Lemma 1. *Let S be a finite set as above, and $\mathcal{A} = \{A_1, \dots, A_r\} \in \mathbb{P}(S)$ an arbitrary partition. If $U \subseteq S$ is non-empty and $\omega \in \mathcal{M}_+(X)$, one has*

$$\pi_U \cdot (R_{\mathcal{A}}^S(\omega)) = R_{\mathcal{A}|_U}^U(\pi_U \cdot \omega),$$

where $\mathcal{A}|_U \in \mathbb{P}(U)$ and the upper index of a recombinator indicates on which measure space it acts.

Proof. We shall show the claimed identity by verifying it for certain ‘rectangular’ measurable sets that suffice for the equality of the two measures as Baire measures, and then rely on the unique extension of Baire measures to regular Borel measures; compare the discussion around [6, Fact 1]. To this end, let $\mathcal{A} \in \mathbb{P}(S)$ be given as $\mathcal{A} = \{A_1, \dots, A_r\}$, and write $\mathcal{A}|_U \in \mathbb{P}(U)$ as $\mathcal{A}|_U = \{V_1, \dots, V_s\}$, where $s \leq r$. Each V_i is contained in precisely one part of \mathcal{A} . Without loss of generality, we may thus assume that $V_i = A_i \cap U$ for $1 \leq i \leq s$, while $A_j \cap U = \emptyset$ for all $j > s$ (when $r = s$, this case does not occur).

Now, let E_i be a Borel set in $X_{V_i} = \pi_{V_i}(X)$, and consider the corresponding ‘rectangular’ set $E = E_1 \times \dots \times E_s \subseteq X_U$, where we again assume that the product observes proper site ordering. When evaluated on E , the right-hand side of our claim gives

$$\begin{aligned} (R_{\mathcal{A}|_U}^U(\pi_U \cdot \omega))(E) &= \frac{1}{\|\pi_U \cdot \omega\|^{s-1}} \bigotimes_{i=1}^s (\pi_{V_i}^{(U)} \cdot (\pi_U \cdot \omega))(E_1 \times \dots \times E_s) \\ &= \frac{1}{\|\omega\|^{s-1}} \bigotimes_{i=1}^s (\pi_{V_i} \cdot \omega)(E_1 \times \dots \times E_s), \end{aligned}$$

where we have used that $\|\pi_U \cdot \omega\| = \|\omega\|$ and where $\pi^{(U)}$ denotes a projector that is only defined on the subspace X_U . This is now to be compared with the evaluation of the left hand side of the claim, which gives (with $U^c = S \setminus U$)

$$\begin{aligned} (\pi_U \cdot R_{\mathcal{A}}^S(\omega))(E) &= R_{\mathcal{A}}^S(\omega)(\pi_U^{-1}(E)) = \frac{1}{\|\omega\|^{r-1}} \bigotimes_{i=1}^r (\pi_{A_i} \cdot \omega)(E \times X_{U^c}) \\ &= \frac{1}{\|\omega\|^{r-1}} \left(\bigotimes_{i=1}^s (\pi_{A_i} \cdot \omega) \right) \otimes \left(\bigotimes_{j=s+1}^r (\pi_{A_j} \cdot \omega) \right) (E_1 \times \dots \times E_s \times X_{U^c}) \\ &= \frac{1}{\|\omega\|^{r-1}} \left(\prod_{i=1}^s \omega(\pi_{A_i}^{-1}(E_i \times X_{A_i \setminus V_i})) \right) \left(\prod_{j=s+1}^r \omega(\pi_{A_j}^{-1}(X_{A_j})) \right) \\ &= \frac{1}{\|\omega\|^{r-1}} \left(\prod_{i=1}^s \omega(\pi_{V_i}^{-1}(E_i)) \right) \left(\prod_{j=s+1}^r \omega(X) \right) \\ &= \frac{\|\omega\|^{r-s}}{\|\omega\|^{r-1}} \prod_{i=1}^s (\pi_{V_i} \cdot \omega)(E_i) = \frac{1}{\|\omega\|^{s-1}} \bigotimes_{i=1}^s (\pi_{V_i} \cdot \omega)(E_1 \times \dots \times E_s), \end{aligned}$$

which agrees with our previous expression and thus proves the claim. \square

Lemma 1 has some interesting consequences for the structure of recombinators. We only state the result and omit the proof, as it is analogous to the corresponding result in [6].

Proposition 2. *Let $\mathcal{A}, \mathcal{B} \in \mathbb{P}(S)$. On $\mathcal{M}_+(X)$, the corresponding recombinators satisfy*

$$R_{\mathcal{A}} R_{\mathcal{B}} = R_{\mathcal{A} \wedge \mathcal{B}}.$$

In particular, each recombinator is an idempotent and any two recombinators commute. \square

Below, we shall also need probability vectors on $\mathbb{P} = \mathbb{P}(S)$, which form the compact space $\mathcal{P}(\mathbb{P})$. An interesting subclass of them can be constructed as follows.

Let $f: \mathbb{P} \rightarrow [0, 1]$ be a function on the lattice \mathbb{P} , and consider

$$c(\mathcal{A}) := \sum_{\substack{\mathbb{G} \subseteq \mathbb{P} \\ \bigwedge \mathbb{G} = \mathcal{A}}} \prod_{\mathcal{B} \in \mathbb{G}} (1 - f(\mathcal{B})) \prod_{\mathcal{C} \in \mathbb{G}^c} f(\mathcal{C}) \quad (5)$$

for $\mathcal{A} \in \mathbb{P}$, which clearly satisfies $c(\mathcal{A}) \geq 0$. Moreover, one has

$$\sum_{\mathcal{A} \in \mathbb{P}} c(\mathcal{A}) = \sum_{\mathbb{G} \subseteq \mathbb{P}} \prod_{\mathcal{B} \in \mathbb{G}} (1 - f(\mathcal{B})) \prod_{\mathcal{C} \in \mathbb{G}^c} f(\mathcal{C}) = \prod_{\mathcal{A} \in \mathbb{P}} ((1 - f(\mathcal{A})) + f(\mathcal{A})) = 1, \quad (6)$$

wherefore $c \in \mathcal{P}(\mathbb{P}(S))$, which means that c is a probability vector on the lattice $\mathbb{P}(S)$.

At this point, we have gathered the core material to embark on the discussion of the recombination equation and its properties.

3. The general recombination equation. With the tools at hand, it is now rather obvious how to generalise the ‘mass balance’ equation (1) of the Introduction to our measure-theoretic setting. Within the Banach space $(\mathcal{M}(X), \|\cdot\|)$, we thus consider the nonlinear ODE

$$\dot{\omega}_t = \sum_{\mathcal{A} \in \mathbb{P}(S)} \varrho(\mathcal{A}) (R_{\mathcal{A}} - \mathbb{1})(\omega_t) \quad (7)$$

with non-negative numbers $\varrho(\mathcal{A})$ that have the meaning of *recombination rates* in our context. They are written in this way because we consider $\varrho: \mathbb{P}(S) \rightarrow \mathbb{R}$ as an element of the *Möbius algebra* over \mathbb{R} ; see [1, 25] for background. We will usually assume that an initial condition $\omega_0 \in \mathcal{M}(X)$ for $t = 0$ is given for the ODE (7), and then speak of the corresponding *Cauchy problem* (or initial value problem). Biologically, Eq. (7) describes the change in composition of a population in which offspring is produced by piecing together sequences from various parents according to the collection of the $\varrho(\mathcal{A})$ with $\mathcal{A} \in \mathbb{P}(S)$. More precisely, with rate $\varrho(\mathcal{A})$ for every $\mathcal{A} = \{A_1, \dots, A_r\}$, the sites in A_1 inherited from parent 1 are reassociated with the sites in A_2 inherited from parent 2, and with the sites in A_i inherited from parent i for $2 < i \leq r$.

Remark 2. Since $R_{\mathbb{1}} = \mathbb{1}$, the value of $\varrho(\mathbb{1})$ is immaterial, and the corresponding term could clearly be omitted on the right-hand side of Eq. (7). We nevertheless keep it here, as it will become useful in connection with the reduction to subsystems.

Another way to write the ODE is

$$\dot{\omega}_t = -\varrho_{\text{tot}} \omega_t + \sum_{\mathcal{A} \in \mathbb{P}(S)} \varrho(\mathcal{A}) R_{\mathcal{A}}(\omega_t) = -(\varrho_{\text{tot}} - \varrho(\mathbb{1})) \omega_t + \sum_{\mathcal{A} \prec \mathbb{1}} \varrho(\mathcal{A}) R_{\mathcal{A}}(\omega_t)$$

with $\varrho_{\text{tot}} := \sum_{\mathcal{A} \in \mathbb{P}(S)} \varrho(\mathcal{A})$ being the total recombination rate. \diamond

With $\Phi := \sum_{\mathcal{A} \in \mathbb{P}(S)} \varrho(\mathcal{A}) (R_{\mathcal{A}} - \mathbb{1})$, we can now simply write

$$\dot{\omega}_t = \Phi(\omega_t), \quad (8)$$

but we must keep in mind that Φ is a nonlinear operator. Nevertheless, one has the following basic result; compare [2] for background on ODEs on Banach spaces.

Proposition 3. *Let S be a finite set and X the corresponding locally compact product space as introduced above. Then, the Cauchy problem of Eq. (7) with initial condition $\omega_0 \in \mathcal{M}(X)$ has a unique solution. Moreover, the cone $\mathcal{M}_+(X)$ is forward invariant, and the flow is norm-preserving on $\mathcal{M}_+(X)$. In particular, $\mathcal{P}(X)$ is forward invariant under the flow.*

Proof. By part (2) of Proposition 1, all $R_{\mathcal{A}}$ are globally Lipschitz on $\mathcal{M}(X)$, which is then also true of Φ . Therefore, the uniqueness statement for the Cauchy problem is clear.

Next, let $\nu \in \mathcal{M}_+(X)$ and consider an arbitrary Borel set $E \subset X$ with $\nu(E) = 0$. Then, with $\varrho_{\text{tot}} = \sum_{\mathcal{A} \in \mathbb{P}(S)} \varrho(\mathcal{A})$ as above in Remark 2, we have

$$(\Phi(\nu))(E) = -\varrho_{\text{tot}} \nu(E) + \sum_{\mathcal{A} \in \mathbb{P}(S)} \varrho(\mathcal{A}) (R_{\mathcal{A}}(\nu))(E) \geq 0$$

because $\nu(E) = 0$ by assumption and all other terms are non-negative as a result of part (4) of Proposition 1 together with $\varrho(\mathcal{A}) \geq 0$ for all $\mathcal{A} \in \mathbb{P}(S)$. Positive invariance of the closed cone $\mathcal{M}_+(X)$ now follows from a classic continuity argument; see [2, Thm. 16.5 and Rem. 16.6] for details and [6, Thm. 1] for the analogous argument in the single-crossover model.

When ν is a positive measure, one has $\nu(X) = \|\nu\| = \|R_{\mathcal{A}}(\nu)\| = (R_{\mathcal{A}}(\nu))(X)$ as a consequence of Proposition 1, and hence

$$(\Phi(\nu))(X) = -\varrho_{\text{tot}} \nu(X) + \sum_{\mathcal{A} \in \mathbb{P}(S)} \varrho(\mathcal{A}) (R_{\mathcal{A}}(\nu))(X) = 0,$$

which implies the preservation of the norm of a positive measure under the forward flow. The last claim is then obvious. \square

In view of our underlying biological problem, we restrict our attention to the investigation of the recombination equation on the cone $\mathcal{M}_+(X)$, and on $\mathcal{P}(X)$ in particular. To proceed, we will first bring the ODE (7) into a simpler form. To this end, observe the structure of Φ , which suggests that, as the flow proceeds forward in time from an initial measure $\nu \in \mathcal{P}(X)$, the solution picks up fractions of components of the form $R_{\mathcal{B}}(\nu)$ for various or even all $\mathcal{B} \in \mathbb{P}(S)$, depending on which recombination rates are positive and which partitions can thus be reached in the course of time.

Let $|S| = n$ and fix some $\nu \in \mathcal{M}_+(X)$. Now, define the (finite-dimensional) set

$$\Delta_{\nu} := \left\{ \sum_{\mathcal{C} \in \mathbb{P}(S)} q(\mathcal{C}) R_{\mathcal{C}}(\nu) \mid q \in \mathcal{P}(\mathbb{P}(S)) \right\},$$

which is the closed convex set that consists of all convex linear combinations of the measures $R_{\mathcal{C}}(\nu)$ with $\mathcal{C} \in \mathbb{P}(S)$. In fact, when the rates $\varrho(\mathcal{A})$ run through all non-negative values, Δ_{ν} is the smallest closed convex set that contains all measures which can be reached in the course of recombination from the initial measure ν . The dimension of Δ_{ν} depends on the nature of ν , and can even be 0. When the dimension is maximal (meaning $B(n) - 1$, which is the generic case), Δ_{ν} is a simplex and the $B(n)$ measures $R_{\mathcal{C}}(\nu)$ are the *extremal* measures of the simplex in the sense of convex analysis. Observe that

$$R_{\mathcal{B}}(\Delta_{\nu}) \subseteq \Delta_{\nu}$$

holds for any $\mathcal{B} \in \mathbb{P}(S)$, which follows from a straightforward though slightly technical calculation (details of which are given below). This now suggests the ansatz

$$\omega_t = \sum_{\mathcal{C} \in \mathbb{P}(S)} a_t(\mathcal{C}) R_{\mathcal{C}}(\omega_0) \tag{9}$$

for the solution of our above Cauchy problem, with (generic) initial condition ω_0 and coefficient functions a_t , the latter thus with $a_0(\mathcal{C}) = \delta_{\mathcal{C}, \mathbf{1}}$. For each t , we view $a_t : \mathbb{P}(S) \rightarrow \mathbb{R}$ as an element of the Möbius algebra. We shall prove a posteriori

that the strategy of Eq. (9) works, and that the original Banach space ODE is thus reduced to a finite-dimensional system of ODEs, whose solution consists of a one-parameter family of probability vectors on $\mathbb{P}(S)$. The case of non-generic initial conditions will be discussed afterwards.

To proceed, we now have to calculate the action of Φ on a general measure $\omega \in \Delta_\nu$. It is convenient to first consider probability measures, as the extension to general positive measures is then immediate via the positive homogeneity of the recombinators. So, let $\nu \in \mathcal{P}(X)$, which means $\Delta_\nu \subseteq \mathcal{P}(X)$. Consider a single partition $\mathcal{B} = \{B_1, \dots, B_r\}$ and observe that

$$\begin{aligned} R_{\mathcal{B}}(\omega) &= \bigotimes_{i=1}^r \pi_{B_i} \cdot \omega = \bigotimes_{i=1}^r \pi_{B_i} \cdot \left(\sum_{\mathcal{C} \in \mathbb{P}(S)} q(\mathcal{C}) R_{\mathcal{C}}(\nu) \right) \\ &= \bigotimes_{i=1}^r \sum_{\mathcal{C} \in \mathbb{P}(S)} q(\mathcal{C}) (\pi_{B_i} \cdot R_{\mathcal{C}}(\nu)) = \sum_{\substack{\mathcal{C}_1, \dots, \mathcal{C}_r \\ \in \mathbb{P}(S)}} \left(\prod_{j=1}^r q(\mathcal{C}_j) \right) \bigotimes_{i=1}^r R_{\mathcal{C}_i|_{B_i}}^{B_i}(\pi_{B_i} \cdot \nu), \end{aligned}$$

where we have used Lemma 1 in the last step. Observing that the product measure in the last expression is a measure of the form $R_{\mathcal{A}}(\nu)$ for some $\mathcal{A} \in \mathbb{P}(S)$ with $\mathcal{A} \preccurlyeq \mathcal{B}$, and that each such \mathcal{A} must occur here, we see that

$$R_{\mathcal{B}}(\omega) = R_{\mathcal{B}} \left(\sum_{\mathcal{C} \in \mathbb{P}(S)} q(\mathcal{C}) R_{\mathcal{C}}(\nu) \right) = \sum_{\mathcal{A} \preccurlyeq \mathcal{B}} \left(\prod_{i=1}^{|\mathcal{B}|} \sum_{\substack{\mathcal{C} \in \mathbb{P}(S) \\ \mathcal{C}|_{B_i} = \mathcal{A}|_{B_i}}} q(\mathcal{C}) \right) R_{\mathcal{A}}(\nu), \quad (10)$$

where we use a dot under a symbol to mark it as the summation variable, thus following the notation of [1]. Next, if $\alpha \geq 0$, Proposition 1 implies that

$$\alpha \omega = \sum_{\mathcal{C} \in \mathbb{P}(S)} \alpha q(\mathcal{C}) R_{\mathcal{C}}(\nu) = \sum_{\mathcal{C} \in \mathbb{P}(S)} q(\mathcal{C}) R_{\mathcal{C}}(\alpha \nu)$$

are two equivalent ways to write $\omega' = \alpha \omega$. Consequently, we extend our previous formula as

$$R_{\mathcal{B}}(\omega) = \sum_{\mathcal{A} \preccurlyeq \mathcal{B}} \gamma(q; \mathcal{A}, \mathcal{B}) R_{\mathcal{A}}(\nu), \quad (11)$$

where we define the function γ as

$$\gamma(q; \mathcal{A}, \mathcal{B}) := \frac{1}{\|q\|_1^{|\mathcal{B}|-1}} \prod_{i=1}^{|\mathcal{B}|} \sum_{\substack{\mathcal{C} \in \mathbb{P}(S) \\ \mathcal{C}|_{B_i} = \mathcal{A}|_{B_i}}} q(\mathcal{C}) \quad (12)$$

for any $\mathcal{A}, \mathcal{B} \in \mathbb{P}(S)$ with $\mathcal{A} \preccurlyeq \mathcal{B}$, and as $\gamma(q; \mathcal{A}, \mathcal{B}) = 0$ otherwise. In Eq. (12), q may be any real vector of dimension $B(n)$, even though we have only considered positive ones above. Also, one has $\gamma(0; \mathcal{A}, \mathcal{B}) = 0$ for consistency.

Remark 3. When q is a probability vector on $\mathbb{P}(S)$, one has $\|q\|_1 = q(\mathbb{P}(S)) = 1$. The right-hand side of Eq. (11) can then be read as a *product of marginalised probabilities* for the subsystems defined by the parts B_i of the partition \mathcal{B} . This structure will be made explicit in Section 5 (see Lemma 1 in particular) and will later pave the way to a recursive solution of the recombination equation. \diamond

For fixed q , the function $\gamma(q; \cdot, \cdot)$ is an element of the *incidence algebra* of our lattice $\mathbb{P}(S)$ over the field \mathbb{R} . Following [1], we denote this algebra by $\mathbb{A}(\mathbb{P}(S))$, with convolution as multiplication. The latter is defined as

$$(\alpha * \beta)(\mathcal{A}, \mathcal{B}) = \sum_{\mathcal{C} \in [\mathcal{A}, \mathcal{B}]} \alpha(\mathcal{A}, \mathcal{C}) \beta(\mathcal{C}, \mathcal{B}).$$

Note that this definition automatically gives $(\alpha * \beta)(\mathcal{A}, \mathcal{B}) = 0$ if $\mathcal{A} \not\leq \mathcal{B}$, as it must. Important elements of $\mathbb{A}(\mathbb{P}(S))$ include the *unit* δ , defined by $\delta(\mathcal{A}, \mathcal{B}) = \delta_{\mathcal{A}, \mathcal{B}}$, the *zeta function* ζ , defined by $\zeta(\mathcal{A}, \mathcal{B}) = 1$ for $\mathcal{A} \leq \mathcal{B}$ and $\zeta(\mathcal{A}, \mathcal{B}) = 0$ otherwise, and the *Möbius function* μ , which is the multiplicative (left and right) inverse of ζ . One then has $\zeta * \mu = \mu * \zeta = \delta$. We refer to [25] for details and more advanced aspects of incidence algebras.

With $\varrho_{\text{tot}} = \sum_{\mathcal{A} \in \mathbb{P}(S)} \varrho(\mathcal{A})$ and $0 \leq \omega \in \Delta_\nu$, Eq. (11) allows us to continue as follows,

$$\begin{aligned} \Phi(\omega) &= \Phi\left(\sum_{\mathcal{C} \in \mathbb{P}(S)} q(\mathcal{C}) R_{\mathcal{C}}(\nu)\right) = \sum_{\mathcal{B} \in \mathbb{P}(S)} \varrho(\mathcal{B}) (R_{\mathcal{B}} - \mathbb{1}) \sum_{\mathcal{C} \in \mathbb{P}(S)} q(\mathcal{C}) R_{\mathcal{C}}(\nu) \\ &= -\varrho_{\text{tot}} \omega + \sum_{\mathcal{B} \in \mathbb{P}(S)} \varrho(\mathcal{B}) \sum_{\mathcal{A} \leq \mathcal{B}} \gamma(q; \mathcal{A}, \mathcal{B}) R_{\mathcal{A}}(\nu) \\ &= \sum_{\mathcal{A} \in \mathbb{P}(S)} R_{\mathcal{A}}(\nu) \left(-\varrho_{\text{tot}} q(\mathcal{A}) + \sum_{\mathcal{B} \geq \mathcal{A}} \gamma(q; \mathcal{A}, \mathcal{B}) \varrho(\mathcal{B})\right), \end{aligned} \quad (13)$$

which is the desired action of Φ on elements of $\mathcal{M}_+(X)$.

Inserting Eq. (9) into the recombination equation (7) now gives the following result.

Lemma 2. *The ansatz (9) for the solution ω_t of the recombination equation (7) on $\mathcal{M}_+(X)$ leads to a system of induced ODEs for the coefficient functions a_t , namely*

$$\dot{a}_t(\mathcal{A}) = -\varrho_{\text{tot}} a_t(\mathcal{A}) + \sum_{\mathcal{B} \geq \mathcal{A}} \gamma(a_t; \mathcal{A}, \mathcal{B}) \varrho(\mathcal{B}) \quad (14)$$

for $\mathcal{A} \in \mathbb{P}(S)$, where γ is defined as in Eq. (12).

Proof. Clearly, Eq. (14) is the result of a comparison of coefficients, based on the original equation (7), its short form (8), and Eq. (13). For a generic $\nu \in \mathcal{M}_+(X)$, this is justified by the fact that the measures $R_{\mathcal{A}}(\nu)$ with $\mathcal{A} \in \mathbb{P}(S)$ then are the extremal measures of the forward-invariant simplex Δ_ν . Since the set of generic ν is dense in $\mathcal{M}_+(X)$, and the solution of the Cauchy problem depends continuously on the initial condition, the extension to all $\nu \in \mathcal{M}_+(X)$ is consistent. \square

Note that our above calculation was based upon the action of recombinators on *positive* measures, because we have used Lemma 1. Consequently, we do not know how Eqs. (7) and (14) are related beyond this case. Fortunately, this is not required, as the next result shows, where we use $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x \geq 0\}$.

Proposition 4. *The ODE system defined by Eq. (14), with γ as in Eq. (12), is of dimension $B(n) = |\mathbb{P}(S)|$, and has a unique solution for its Cauchy problem.*

The closed cone $\mathcal{M}_+(\mathbb{P}(S)) \simeq \mathbb{R}_+^{B(n)}$ is invariant in forward time, and the $\|\cdot\|_1$ -norm of non-negative initial conditions is preserved. In particular, the simplex $\mathcal{P}(\mathbb{P}(S))$ of probability vectors on $\mathbb{P}(S)$ is invariant in forward time.

Proof. The ODE system of Eq. (14) can be written as $\dot{a}_t = \Psi(a_t)$, where Ψ is nonlinear. The first claim is clear, while solution uniqueness once again follows from Lipschitz continuity of Ψ , which is obvious here.

Let $q \in \mathcal{M}_+(\mathbb{P}(S)) = \mathbb{R}_+^{\mathbb{P}(S)} \simeq \mathbb{R}_+^{B(n)}$ be arbitrary and consider any subset $\mathbb{G} \subset \mathbb{P}(S)$ such that $q(\mathbb{G}) := \sum_{\mathcal{A} \in \mathbb{G}} q(\mathcal{A}) = 0$. Then, one has

$$(\Psi(q))(\mathbb{G}) = \sum_{\mathcal{A} \in \mathbb{G}} (\Psi(q))(\mathcal{A}) = -\varrho_{\text{tot}} q(\mathbb{G}) + \sum_{\mathcal{A} \in \mathbb{G}} \sum_{\mathcal{B} \succ \mathcal{A}} \gamma(q; \mathcal{A}, \mathcal{B}) \varrho(\mathcal{B}) \geq 0,$$

since $q(\mathbb{G}) = 0$ by assumption and all $\gamma(q; \mathcal{A}, \mathcal{B})$ as well as all $\varrho(\mathcal{B})$ are non-negative. This gives the invariance of $\mathcal{M}_+(\mathbb{P}(S))$ under the forward flow by standard arguments [2, Thm. 16.5].

Repeating the calculation for $\mathbb{G} = \mathbb{P}(S)$ yields

$$(\Psi(q))(\mathbb{P}(S)) = -\varrho_{\text{tot}} q(\mathbb{P}(S)) + \sum_{\mathcal{B} \in \mathbb{P}(S)} \varrho(\mathcal{B}) \sum_{\mathcal{A} \preccurlyeq \mathcal{B}} \gamma(q; \mathcal{A}, \mathcal{B}).$$

Now, the last sum (for any fixed \mathcal{B}) is

$$\sum_{\mathcal{A} \preccurlyeq \mathcal{B}} \gamma(q; \mathcal{A}, \mathcal{B}) = \|q\|_1 = q(\mathbb{P}(S)), \quad (15)$$

as follows from Eq. (11) by taking norms on both sides and observing that $\|R_{\mathcal{A}}(\nu)\| = \|\nu\| = 1$ for all $\mathcal{A} \preccurlyeq \mathcal{B}$ as well as $\|R_{\mathcal{B}}(\omega)\| = 1$, while one has $\gamma(q; \mathcal{A}, \mathcal{B}) \geq 0$ for any $q \in \mathcal{M}_+(\mathbb{P}(S))$; alternatively, one can also verify this with a direct calculation on the basis of the definition of γ . Consequently, $(\Psi(q))(\mathbb{P}(S)) = 0$, which implies the claimed norm preservation. The positive invariance of $\mathcal{P}(\mathbb{P}(S))$ is then clear. \square

Remark 4. Observing that $\gamma(q; \mathcal{A}, \mathbb{1}) = q(\mathcal{A})$ holds for any $q \in \mathbb{R}^{\mathbb{P}(S)}$, one can rewrite the ODE of Eq. (7) as

$$\dot{a}_t(\mathcal{A}) = -(\varrho_{\text{tot}} - \varrho(\mathbb{1}))a_t(\mathcal{A}) + \sum_{\mathcal{A} \preccurlyeq \mathcal{B} \prec \mathbb{1}} \gamma(a_t; \mathcal{A}, \mathcal{B}) \varrho(\mathcal{B}),$$

which corresponds to the observation made in Remark 2. \diamond

With Proposition 4, we may now conclude that our ansatz (14) is consistent and suitable for the reduction of the original problem to a finite-dimensional one.

Theorem 1. *The one-parameter family of measures $\{\omega_t \mid t \geq 0\}$ is a solution of the Cauchy problem of Eq. (7) with initial condition $\omega_0 \in \mathcal{M}_+(X)$ if and only if it is of the form*

$$\omega_t = \sum_{\mathcal{B} \in \mathbb{P}(S)} a_t(\mathcal{B}) R_{\mathcal{B}}(\omega_0)$$

where the coefficient functions satisfy the Cauchy problem of Eq. (14) with initial condition $a_0(\mathcal{B}) = \delta(\mathcal{B}, \mathbb{1})$.

Proof. The solution property is clear from our above calculations, while the correspondence of the initial conditions is obvious for the generic case, and extends to the general case by a standard continuity argument. The claim now follows from the uniqueness statements for the two Cauchy problems. \square

Remark 5. Let us mention that an alternative path to Theorem 1 is possible via [2, Thm. 16.5 and Rem. 16.6], by showing that the convex set Δ_ν is forward invariant for any $\nu \in \mathcal{M}_+(X)$ as initial condition. This requires the verification of the ‘inside reflection property’ for any piece of the boundary of Δ_ν , which is

somewhat tedious in view of the possible degeneracies. This is the reason why we opted for our approach above. \diamond

Let us briefly discuss the structure of Δ_ν for a general $\nu \in \mathcal{M}_+(X)$, including the non-generic cases. Each ν gives rise to a set of partitions

$$\mathbb{H}_\nu := \{\mathcal{A} \in \mathbb{P}(S) \mid R_{\mathcal{A}}(\nu) = \nu\},$$

which is non-empty (since $\underline{1} \in \mathbb{H}_\nu$ for all ν) and defines a unique partition

$$\mathcal{U}_\nu := \bigwedge \mathbb{H}_\nu = \{U_1, \dots, U_r\},$$

so that $\mathbb{H}_\nu = \{\mathcal{A} \in \mathbb{P}(S) \mid \mathcal{A} \succcurlyeq \mathcal{U}_\nu\}$. The partition \mathcal{U}_ν also defines the sublattice $\mathbb{P}_\nu := \mathbb{P}(U_1) \times \dots \times \mathbb{P}(U_r)$ of product form, with $B(|U_1|) \cdot \dots \cdot B(|U_r|)$ elements. It is precisely this sublattice of $\mathbb{P}(S)$ that determines the structure of the convex set Δ_ν , which is now the Cartesian product of r simplices, and of total dimension $(B(|U_1|) - 1) + \dots + (B(|U_r|) - 1)$.

It is clear that the time evolution of a measure $\nu \in \mathcal{M}_+(X)$ under the flow of the recombination equation (7) can thus be reduced to a smaller ODE system, which is fully consistent with our above treatment, due to Eq. (10), as $\Phi(\nu) = \Phi(R_{\mathcal{U}_\nu}(\nu))$ then means that effectively only partitions from the set \mathbb{P}_ν are involved. We leave it to the reader to spell out the details for the modified correspondence and the appropriate initial conditions for the finite-dimensional ODE system.

We are now in the position to approach a solution of the recombination equation.

4. Solution under a linearity assumption. The relatively simple solution structure in the case of single-crossover dynamics (see references [6, 5]) was due to the fact that the nonlinear recombinators acted *linearly* along solutions. Since this is a special case of our general model, via setting $\varrho(\mathcal{A}) = 0$ for any \mathcal{A} that fails to be an ordered partition with (at most) two parts, it is reasonable to consider this point of view also more generally. It will turn out that the linearity assumption is false in general, but we can still learn some interesting things along the way.

Thus, let us assume that also the more general recombinators act linearly along the solution. A simple calculation shows that our coefficient functions then have to satisfy the ODEs

$$\frac{d}{dt} a_t^{\text{lin}}(\mathcal{A}) = -\varrho_{\text{tot}} a_t^{\text{lin}}(\mathcal{A}) + \sum_{\mathcal{B} \succcurlyeq \mathcal{A}} \beta(a_t^{\text{lin}}; \mathcal{A}, \mathcal{B}) \varrho(\mathcal{B}) \quad (16)$$

for all $\mathcal{A} \in \mathbb{P}(S)$, where, for any fixed $q \in \mathcal{M}(\mathbb{P}(S))$,

$$\beta(q; \mathcal{A}, \mathcal{B}) := \begin{cases} \sum_{\mathcal{C} \wedge \mathcal{B} = \mathcal{A}} q(\mathcal{C}), & \text{if } \mathcal{A} \preccurlyeq \mathcal{B}, \\ 0, & \text{otherwise,} \end{cases} \quad (17)$$

is another element of the incidence algebra $\mathbb{A}(\mathbb{P}(S))$. The difference of Eq. (16) to the general equation (14) thus lies in the replacement of the function γ by the significantly simpler linear function β .

Proposition 5. *The Cauchy problem defined by Eq. (16) together with the initial condition $a_0^{\text{lin}}(\mathcal{A}) = \delta(\mathcal{A}, \underline{1})$ has the unique solution given by*

$$a_t^{\text{lin}}(\mathcal{A}) = \sum_{\substack{\mathbb{G} \subseteq \mathbb{P}(S) \\ \bigwedge \mathbb{G} = \mathcal{A}}} \prod_{\sigma \in \mathbb{G}} (1 - e^{-\varrho(\sigma)t}) \prod_{\tau \in \mathbb{G}^c} e^{-\varrho(\tau)t},$$

which, for $t \geq 0$, constitutes a one-parameter family of probability vectors on $\mathbb{P}(S)$.

Proof. Consider the (upper) summatory function $(Fa_t^{\text{lin}})(\mathcal{A}) := \sum_{\mathcal{B} \succcurlyeq \mathcal{A}} a_t^{\text{lin}}(\mathcal{B})$, which is

$$\begin{aligned} (Fa_t^{\text{lin}})(\mathcal{A}) &= \sum_{\substack{\mathbb{G} \subseteq \mathbb{P}(S) \\ \wedge \mathbb{G} \succcurlyeq \mathcal{A}}} \prod_{\sigma \in \mathbb{G}} (1 - e^{-\varrho(\sigma)t}) \prod_{\tau \in \mathbb{G}^c} e^{-\varrho(\tau)t} \\ &= \prod_{\mathcal{B} \notin [\mathcal{A}, \mathbb{1}]} e^{-\varrho(\mathcal{B})t} \sum_{\mathbb{G} \subseteq [\mathcal{A}, \mathbb{1}]} \prod_{\sigma \in \mathbb{G}} (1 - e^{-\varrho(\sigma)t}) \prod_{\tau \in \mathbb{G}^c \cap [\mathcal{A}, \mathbb{1}]} e^{-\varrho(\tau)t} \\ &= \prod_{\mathcal{B} \notin [\mathcal{A}, \mathbb{1}]} e^{-\varrho(\mathcal{B})t} \prod_{\mathcal{C} \in [\mathcal{A}, \mathbb{1}]} (1 - e^{-\varrho(\mathcal{C})t} + e^{-\varrho(\mathcal{C})t}) = \prod_{\mathcal{B} \notin [\mathcal{A}, \mathbb{1}]} e^{-\varrho(\mathcal{B})t} \end{aligned}$$

and hence satisfies the simple ODE

$$\frac{d}{dt} (Fa_t^{\text{lin}})(\mathcal{A}) = - \left(\sum_{\mathcal{B} \notin [\mathcal{A}, \mathbb{1}]} \varrho(\mathcal{B}) \right) (Fa_t^{\text{lin}})(\mathcal{A}), \quad (18)$$

together with the initial condition $(Fa_0^{\text{lin}})(\mathcal{A}) = 1$ for all $\mathcal{A} \in \mathbb{P}(S)$.

On the other hand, for fixed $\mathcal{A}, \mathcal{B} \in \mathbb{P}(S)$ with $\mathcal{A} \preccurlyeq \mathcal{B}$, we have

$$\begin{aligned} \sum_{\mathcal{C} \in [\mathcal{A}, \mathcal{B}]} \beta(a_t^{\text{lin}}; \mathcal{C}, \mathcal{B}) &= \sum_{\mathcal{C} \in [\mathcal{A}, \mathcal{B}]} \sum_{\mathcal{D} \wedge \mathcal{B} = \mathcal{C}} a_t^{\text{lin}}(\mathcal{D}) = \sum_{\mathcal{D} \wedge \mathcal{B} \succcurlyeq \mathcal{A}} a_t^{\text{lin}}(\mathcal{D}) \\ &= \sum_{\mathcal{D} \succcurlyeq \mathcal{A}} a_t^{\text{lin}}(\mathcal{D}) = (Fa_t^{\text{lin}})(\mathcal{A}), \end{aligned}$$

wherefore the summatory function of the right-hand side of Eq. (16), evaluated at \mathcal{A} , becomes

$$-\varrho_{\text{tot}}(Fa_t^{\text{lin}})(\mathcal{A}) + \left(\sum_{\mathcal{B} \succcurlyeq \mathcal{A}} \varrho(\mathcal{B}) \right) (Fa_t^{\text{lin}})(\mathcal{A}) = - \left(\sum_{\mathcal{B} \notin [\mathcal{A}, \mathbb{1}]} \varrho(\mathcal{B}) \right) (Fa_t^{\text{lin}})(\mathcal{A}),$$

which agrees with the right-hand side of Eq. (18). The claim now follows from (upper) Möbius inversion, because, for all $t \geq 0$ and for all $\mathcal{A} \in \mathbb{P}(S)$, one has

$$a_t^{\text{lin}}(\mathcal{A}) = \sum_{\mathcal{B} \succcurlyeq \mathcal{A}} \mu(\mathcal{A}, \mathcal{B}) (Fa_t^{\text{lin}})(\mathcal{B}), \quad (19)$$

where $\mu \in \mathbb{A}(\mathbb{P}(S))$ denotes the Möbius function for the lattice $\mathbb{P}(S)$.

The fact that a_t , for each $t \geq 0$, is a probability vector on $\mathbb{P}(S)$ follows from our earlier calculation in Eq. (5), with $f(x) = e^x$. \square

This approach clearly has a lattice-theoretic basis, which lends itself to a number of interesting further aspects and insights [9].

Remark 6. Let us note that Eqs. (18) and (19) allow a re-interpretation of the coefficient formula from Proposition 5 as

$$a_t^{\text{lin}}(\mathcal{A}) = \sum_{\mathcal{B} \succcurlyeq \mathcal{A}} \mu(\mathcal{A}, \mathcal{B}) e^{-\chi(\mathcal{B})t} \quad (20)$$

with $\chi(\mathcal{B}) = \sum_{\mathcal{C} \notin [\mathcal{B}, \mathbb{1}]} \varrho(\mathcal{C})$ being the *decay rate* of the corresponding (exponential) term. This relation means $\chi(\mathcal{B}) = \varrho_{\text{tot}} - \sum_{\mathcal{C} \succcurlyeq \mathcal{B}} \varrho(\mathcal{C})$, so that the recombination rates are obtained from the decay rates by means of (upper) Möbius inversion as

$$\varrho(\mathcal{B}) = \delta(\mathcal{B}, \mathbb{1}) \varrho_{\text{tot}} - \sum_{\mathcal{C} \succcurlyeq \mathcal{B}} \mu(\mathcal{B}, \mathcal{C}) \chi(\mathcal{C}),$$

where it is assumed that the total recombination rate ϱ_{tot} is known. This detail corresponds to the fact that $\varrho(\underline{1})$ does not contribute to any of the $\chi(\mathcal{A})$.

Either version of a_t^{lin} permits the determination of the asymptotic properties of the coefficients a_t^{lin} as $t \rightarrow \infty$, and hence that of the measure ω_t . In particular, when the partition subset $\mathbb{G} := \{\mathcal{A} \in \mathbb{P}(S) \mid \varrho(\mathcal{A}) > 0\}$ satisfies $\bigwedge \mathbb{G} = \underline{0}$, one has

$$\lim_{t \rightarrow \infty} \omega_t = R_0(\omega_0),$$

with convergence in the $\|\cdot\|$ -topology. This is the obvious generalisation of the known asymptotic properties in the special cases treated in [11, 6]. \diamond

Remark 7. Let us briefly mention that the linear system of ODEs defined by Eq. (16) can still be solved when the recombination rates become time-dependent, as was previously observed in [5, Addendum] for single-crossover recombination. Indeed, if all $\varrho_t(\mathcal{A})$ are non-negative functions of time (which is needed to preserve all claims of Proposition 4 also for Eq. (16) with time-dependent rates), the solution formula from Proposition 5 becomes

$$a_t^{\text{lin}}(\mathcal{A}) = \sum_{\substack{\mathbb{G} \subseteq \mathbb{P}(S) \\ \bigwedge \mathbb{G} = \mathcal{A}}} \prod_{\sigma \in \mathbb{G}} \left(1 - \exp(-\int_0^t \varrho_s(\sigma) ds)\right) \prod_{\tau \in \mathbb{G}^c} \exp(-\int_0^t \varrho_s(\tau) ds),$$

for the same initial conditions. The proof is completely analogous to that of Proposition 5, now with

$$(Fa_t^{\text{lin}}(\mathcal{A})) = \prod_{\mathcal{B} \notin [\mathcal{A}, \underline{1}]} \exp(-\int_0^t \varrho_s(\mathcal{B}) ds).$$

This summatory function then satisfies the ODE

$$\frac{d}{dt}(Fa_t^{\text{lin}})(\mathcal{A}) = -\left(\sum_{\mathcal{B} \notin [\mathcal{A}, \underline{1}]} \varrho_t(\mathcal{B})\right) (Fa_t^{\text{lin}})(\mathcal{A}),$$

which replaces Eq. (18) in this generalisation. \diamond

Theorem 2. *Let $\mathcal{A} \in \mathbb{P}(S)$ be $\mathcal{A} = \underline{1}$ or a partition into two parts, one of which is a singleton set. Then, the coefficient formula from Proposition 5 gives the correct solution also for the Cauchy problem of the general recombination equation from Eq. (14).*

Proof. Let a probability vector q on $\mathbb{P}(S)$ be given. The definitions of β and γ from Eqs. (15) and (17) imply, via a simple calculation, that

$$\gamma(q; \mathcal{A}, \underline{1}) = q(\mathcal{A}) = \beta(q; \mathcal{A}, \underline{1})$$

holds for any $\mathcal{A} \in \mathbb{P}(S)$. This gives the claim for $\mathcal{A} = \underline{1}$, because Eqs. (14) and (16) are equal in this case, hence $a_t(\underline{1}) = a_t^{\text{lin}}(\underline{1})$.

More generally, when $\mathcal{A} = \{\{i\}, S \setminus \{i\}\}$ for some $i \in S$, one finds that

$$\gamma(q; \mathcal{A}, \mathcal{A}) = q(\underline{1}) + q(\mathcal{A}) = \beta(q; \mathcal{A}, \mathcal{A}),$$

so that the ODEs from Eqs. (14) and (16) coincide also for such partitions \mathcal{A} , which proves the second claim. \square

Corollary 2. *The coefficient formula from Proposition 5 gives the correct solution of Eq. (14) for all $\mathcal{A} \in \mathbb{P}(S)$ whenever S is a finite set with $1 \leq |S| \leq 3$.*

Proof. There is nothing to prove for $|S| = 1$. The claim for $|S| = 2$ is obvious, and also follows from [5, Prop. 3]. When $|S| = 3$, all partitions except $\mathcal{A} = \emptyset$ satisfy the conditions of Theorem 2. Since $a_t(0) = 1 - \sum_{\mathcal{B} \neq \emptyset} a_t(\mathcal{B})$, the claim follows. \square

Remark 8. In the special situation of single-crossover recombination, where we have $\varrho(\mathcal{A}) > 0$ only for *ordered* partitions \mathcal{A} into two parts, the solution formula of Proposition 5 reduces to the known solution for this case from [6, 5]. In particular, the linearity assumption is satisfied, and the solution holds for all system sizes and all values of the single-crossover rates. \diamond

Note, however, that already for $|S| = 4$, when we are beyond the single-crossover case, the coefficients a_t and a_t^{lin} can differ, for instance for $\mathcal{A} = \{\{1, 2\}, \{3, 4\}\}$, which is a biologically relevant partition. We thus need to proceed without the linearisation assumption.

5. General case: Marginalisation consistency. It is clear that our general recombination equation can only be considered a reasonable model if it is *marginalisation consistent*. By this we mean that the restriction to a subsystem, via appropriate marginalisation, gives a solution of the recombination equation for the subsystem. We now discuss this in more detail, and then establish this consistency property for our model, both in the measure-theoretic and in the finite-dimensional version. The latter case will depend on an interesting interplay between elements of the Möbius and the incidence algebras at hand.

Let S be as above, or any other finite set with n elements, and consider a subsystem as specified by $\emptyset \neq U \subset S$. When ω_t^S is the solution of the general recombination equation (7) with initial condition $\omega_0^S \in \mathcal{M}_+(X)$ according to Proposition 3, it is natural to define

$$\omega_t^U = \pi_U \cdot \omega_t^S \quad (21)$$

as the corresponding (marginalised) measure for the subsystem defined by U . Then, recalling that the projector π_U is linear, we get

$$\begin{aligned} \frac{d}{dt} \omega_t^U &= \pi_U \cdot \left(\frac{d}{dt} \omega_t^S \right) = \pi_U \cdot \left(\sum_{\mathcal{B} \in \mathbb{P}(S)} \varrho^S(\mathcal{B}) (R_{\mathcal{B}}^S - \mathbb{1})(\omega_t^S) \right) \\ &= \sum_{\mathcal{B} \in \mathbb{P}(S)} \varrho^S(\mathcal{B}) \left(\pi_U \cdot (R_{\mathcal{B}}^S(\omega_t^S)) - \omega_t^U \right) = \sum_{\mathcal{B} \in \mathbb{P}(S)} \varrho^S(\mathcal{B}) \left(R_{\mathcal{B}|_U}^U(\omega_t^U) - \omega_t^U \right) \\ &= \sum_{\mathcal{A} \in \mathbb{P}(U)} \sum_{\substack{\mathcal{B} \in \mathbb{P}(S) \\ \mathcal{B}|_U = \mathcal{A}}} \varrho^S(\mathcal{B}) \left(R_{\mathcal{A}}^U(\omega_t^U) - \omega_t^U \right) = \sum_{\mathcal{A} \in \mathbb{P}(U)} \varrho^U(\mathcal{A}) (R_{\mathcal{A}}^U - \mathbb{1})(\omega_t^U), \end{aligned}$$

where the fourth step is an application of Lemma 1, while the last step anticipates the definition of the *induced* (or *marginal*) *recombination rates* for the subsystem as

$$\varrho^U(\mathcal{A}) := \sum_{\substack{\mathcal{D} \in \mathbb{P}(S) \\ \mathcal{D}|_U = \mathcal{A}}} \varrho^S(\mathcal{D}) \quad (22)$$

for any $\mathcal{A} \in \mathbb{P}(U)$. It is obvious that non-negativity of the rates ϱ^S implies that of the marginal rates ϱ^U . Our little calculation proves the following result.

Proposition 6. *Let S be a finite set and $U \subset S$ a non-empty subset. If ω_t^S is a solution of the recombination equation (7), with recombination rates $\varrho^S(\mathcal{B}) \geq 0$*

for $\mathcal{B} \in \mathbb{P}(S)$ and with initial condition $\omega_0^S \in \mathcal{M}_+(X)$, the marginalised measure ω_t^U from Eq. (21) solves the recombination equation for the subsystem defined by U , provided the recombination rates $\rho^U(\mathcal{A})$ for $\mathcal{A} \in \mathbb{P}(U)$ are defined according to Eq. (22). \square

Remark 9. The problem of marginalisation consistency has been observed early on in mathematical population genetics. Ewens and Thomson [16] have tackled it in 1977 in models that describe the combined action of recombination *and selection* for the two-parent case in discrete time; see also the review in [11, pp. 69–72]. The corresponding dynamics is, in general, *not* marginalisation consistent. However, it is obvious from the calculations in [16] that consistency does apply in the case without selection, and the dynamics is then governed by the marginal recombination probabilities, which are the two-parent, discrete-time analogues of our marginal recombination rates. \diamond

Let us see how this result translates to the finite-dimensional ODE systems at the level of the coefficient functions $a_t^U(\mathcal{A})$. Given a general probability vector $q = q^S$ on $\mathbb{P}(S)$, we define the corresponding *marginal probabilities* via

$$q^U(\mathcal{A}) := \sum_{\substack{\mathcal{D} \in \mathbb{P}(S) \\ \mathcal{D}|_U = \mathcal{A}}} q^S(\mathcal{D}), \quad (23)$$

for any partition $\mathcal{A} \in \mathbb{P}(U)$, in complete analogy to Eq. (22). Clearly, one has $q^U(\mathcal{A}) \geq 0$, while proper normalisation follows from

$$\sum_{\mathcal{A} \in \mathbb{P}(U)} q^U(\mathcal{A}) = \sum_{\mathcal{A} \in \mathbb{P}(U)} \sum_{\substack{\mathcal{D} \in \mathbb{P}(S) \\ \mathcal{D}|_U = \mathcal{A}}} q^S(\mathcal{D}) = \sum_{\mathcal{B} \in \mathbb{P}(S)} \sum_{\substack{\mathcal{C} \in \mathbb{P}(U) \\ \mathcal{C} = \mathcal{B}|_U}} q^S(\mathcal{B}) = \sum_{\mathcal{B} \in \mathbb{P}(S)} q^S(\mathcal{B}) = 1, \quad (24)$$

where the penultimate step follows from the observation that the inner sum over \mathcal{C} consists of precisely one term because the restriction of a partition to a subset $U \subset S$ is unique.

Let us note for later use that, for $\emptyset \neq U \subseteq V \subseteq S$ and $\mathcal{A} \in \mathbb{P}(U)$, we also have

$$q^U(\mathcal{A}) = \sum_{\substack{\mathcal{E} \in \mathbb{P}(V) \\ \mathcal{E}|_U = \mathcal{A}}} q^V(\mathcal{E}), \quad (25)$$

because

$$\sum_{\substack{\mathcal{E} \in \mathbb{P}(V) \\ \mathcal{E}|_U = \mathcal{A}}} q^V(\mathcal{E}) = \sum_{\substack{\mathcal{E} \in \mathbb{P}(V) \\ \mathcal{E}|_U = \mathcal{A}}} \sum_{\substack{\mathcal{D} \in \mathbb{P}(S) \\ \mathcal{D}|_V = \mathcal{E}}} q^S(\mathcal{D}) = \sum_{\substack{\mathcal{D} \in \mathbb{P}(S) \\ \mathcal{D}|_U = \mathcal{A}}} q^S(\mathcal{D}) = q^U(\mathcal{A}).$$

As special cases of Eqs. (23) and (25), let us also note that, for $\mathcal{B} \in \mathbb{P}(S)$ and $\emptyset \neq U \subseteq V \subseteq S$, one has

$$q^U(\mathcal{B}|_U) = \sum_{\substack{\mathcal{D} \in \mathbb{P}(S) \\ \mathcal{D}|_U = \mathcal{B}|_U}} q^S(\mathcal{D}) = \sum_{\substack{\mathcal{E} \in \mathbb{P}(V) \\ \mathcal{E}|_U = \mathcal{B}|_U}} q^V(\mathcal{E}). \quad (26)$$

Remark 10. The (upper) summatory function defined by $(Fq^U)(\mathcal{A}) = \sum_{\mathcal{B} \succ \mathcal{A}} q^U(\mathcal{B})$ can also be calculated by marginalisation, namely as

$$\begin{aligned} (Fq^U)(\mathcal{A}) &= \sum_{\mathcal{B} \succ \mathcal{A}} q^U(\mathcal{B}) = \sum_{\mathcal{B} \succ \mathcal{A}} \sum_{\substack{\mathcal{D} \in \mathbb{P}(S) \\ \mathcal{D}|_U = \mathcal{B}}} q^S(\mathcal{D}) = \sum_{\substack{\mathcal{D} \in \mathbb{P}(S) \\ \mathcal{D}|_U \succ \mathcal{A}}} q^S(\mathcal{D}) \\ &= \sum_{\substack{\mathcal{C} \in \mathbb{P}(S) \\ \mathcal{C}|_U = \mathcal{A}}} \sum_{\mathcal{D} \succ \mathcal{C}} q^S(\mathcal{D}) = \sum_{\substack{\mathcal{C} \in \mathbb{P}(S) \\ \mathcal{C}|_U = \mathcal{A}}} (Fq^S)(\mathcal{C}), \end{aligned}$$

wherefore consistency at this level is obvious. \diamond

Let us return to the recombination rates $\varrho^S(\mathcal{B})$ for $\mathcal{B} \in \mathbb{P}(S)$ together with

$$\varrho_{\text{tot}}^S = \sum_{\mathcal{B} \in \mathbb{P}(S)} \varrho^S(\mathcal{B}),$$

compare Remark 2, and consider the marginal rates $\varrho^U(\mathcal{A})$ with $\mathcal{A} \in \mathbb{P}(U)$ and $\emptyset \neq U \subseteq S$. Repeating the above calculation of the normalisation condition, one finds

$$\varrho_{\text{tot}}^U = \varrho_{\text{tot}}^S = \varrho_{\text{tot}}, \quad (27)$$

as it should be. The total recombination rate is thus the same on all levels, and independent of U , as it must. Note that, in this process, $\varrho^U(\{U\}) = \varrho^U(\underline{1})$ need not vanish. This does not matter because the corresponding recombinator (on the subsystem) is the identity, and hence does not affect the solution; compare Remark 2.

Let us now see how the result of Proposition 6 translates to properties of the coefficients for the ODE system and its subsystems. Here, the desired marginalisation consistency will depend on the following slightly technical, but somewhat surprising identity.

Lemma 3. *Let $U \subseteq S$ be as before, and let $\mathcal{A}, \mathcal{B} \in \mathbb{P}(U)$ with $\mathcal{A} \preceq \mathcal{B}$ be arbitrary, but fixed. Then, for any $\mathcal{D} \in \mathbb{P}(S)$ with $\mathcal{D}|_U = \mathcal{B}$, one has the product representation and reduction relation*

$$\sum_{\substack{\mathbb{P}(S) \ni \mathcal{C} \preceq \mathcal{D} \\ \mathcal{C}|_U = \mathcal{A}}} \gamma^S(q^S; \mathcal{C}, \mathcal{D}) = \sum_{\substack{\mathbb{P}(S) \ni \mathcal{C} \preceq \mathcal{D} \\ \mathcal{C}|_U = \mathcal{A}}} \prod_{i=1}^{|\mathcal{D}|} q^{D_i}(\mathcal{C}|_{D_i}) = \prod_{i=1}^{|\mathcal{B}|} q^{B_i}(\mathcal{A}|_{B_i}) = \gamma^U(q^U; \mathcal{A}, \mathcal{B}),$$

where q^S is any probability vector on $\mathbb{P}(S)$, q^U its marginalisation according to Eq. (23), and γ^U is defined as in Eq. (12), but with S replaced by U .

Proof. Let $\mathcal{A}, \mathcal{B} \in \mathbb{P}(U)$ with $\mathcal{A} \preceq \mathcal{B}$ be given, and assume $\mathcal{B} = \{B_1, \dots, B_r\}$. With our definition of the γ -function in Eq. (12) together with Eqs. (26) and (23), the right-hand side evaluates as

$$\gamma^U(q^U; \mathcal{A}, \mathcal{B}) = \prod_{i=1}^r \sum_{\substack{\mathcal{C} \in \mathbb{P}(U) \\ \mathcal{C}|_{B_i} = \mathcal{A}|_{B_i}}} q^U(\mathcal{C}) = \prod_{i=1}^r q^{B_i}(\mathcal{A}|_{B_i}). \quad (28)$$

For an arbitrary $\mathcal{D} \in \mathbb{P}(S)$ with $\mathcal{D} = \{D_1, \dots, D_{|\mathcal{D}|}\}$ and $\mathcal{D}|_U = \mathcal{B}$, this is now to be compared with

$$\sum_{\substack{\mathbb{P}(S) \ni \mathcal{C} \preceq \mathcal{D} \\ \mathcal{C}|_U = \mathcal{A}}} \gamma^S(q^S; \mathcal{C}, \mathcal{D}) = \sum_{\substack{\mathbb{P}(S) \ni \mathcal{C} \preceq \mathcal{D} \\ \mathcal{C}|_U = \mathcal{A}}} \prod_{i=1}^{|\mathcal{D}|} \sum_{\substack{\mathcal{E} \in \mathbb{P}(S) \\ \mathcal{E}|_{D_i} = \mathcal{C}|_{D_i}}} q^S(\mathcal{E}) = \sum_{\substack{\mathbb{P}(S) \ni \mathcal{C} \preceq \mathcal{D} \\ \mathcal{C}|_U = \mathcal{A}}} \prod_{i=1}^{|\mathcal{D}|} q^{D_i}(\mathcal{C}|_{D_i}), \quad (29)$$

where the last step is once again a consequence of Eq. (23). To proceed, we rewrite the partition \mathcal{D} as $\mathcal{D} = \{D'_1, \dots, D'_r, D''_1, \dots, D''_s\}$ with $r = |\mathcal{B}|$ and $D'_i \cap U = B_i$ for $1 \leq i \leq r$ (which is without loss of generality) and $D''_j \cap U = \emptyset$ for $1 \leq j \leq s$. Note that $s = 0$ is possible, in which case no D''_j is present and $r = |\mathcal{D}|$.

Now, any partition $\mathcal{C} \in \mathbb{P}(S)$ in the summation in Eq. (29) must be a joining of the form

$$\mathcal{C} = \mathcal{C}'_1 \sqcup \mathcal{C}'_2 \sqcup \dots \sqcup \mathcal{C}'_r \sqcup \mathcal{C}''_1 \sqcup \dots \sqcup \mathcal{C}''_s$$

with $\mathcal{C}'_i \in \mathbb{P}(D'_i)$ and $\mathcal{C}''_j \in \mathbb{P}(D''_j)$, subject to the additional condition that we always have $\mathcal{C}|_U = \mathcal{A}$, which means $\mathcal{C}'_i|_{B_i} = \mathcal{A}|_{B_i}$ for all $1 \leq i \leq r$. The summation on the right-hand side of Eq. (29) can now be broken into smaller sums that can be absorbed into the factors, which amounts to refining each block of \mathcal{D} individually. This turns the right-hand side of Eq. (29) into a product of two terms, namely

$$\left(\prod_{i=1}^r \sum_{\substack{\sigma \in \mathbb{P}(D'_i) \\ \sigma|_{B_i} = \mathcal{A}|_{B_i}}} q^{D'_i}(\sigma) \right) \left(\prod_{j=1}^s \sum_{\tau \in \mathbb{P}(D''_j)} q^{D''_j}(\tau) \right).$$

Now, the sum in each factor of the second product is clearly 1 because $q^{D''_j}$ is a probability vector on $\mathbb{P}(D''_j)$, compare Eq. (24). Consequently, the entire second term is 1, which is also true if $s = 0$ (in which case we have the empty product here). Likewise, the sum in the i th factor of the first product equals $q^{B_i}(\mathcal{A}|_{B_i})$ by Eq. (26), because $B_i \subseteq D'_i$ by our assumptions. We thus get

$$\sum_{\substack{\mathbb{P}(S) \ni \mathcal{C} \preceq \mathcal{D} \\ \mathcal{C}|_U = \mathcal{A}}} \gamma^S(q^S; \mathcal{C}, \mathcal{D}) = \prod_{i=1}^r q^{B_i}(\mathcal{A}|_{B_i}).$$

Together with Eqs. (28) and (29), this proves the lemma. \square

The marginalisation consistency can now be stated as follows.

Proposition 7. *Let S be a finite set and $U \subset S$ a non-empty subset. If the family of probability vectors $\{a_t^S \mid t \geq 0\}$ is a solution of the Cauchy problem of Eq. (14) with initial condition $a_0^S(\mathcal{B}) = \delta^S(\mathcal{B}, \mathbb{1})$, the marginalised family $\{a_t^U \mid t \geq 0\}$, with $a_t^U(\mathcal{A})$ defined according to Eq. (23) for all $\mathcal{A} \in \mathbb{P}(U)$, solves the corresponding Cauchy problem for the subsystem, with initial condition $a_0^U(\mathcal{A}) = \delta^U(\mathcal{A}, \mathbb{1})$ and the marginal recombination rates of Eq. (22). Explicitly, it satisfies the ODE*

$$\dot{a}_t^U(\mathcal{A}) = -(\varrho_{\text{tot}} - \varrho(\mathbb{1}))a_t^U(\mathcal{A}) + \sum_{\mathcal{A} \preceq \mathcal{B} \prec \mathbb{1}} \varrho^U(\mathcal{B}) \prod_{i=1}^{|\mathcal{B}|} a_t^{B_i}(\mathcal{A}|_{B_i}). \quad (30)$$

The analogous statement remains true for a general probability vector a_0^S as initial condition, with the marginalised initial condition a_0^U according to Eq. (23) on the subsystem.

Proof. Let $\{a_t^S \mid t \geq 0\}$ be a solution of Eq. (14), and let $\mathcal{A} \in \mathbb{P}(U)$ be fixed. Then, we have

$$\begin{aligned}
\dot{a}_t^U(\mathcal{A}) &= \sum_{\substack{\mathcal{C} \in \mathbb{P}(S) \\ \mathcal{C}|_U = \mathcal{A}}} \dot{a}_t^S(\mathcal{C}) = \sum_{\substack{\mathcal{C} \in \mathbb{P}(S) \\ \mathcal{C}|_U = \mathcal{A}}} \left(-\varrho_{\text{tot}} a_t^S(\mathcal{C}) + \sum_{\mathcal{D} \succ \mathcal{C}} \gamma^S(a_t^S; \mathcal{C}, \mathcal{D}) \varrho^S(\mathcal{D}) \right) \\
&= -\varrho_{\text{tot}} a_t^U(\mathcal{A}) + \sum_{\substack{\mathcal{C} \in \mathbb{P}(S) \\ \mathcal{C}|_U = \mathcal{A}}} \sum_{\mathcal{D} \succ \mathcal{C}} \gamma^S(a_t^S; \mathcal{C}, \mathcal{D}) \varrho^S(\mathcal{D}) \\
&= -\varrho_{\text{tot}} a_t^U(\mathcal{A}) + \sum_{\substack{\mathcal{D} \in \mathbb{P}(S) \\ \mathcal{D}|_U \succ \mathcal{A}}} \left(\sum_{\substack{\mathcal{C} \preccurlyeq \mathcal{D} \\ \mathcal{C}|_U = \mathcal{A}}} \gamma^S(a_t^S; \mathcal{C}, \mathcal{D}) \right) \varrho^S(\mathcal{D}) \\
&= -\varrho_{\text{tot}} a_t^U(\mathcal{A}) + \sum_{\mathcal{B} \succ \mathcal{A}} \sum_{\substack{\mathcal{D} \in \mathbb{P}(S) \\ \mathcal{D}|_U = \mathcal{B}}} \varrho^S(\mathcal{D}) \sum_{\substack{\mathcal{C} \preccurlyeq \mathcal{D} \\ \mathcal{C}|_U = \mathcal{A}}} \gamma^S(a_t^S; \mathcal{C}, \mathcal{D}) \\
&= -\varrho_{\text{tot}} a_t^U(\mathcal{A}) + \sum_{\mathcal{B} \succ \mathcal{A}} \gamma^U(a_t^U; \mathcal{A}, \mathcal{B}) \sum_{\substack{\mathcal{D} \in \mathbb{P}(S) \\ \mathcal{D}|_U = \mathcal{B}}} \varrho^S(\mathcal{D}) \\
&= -(\varrho_{\text{tot}} - \varrho(\underline{1})) a_t^U(\mathcal{A}) + \sum_{\mathcal{A} \preccurlyeq \mathcal{B} \prec \underline{1}} \varrho^U(\mathcal{B}) \prod_{i=1}^{|\mathcal{B}|} a_t^{B_i}(\mathcal{A}|_{B_i}),
\end{aligned}$$

where we have used Lemma 3 in the second-last step. This made the γ -term independent of \mathcal{D} , which in turn allowed the last step on the basis of Eq. (22) and another application of Lemma 3. The second-last step shows that the marginalised family indeed satisfies the proper ODE for the subsystem as defined for $\emptyset \neq U \subset S$ via Eq. (14), as it must; the last step then leads to Eq. (30).

It is an easy exercise that $\delta^U(\mathcal{A}, \underline{1})$ is the initial condition for the subsystem that emerges as the marginalisation of the original initial condition $\delta^S(\mathcal{B}, \underline{1})$. Since $\mathcal{A} \in \mathbb{P}(U)$ was arbitrary, the main statement is proved. The last claim is an obvious generalisation. \square

Note that Proposition 7 can also be viewed as a consequence of Proposition 6 and Theorem 1. We have nevertheless opted for an explicit verification because several steps of the above proof will reappear when we proceed to a solution of the recombination equation.

6. The backward point of view: Partitioning process. Now that we have understood the structure of the ODE for a_t in the usual (forward) direction of time, let us consider a related (stochastic) process that will provide an additional meaning for a_t . Let $\{\Sigma_t^S\}_{t \geq 0}$ be a Markov chain in continuous time with values in $\mathbb{P}(S)$ that is constructed as follows. Start with $\Sigma_0^S = \underline{1}$. If the current state is $\Sigma_t^S = \mathcal{C}$, then part C_i of \mathcal{C} is replaced by $\sigma \in \mathbb{P}(C_i)$ at rate $\varrho^{C_i}(\sigma)$, independently of all other parts. That is, the transition from \mathcal{C} to $(\mathcal{C} \setminus C_i) \sqcup \sigma$ occurs at rate $\varrho^{C_i}(\sigma)$ for for all $\underline{1} \neq \sigma \in \mathbb{P}(C_i)$ and $1 \leq i \leq |\mathcal{C}|$. Obviously, $\{\Sigma_t^S\}_{t \geq 0}$ is a process of progressive refinements, which we call the *partitioning process*. Likewise, we define the partitioning process $\{\Sigma_t^U\}_{t \geq 0}$ on $\mathbb{P}(U)$ for $\emptyset \neq U \subset S$ in the same way as $\{\Sigma_t^S\}_{t \geq 0}$, but based on the marginal recombination rates ϱ^U .

Now let $P_t^U(\mathcal{C}, \mathcal{D}) := \mathbf{P}(\Sigma_t^U = \mathcal{D} \mid \Sigma_0^U = \mathcal{C})$, where \mathbf{P} denotes probability. That is, $P_t^U(\mathcal{C}, \mathcal{D})$ is the transition probability (in standard notation) from ‘state’ \mathcal{C} to

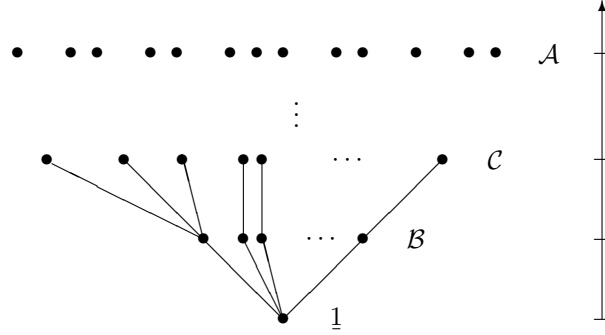


FIGURE 2. A sketch of the partitioning process. Each dot represents one part of the corresponding partition, while the lines indicate the partitioning.

‘state’ \mathcal{D} during a time interval of length t . Since $\{\Sigma_t^U\}_{t \geq 0}$ is a process of progressive refinements, it is obvious that

$$P_t^U(\mathcal{C}, \mathcal{D}) = 0 \quad \text{for } \mathcal{D} \not\preceq \mathcal{C}. \quad (31)$$

Furthermore, since the parts are (conditionally) independent once they appear, we have

$$P_t^U(\mathcal{C}, \mathcal{D}) = \prod_{i=1}^{|\mathcal{C}|} P_t^{C_i}(\mathbb{1}, \mathcal{D}|_{C_i}) \quad \text{for } \mathcal{D} \preceq \mathcal{C}. \quad (32)$$

Let us now consider the actual distribution of Σ_t^U , that is, the collection $\{P_t^U(\mathbb{1}, \mathcal{A})\}_{\mathcal{A} \in \mathbb{P}(U)}$. Clearly, the initial value is $P_0^U(\mathbb{1}, \mathcal{A}) = \delta(\mathcal{A}, \mathbb{1})$. The time evolution is given by

$$\begin{aligned} \frac{d}{dt} P_t^U(\mathbb{1}, \mathcal{A}) &= -(\varrho_{\text{tot}} - \varrho^U(\mathbb{1})) P_t^U(\mathbb{1}, \mathcal{A}) + \sum_{\mathcal{B} \prec \mathbb{1}} \varrho^U(\mathcal{B}) P_t^U(\mathcal{B}, \mathcal{A}) \\ &= -(\varrho_{\text{tot}} - \varrho^U(\mathbb{1})) P_t^U(\mathbb{1}, \mathcal{A}) + \sum_{\mathcal{A} \preceq \mathcal{B} \prec \mathbb{1}} \varrho^U(\mathcal{B}) \prod_{i=1}^{|\mathcal{B}|} P_t^{B_i}(\mathbb{1}, \mathcal{A}|_{B_i}). \end{aligned} \quad (33)$$

Here, the first step is an application of the *Kolmogorov backward equation* (for background, see [17, Ch. XVII.8] or [21, Ch. 2.1]), namely, the decomposition according to the first transition away from $\mathbb{1}$, which is to state \mathcal{B} with rate $\varrho^U(\mathcal{B})$ for all $\mathcal{B} \prec \mathbb{1}$. The second step uses Eqs. (31) and (32). The argument is illustrated in Figure 2.

Comparing Eq. (33) with Eq. (30), we see that the quantities $P_t^U(\mathbb{1}, \mathcal{A})$ and $a_t^U(\mathcal{A})$, with $U \subseteq S$ and $\mathcal{A} \in \mathbb{P}(U)$, satisfy the same collection of ODEs, and the initial values also agree at $P_0^U(\mathbb{1}, \mathcal{A}) = \delta_{\mathcal{A}, \mathbb{1}} = a_0^U(\mathcal{A})$. By an obvious inductive argument, the two families can thus be identified. We have therefore shown the following result.

Theorem 3. *The probability vector a_t^U agrees with the distribution of the partitioning process Σ_t^U . Explicitly, we have*

$$a_t^U(\mathcal{A}) = P_t^U(\mathbb{1}, \mathcal{A}) = \mathbf{P}(\Sigma_t^U = \mathcal{A} \mid \Sigma_0^U = \mathbb{1})$$

for all $U \subseteq S$, $\mathcal{A} \in \mathbb{P}(U)$ and all $t \geq 0$. \square

As the reader may have noticed, the partitioning process was defined so as to reflect the action of recombination on the ancestry of the genetic material of an individual *backward in time*. Namely, if a sequence is pieced together according to a partition $\mathcal{A} = \{A_1, \dots, A_r\}$ from various parents forwards in time, this implies that the sequence is partitioned into the parts of \mathcal{A} when we look backwards in time, where each part is associated with a different parent. In this light, Theorem 3 means the following: If we follow the ancestry of the genetic material of an individual from the present population (that is, starting at time t) backward in time, then $a_t^S(\mathcal{A})$ is the probability that the sites are partitioned into different parents according to \mathcal{A} at time t before the present. That is, the sites in A_1 go back to one individual in the initial population, the sites in A_2 go back to a second individual and so on, and the sites in A_r go back to an r th individual. The partitioning process is the deterministic limit of the corresponding stochastic process in finite populations, namely, the *ancestral recombination graph* (ARG); see [14, Ch. 3.4], but note that the ARG is usually described for single crossovers only.

At the same time, the above suggests a nice interpretation of the solution of the recombination equation in Theorem 1. Indeed, the type distribution at present may be obtained in a two-step procedure. In the first step, one decides how the sites of a present individual have been pieced together from different parents from the initial population. In the second step, one assigns letters to the various parts (that is, parents): If the corresponding partition is $\mathcal{A} = \{A_1, \dots, A_r\}$, then the letters for the sites in A_1 are drawn from $\pi_{A_1} \cdot \omega_0$, and so on, until the letters for the sites in A_r are drawn from $\pi_{A_r} \cdot \omega_0$ — and this independently, due to the infinite population. The resulting type distribution of our individual today is $R_{\mathcal{A}}(\omega_0)$. What we have described here in words amounts to a *duality relation* between the recombination dynamics forward in time and the partitioning process backward in time. We do not go into detail here; for more, see [8] and [15].

7. Generic case: Recursive solution. Motivated by the Möbius formula in Eq. (20) and by a reminiscent structure in the solution of the discrete recombination equation, compare [3], we now restrict Eq. (14) to the forward-invariant simplex $\mathcal{P}(\mathbb{P}(S))$, compare Proposition 4, and make the *ansatz*

$$a_t^U(\mathcal{A}) = \sum_{\mathcal{B} \neq \mathcal{A}} \theta^U(\mathcal{A}, \mathcal{B}) e^{-\psi^U(\mathcal{B})t} \quad (34)$$

for any non-empty $U \subseteq S$, with decay rates ψ^U and coefficient functions θ^U . Our aim is to determine θ^U by a comparison of coefficients once we know a (recursive) formula for the rate function ψ^U together with linear independence of the exponentials in Eq. (34). Note that this will then also solve the ODE (14) with an initial condition from $\mathcal{M}_+(\mathbb{P}(S))$ via multiplying the right-hand side of Eq. (34) by the norm of the initial condition.

From our results of Section 4, see Theorem 2 and Remark 6 in comparison with Eq. (34), we know that

$$\psi^U(\mathbb{1}) = \chi^U(\mathbb{1}) = \sum_{\mathcal{A} \neq \mathbb{1}} \varrho^U(\mathcal{A}) = \varrho_{\text{tot}} - \varrho^U(\mathbb{1}). \quad (35)$$

Note that we have used the fact that $\mu^U(\mathbb{1}, \mathbb{1}) = \theta^U(\mathbb{1}, \mathbb{1}) = 1$ for all $\emptyset \neq U \subseteq S$, which is obvious. In particular, this gives $\psi^U(\mathcal{A}) = 0$ for any U with $|U| = 1$, which is the trivial limiting case with an empty sum on the right-hand side. All other

values are now defined recursively by

$$\psi^U(\mathcal{A}) := \sum_{i=1}^{|\mathcal{A}|} \psi^{A_i}(\mathbb{1}), \quad (36)$$

which is motivated by the well-known eigenvalues in the two-parent, discrete-time analogue of our model; see [19, Theorem 6.4.3] or [20]. In the context of our partitioning process, $\psi^{A_i}(\mathbb{1})$ is the total rate of any further partitioning of part A_i , and so, due to the independence of the parts, $\psi^U(\mathcal{A})$ is the total rate of transitions out of state \mathcal{A} . For any U with $|U| = 2$, one has $\psi^U(\mathcal{A}) = \delta^U(\mathcal{A}, \mathbb{1})$, while one finds $\psi^U(\mathbb{0}) = 0$ for all non-empty $U \subseteq S$. This is consistent with the fact that no further partitioning (or ‘decay’) is possible when starting from the partition $\mathbb{0}$ of U .

The decay rates defined this way have important summation properties as follows.

Lemma 4. *Let S be a finite set and let $\emptyset \neq U \subseteq S$ be arbitrary, but fixed. If $\mathcal{A} = \{A_1, \dots, A_r\} \in \mathbb{P}(U)$, with $r = |\mathcal{A}|$, one has $\psi^U(\mathcal{A}) + \sum_{i=1}^r \varrho^{A_i}(\mathbb{1}) = r \varrho_{\text{tot}}$. If $s_{\mathcal{A}}$ is the number of the parts of \mathcal{A} that are singletons, this can be simplified as*

$$\psi^U(\mathcal{A}) + \sum_{\substack{i=1 \\ |A_i| > 1}}^r \varrho^{A_i}(\mathbb{1}) = (r - s_{\mathcal{A}}) \varrho_{\text{tot}}.$$

In particular, $\psi^U(\mathbb{0}) = 0$ holds for all non-empty $U \subseteq S$.

Moreover, for arbitrary $\mathcal{B}, \mathcal{C} \in \mathbb{P}(U)$ with $\mathcal{C} \preceq \mathcal{B}$, one has

$$\psi^U(\mathcal{C}) = \sum_{i=1}^{|\mathcal{B}|} \psi^{B_i}(\mathcal{C}|_{B_i}).$$

Proof. The first claim follows from the relation $\psi^{A_i}(\mathbb{1}) = \varrho_{\text{tot}} - \varrho^{A_i}(\mathbb{1})$, where ϱ_{tot} does not depend on U as a consequence of Eq. (27). For any part A_i that is a singleton, one has $\varrho^{A_i}(\mathbb{1}) = \varrho_{\text{tot}}$, which implies the second identity. The latter, in turn, confirms that $\psi^U(\mathbb{0}) = 0$ for all $U \subseteq S$, since $s_{\mathcal{A}} = r$ in this case.

For the final identity, fix $\mathcal{B} = \{B_1, \dots, B_r\} \in \mathbb{P}(U)$. If $\mathcal{C} = \{C_1, \dots, C_s\} \in \mathbb{P}(U)$ is a refinement of \mathcal{B} , we have $s \geq r$ and \mathcal{C} must be a joining of the form

$$\mathcal{C} = \mathcal{C}_1 \sqcup \dots \sqcup \mathcal{C}_r$$

with $\mathcal{C}_i = \mathcal{C}|_{B_i} \in \mathbb{P}(B_i)$ for all $1 \leq i \leq r$. Writing $\mathcal{C}_i = \{C_1^{(i)}, \dots, C_{k_i}^{(i)}\}$, we have $s = \sum_{i=1}^r k_i$ because $\mathcal{C} \preceq \mathcal{B}$ implies the existence of a bijection between the two families $\{C_k \mid 1 \leq k \leq s\}$ and $\{C_j^{(i)} \mid 1 \leq i \leq r \text{ and } 1 \leq j \leq k_i\}$. Eq. (36) now implies that the right-hand side of our claim is

$$\sum_{i=1}^{|\mathcal{B}|} \psi^{B_i}(\mathcal{C}_i) = \sum_{i=1}^r \sum_{j=1}^{k_i} \psi^{C_j^{(i)}}(\mathbb{1}) = \sum_{\ell=1}^s \psi^{C_\ell}(\mathbb{1}) = \psi^U(\mathcal{C}),$$

and the argument is complete. \square

If $U \neq \emptyset$, with $|U| \leq 3$, one sees with Corollary 2 that

$$\psi^U(\mathcal{A}) = \chi^U(\mathcal{A}) = \sum_{\mathcal{B} \notin [\mathcal{A}, \mathbb{1}]} \varrho^U(\mathcal{B}) \quad (37)$$

holds for any $\mathcal{A} \in \mathbb{P}(U)$. For larger U , this relation holds for $\mathcal{A} = \mathbb{1}$ and any $\mathcal{A} \in \mathbb{P}(U)$ with $|\mathcal{A}| = 2$ if one of the parts is a singleton set, as follows from Theorem 2.

Lemma 5. *Let the decay rates $\psi^S(\mathcal{A})$ with $\mathcal{A} \in \mathbb{P}(S)$ be distinct and assume that the recombination rates for subsystems are calculated according to Eq. (22). Then, for any non-empty $U \subseteq S$, the decay rates $\psi^U(\mathcal{B})$ with $\mathcal{B} \in \mathbb{P}(U)$ are distinct as well. In this case, the exponential functions $e^{-\psi^U(\mathcal{B})t}$ with $\mathcal{B} \in \mathbb{P}(U)$ are linearly independent over \mathbb{R} .*

Proof. Assume distinctness of the rates $\psi^S(\mathcal{A})$ with $\mathcal{A} \in \mathbb{P}(S)$ and fix a non-empty $U \subseteq S$. If $\mathcal{B}, \mathcal{B}' \in \mathbb{P}(U)$ exist with $\mathcal{B} \neq \mathcal{B}'$ but $\psi^U(\mathcal{B}) = \psi^U(\mathcal{B}')$, we may choose $\mathcal{A}, \mathcal{A}' \in \mathbb{P}(S)$ with $\mathcal{A}|_U = \mathcal{B}$, $\mathcal{A}'|_U = \mathcal{B}'$ (hence $\mathcal{A} \neq \mathcal{A}'$) and $\mathcal{A}|_{U^c} = \mathcal{A}'|_{U^c}$. Since $S = U \dot{\cup} U^c$, we can employ Lemma 5 to derive that

$$\psi^S(\mathcal{A}) = \psi^U(\mathcal{A}|_U) + \psi^{U^c}(\mathcal{A}|_{U^c}) = \psi^U(\mathcal{A}'|_U) + \psi^{U^c}(\mathcal{A}'|_{U^c}) = \psi^S(\mathcal{A}'),$$

which is a contradiction to the distinctness assumption. This proves the first claim, while the second assertion is standard. \square

Below, we need to understand the decay rates in various ways. Let us thus expand on their relation with the recombination rates. Starting from the definition in Eq. (36), with $\mathcal{A} = \{A_1, \dots, A_{|\mathcal{A}|}\}$, one obtains

$$\begin{aligned} \psi^U(\mathcal{A}) &= \sum_{i=1}^{|\mathcal{A}|} (\varrho_{\text{tot}} - \varrho^{A_i}(\underline{1})) = |\mathcal{A}| \varrho_{\text{tot}} - \sum_{i=1}^{|\mathcal{A}|} \sum_{\mathcal{B} \in \mathbb{P}(U)} \mathbf{1}_{\{\mathcal{B}|_{A_i} = \underline{1}\}} \varrho^U(\mathcal{B}) \\ &= |\mathcal{A}| \varrho_{\text{tot}} - \sum_{\mathcal{B} \in \mathbb{P}(U)} \left(\sum_{i=1}^{|\mathcal{A}|} \mathbf{1}_{\{\mathcal{B}|_{A_i} = \underline{1}\}} \right) \varrho^U(\mathcal{B}) \\ &= \sum_{\mathcal{B} \in \mathbb{P}(U)} \left(|\mathcal{A}| - \sum_{i=1}^{|\mathcal{A}|} \mathbf{1}_{\{\mathcal{B}|_{A_i} = \underline{1}\}} \right) \varrho^U(\mathcal{B}), \end{aligned} \quad (38)$$

where $\mathbf{1}_{\{\dots\}}$ denotes the characteristic function of the condition in its subscript.

Lemma 6. *The decay rates defined in Eq. (36) are linear functions of the recombination rates of the form $\psi^U(\mathcal{A}) = \sum_{\mathcal{B} \in \mathbb{P}(U)} \kappa^U(\mathcal{A}, \mathcal{B}) \varrho^U(\mathcal{B})$, with*

$$\kappa^U(\mathcal{A}, \mathcal{B}) = |\mathcal{A}| - \sum_{i=1}^{|\mathcal{A}|} \mathbf{1}_{\{\mathcal{B}|_{A_i} = \underline{1}\}}.$$

In particular, the coefficients $\kappa^U(\mathcal{A}, \mathcal{B})$ are non-negative integers, where one has $\kappa^U(\mathcal{A}, \mathcal{B}) = 0$ if and only if $\mathcal{B} \succcurlyeq \mathcal{A}$.

Proof. The first claim summarises our calculation in Eq. (38). It is obvious from the explicit formula that $\kappa^U(\mathcal{A}, \mathcal{B}) \in \mathbb{N}_0$, for all $\mathcal{A}, \mathcal{B} \in \mathbb{P}(U)$. The coefficient vanishes if and only if each condition under the summation is true, which means $\mathcal{B}|_{A_i} = \underline{1} = \{A_i\}$ for all $1 \leq i \leq |\mathcal{A}|$. But the latter condition, in turn, is equivalent with $\mathcal{B} \succcurlyeq \mathcal{A}$, which proves the assertion. \square

Remark 11. One important consequence of Lemma 6 is that, as linear functions of the recombination rates, $\psi^U(\mathcal{A})$ and $\psi^U(\mathcal{B})$ are different whenever $\mathcal{A} \neq \mathcal{B}$, because the intervals $[\mathcal{A}, \underline{1}]$ and $[\mathcal{B}, \underline{1}]$ are then different, too. By a standard Baire category argument, the situation that the rates $\psi^U(\mathcal{A})$ with $\mathcal{A} \in \mathbb{P}(U)$ are distinct is then generic. In fact, the exceptional set with any degeneracy among the decay rates is a union of true hyperplanes in parameter space $\{\varrho^U(\mathcal{A}) \geq 0 \mid \mathcal{A} \in \mathbb{P}(U)\} \simeq \mathbb{R}_+^{B(|U|)}$.

Consequently, it is both a nowhere dense set and a null set. The corresponding statement remains true if we restrict our attention to the reduced parameter space $\{\varrho^U(\mathcal{A}) \geq 0 \mid \mathcal{A} \in \mathbb{P}_2(U)\} \simeq \mathbb{R}_+^{2^{|U|-1}-1}$. \diamond

Let us now return to the original ansatz (34). The coefficient function θ^U is an element of the incidence algebra $\mathbb{A}(\mathbb{P}(U))$. Since $a_0^U(\mathcal{A}) = \delta^U(\mathcal{A}, \underline{1})$ is our initial condition, we see that, for $\emptyset \neq U \subseteq S$ and any $\mathcal{A} \in \mathbb{P}(U)$, we must have

$$\sum_{\mathcal{B} \succ \mathcal{A}} \theta^U(\mathcal{A}, \mathcal{B}) = \delta^U(\mathcal{A}, \underline{1}),$$

hence $\theta^U(\underline{1}, \underline{1}) = 1$ (as mentioned above) together with

$$\theta^U(\mathcal{A}, \underline{1}) = - \sum_{\mathcal{A} \prec \mathcal{C} \prec \underline{1}} \theta^U(\mathcal{A}, \mathcal{C}) \quad (39)$$

for $\mathcal{A} \prec \underline{1}$. This way, for non-empty $U \subseteq S$ and $\mathcal{A} \in \mathbb{P}(U)$, the coefficients $\theta^U(\mathcal{A}, \underline{1})$ are either fixed or recursively determined from $\theta^U(\mathcal{A}, \mathcal{C})$ with $\mathcal{A} \prec \mathcal{C} \prec \underline{1}$.

Taking the derivative of our ansatz (34) leads to

$$\dot{a}_t^U(\mathcal{A}) = - \sum_{\mathcal{B} \succ \mathcal{A}} \psi^U(\mathcal{B}) \theta^U(\mathcal{A}, \mathcal{B}) e^{-\psi^U(\mathcal{B})t}, \quad (40)$$

which has to be compared with the right-hand side of Eq. (30). Inserting the corresponding expression for the coefficients of the subsystem according to our ansatz yields the expression

$$\begin{aligned} \dot{a}_t^U(\mathcal{A}) &= -\varrho_{\text{tot}} a_t^U(\mathcal{A}) + \sum_{\mathcal{B} \succ \mathcal{A}} \varrho^U(\mathcal{B}) \prod_{i=1}^{|\mathcal{B}|} \sum_{\substack{\mathcal{C}_i \in \mathbb{P}(B_i) \\ \mathcal{C}_i \succ \mathcal{A}|_{B_i}}} \theta^{B_i}(\mathcal{A}|_{B_i}, \mathcal{C}_i) e^{-\psi^{B_i}(\mathcal{C}_i)t} \\ &= -\varrho_{\text{tot}} a_t^U(\mathcal{A}) + \sum_{\mathcal{B} \succ \mathcal{A}} \varrho^U(\mathcal{B}) \sum_{\mathcal{C} \in [\mathcal{A}, \mathcal{B}]} \exp\left(-\sum_{i=1}^{|\mathcal{B}|} \psi^{B_i}(\mathcal{C}|_{B_i})t\right) \prod_{j=1}^{|\mathcal{B}|} \theta^{B_j}(\mathcal{A}|_{B_j}, \mathcal{C}|_{B_j}) \\ &= -\varrho_{\text{tot}} a_t^U(\mathcal{A}) + \sum_{\mathcal{B} \succ \mathcal{A}} \varrho^U(\mathcal{B}) \sum_{\mathcal{C} \in [\mathcal{A}, \mathcal{B}]} e^{-\psi^U(\mathcal{C})t} \prod_{j=1}^{|\mathcal{B}|} \theta^{B_j}(\mathcal{A}|_{B_j}, \mathcal{C}|_{B_j}), \end{aligned}$$

where the second part of Lemma 4 was used in the last step. Inserting the expression (34) for $a_t^U(\mathcal{A})$, changing summation variables and regrouping terms gives

$$\begin{aligned} \dot{a}_t^U(\mathcal{A}) &= \sum_{\mathcal{B} \succ \mathcal{A}} e^{-\psi^U(\mathcal{B})t} \left(\varrho_{\text{tot}} \theta^U(\mathcal{A}, \mathcal{B}) - \sum_{\mathcal{C} \succ \mathcal{B}} \varrho^U(\mathcal{C}) \prod_{i=1}^{|\mathcal{C}|} \theta^{C_i}(\mathcal{A}|_{C_i}, \mathcal{B}|_{C_i}) \right) \\ &= - \sum_{\mathcal{B} \succ \mathcal{A}} e^{-\psi^U(\mathcal{B})t} \left(\psi^U(\underline{1}) \theta^U(\mathcal{A}, \mathcal{B}) - \sum_{\mathcal{B} \prec \mathcal{C} \prec \underline{1}} \varrho^U(\mathcal{C}) \prod_{i=1}^{|\mathcal{C}|} \theta^{C_i}(\mathcal{A}|_{C_i}, \mathcal{B}|_{C_i}) \right), \quad (41) \end{aligned}$$

where the second step effectively removes all terms with $\varrho^U(\underline{1})$; compare Remark 2.

Let $U \subseteq S$ be non-empty and let us assume that the exponential functions on the right-hand side of Eq. (34) are linearly independent. By Lemma 5, this is certainly the case when the rates $\psi^S(\mathcal{A})$ with $\mathcal{A} \in \mathbb{P}(S)$ are distinct; compare Remark 11.

Now, a comparison of coefficients in Eq. (41) with those of the right-hand side of Eq. (40) yields the relations

$$\theta^U(\mathcal{A}, \mathcal{B})(\psi^U(\underline{1}) - \psi^U(\mathcal{B})) = \sum_{\mathcal{B} \preccurlyeq \mathcal{C} \prec \underline{1}} \varrho^U(\mathcal{C}) \prod_{i=1}^{|\mathcal{C}|} \theta^{C_i}(\mathcal{A}|_{C_i}, \mathcal{B}|_{C_i}),$$

for all $\mathcal{A}, \mathcal{B} \in \mathbb{P}(U)$ with $\mathcal{A} \preccurlyeq \mathcal{B}$. Note that our assumption entails the condition that $\psi^U(\underline{1}) \neq \psi^U(\mathcal{B})$ for all $\underline{1} \neq \mathcal{B} \in \mathbb{P}(U)$, so that we get

$$\theta^U(\mathcal{A}, \mathcal{B}) = \sum_{\mathcal{B} \preccurlyeq \mathcal{C} \prec \underline{1}} \frac{\varrho^U(\mathcal{C})}{\psi^U(\underline{1}) - \psi^U(\mathcal{B})} \prod_{i=1}^{|\mathcal{C}|} \theta^{C_i}(\mathcal{A}|_{C_i}, \mathcal{B}|_{C_i}) \quad (42)$$

for all $\mathcal{A} \preccurlyeq \mathcal{B} \prec \underline{1}$, where the θ -coefficients for the subsystems are themselves determined recursively, under the distinctness condition for the subsystem decay rates, which follows from our assumption by Lemma 5. Since we know θ^U for any U with $1 \leq |U| \leq 3$ from Corollary 2 and Remark 6, see also Corollary 5 below, the recursion (42) uniquely determines all θ -coefficients.

We have thus shown the following result for the generic situation of Remark 11.

Theorem 4. *Let $S \neq \emptyset$ be a finite set, and assume that the decay rates $\psi^S(\mathcal{A})$ with $\mathcal{A} \in \mathbb{P}(S)$ are distinct. Let the decay rates for non-empty subsets $U \subseteq S$ be defined via Eqs. (35) and (36), where the marginal recombination rates on subsystems are given by Eq. (22).*

Then, our exponential ansatz (34) solves the Cauchy problem of Eq. (14) with initial condition $a_0^S(\mathcal{A}) = \delta^S(\mathcal{A}, \underline{1})$ if and only if the coefficients θ^S are recursively determined by Eq. (42) together with $\theta^U(\underline{1}, \underline{1}) = 1$ and Eq. (39). \square

Let us mention in passing that a similar result can be derived for other initial conditions as well. We concentrate on this one as it fits Theorem 1 and thus also provides a solution of the original recombination equation.

Remark 12. The recursion in Eq. (42), and the solution thus constructed for the recombination dynamics, is similar in structure to the corresponding relations (6.5.1) and (6.5.2) in Lyubich's book [19] for the two-parent case in discrete time. However, in contrast to our θ 's, Lyubich's coefficients depend on ω_0 . As noted in [19], this dependence on the initial condition poses a serious difficulty to the solution, since it makes an explicit iteration impossible. Put differently, our approach achieves the complete separation of the recombination structure from the types, as already inherent in Theorem 4 and further discussed in Section 6; this entails a crucial simplification. \diamond

In the situation of Theorem 4, we thus know that our original ansatz (34) leads to a solution. Since we then also know that $\psi^S(\mathcal{A}) = 0$ holds only for $\mathcal{A} = \emptyset$, and $\psi^S(\mathcal{A}) > 0$ otherwise, we have the following immediate consequence for the asymptotic behaviour of the solution.

Corollary 3. *Under the assumptions of Theorem 4, one has*

$$\lim_{t \rightarrow \infty} a_t^S(\mathcal{A}) = \delta^S(\emptyset, \mathcal{A}),$$

which means that the corresponding solution ω_t of the ODE (7), with initial condition ω_0 , $\|\cdot\|$ -converges to the equilibrium $R_0(\omega_0)$ as $t \rightarrow \infty$. Moreover, the convergence is exponentially fast, with the rate given by $\min\{\psi^S(\mathcal{A}) \mid \mathcal{A} \neq \emptyset\}$, while the individual rate for the convergence of $a_t^S(\mathcal{A})$ is given by $\min\{\psi^S(\mathcal{B}) \mid \emptyset \neq \mathcal{B} \succcurlyeq \mathcal{A}\}$. \square

The recursive nature of our solution does not immediately help to understand its structure and meaning. Let us thus study some aspects of it in more detail.

8. Generic solution: Structure and further properties. To begin, let us observe that Theorem 4 has an interesting consequence which contrasts the result of Lemma 6 and shows a surprising structural similarity with the corresponding relation of Remark 6.

Corollary 4. *Let S be as in Theorem 4 and assume that the decay rates $\psi^S(\mathcal{A})$ with $\mathcal{A} \in \mathbb{P}(S)$ are distinct. Then, the recombination and the decay rates are related by*

$$\varrho^S(\mathcal{A}) = \varrho_{\text{tot}} \delta^S(\mathcal{A}, \underline{1}) - \sum_{\mathcal{B} \succ \mathcal{A}} \theta^S(\mathcal{A}, \mathcal{B}) \psi^S(\mathcal{B}),$$

and the corresponding relation holds for any subsystem that is defined by a non-empty $U \subseteq S$.

Proof. By Theorem 4 and Proposition 4, we know that the ansatz of Eq. (34), for $U = S$, leads to the unique solution of the ODE (14) with initial condition $a_0^S(\mathcal{A}) = \delta^S(\mathcal{A}, \underline{1})$. Thus, we may equate the right-hand side of Eq. (14) with that of Eq. (40), again for $U = S$, and consider the resulting identity at $t = 0$. Observing that

$$\gamma(a_0^S; \mathcal{A}, \mathcal{B}) = \delta^S(\mathcal{A}, \mathcal{B})$$

holds for our probability vector a_0^S as a result of Eq. (12), the first claim is clear.

The second assertion follows by the corresponding calculation with a_t^U , which is justified by Proposition 7 together with Lemma 5. \square

Note that the relation of Corollary 4 is a nonlinear one, because the θ -coefficients themselves generally depend on the recombination rates. To make any further progress, we need to better understand these coefficients, beyond $\theta^U(\underline{1}, \underline{1}) = 1$ and the relations in Eq. (39).

Proposition 8. *Under the assumptions of Theorem 4, we have the following properties.*

1. $\theta^U(\underline{0}, \underline{0}) = 1$ for all non-empty $U \subseteq S$;
2. $\sum_{\mathcal{A} \preceq \mathcal{B}} \theta^U(\mathcal{A}, \mathcal{B}) = \delta^U(\underline{0}, \mathcal{B})$ holds for all $\mathcal{B} \in \mathbb{P}(U)$;
3. Assume further that ϱ^S is strictly positive on all partitions of S with two parts. Then, $\theta^U(\mathcal{A}, \mathcal{A}) \neq 0$ for all $\mathcal{A} \in \mathbb{P}(U)$ and all non-empty $U \subseteq S$; In particular, θ^U is then an invertible element of the incidence algebra for $\mathbb{P}(U)$ over \mathbb{R} .

Proof. For the first claim, observe that $\underline{0} = \underline{1}$ for any U with $|U| = 1$, so $\theta^U(\underline{0}, \underline{0}) = \theta^U(\underline{1}, \underline{1}) = 1$ in this case. Assume now that the claim is true for all U with $|U| \leq r$, and consider a larger set, $U = \{u_1, \dots, u_{r+1}\}$ say. With $\psi^U(\underline{0}) = 0$, we then get from Eq. (42) that

$$\theta^U(\underline{0}, \underline{0}) = \frac{1}{\psi^U(\underline{1})} \sum_{\mathcal{C} \neq \underline{1}} \varrho^U(\mathcal{C}) \prod_{i=1}^{|\mathcal{C}|} \theta^{C_i}(\underline{0}, \underline{0}) = \frac{1}{\psi^U(\underline{1})} \sum_{\mathcal{C} \neq \underline{1}} \varrho^U(\mathcal{C}) = 1,$$

where the second step uses the induction hypothesis (note that we always have $|C_i| \leq r$), while the last step employs Eq. (37) which applies here.

Next, the normalisation property of the $a_t^U(\mathcal{A})$ implies

$$1 = \sum_{\mathcal{A} \in \mathbb{P}(U)} a_t^U(\mathcal{A}) = \sum_{\mathcal{A} \in \mathbb{P}(U)} \sum_{\mathcal{B} \succ \mathcal{A}} \theta^U(\mathcal{A}, \mathcal{B}) e^{-\psi^U(\mathcal{B})t} = \sum_{\mathcal{B} \in \mathbb{P}(U)} e^{-\psi^U(\mathcal{B})t} \sum_{\mathcal{A} \preccurlyeq \mathcal{B}} \theta^U(\mathcal{A}, \mathcal{B}).$$

Since $\psi^U(0) = 0$ and since the decay rates $\psi^U(\mathcal{B})$ with $\mathcal{B} \in \mathbb{P}(U)$ are distinct by our assumption together with Lemma 5, the functions $e^{-\psi^U(\mathcal{B})t}$ are linearly independent. Consequently, the last identity is equivalent to the second claim.

For the third claim, recall that the dimension of the ODE system (14) is $B(n)$ if the set of partitions $\mathcal{A} \in \mathbb{P}(S)$ with $\varrho^S(\mathcal{A}) > 0$ generates the entire lattice $\mathbb{P}(S)$, which is the case under our assumptions. The convex set that is spanned by the solution functions a_t^S is then a simplex of dimension $B(n) - 1$, and the normalisation condition $\sum_{\mathcal{A} \in \mathbb{P}(S)} a_t^S(\mathcal{A}) = 1$ is the only linear relation between these functions.

We already know that $\theta^S(\underline{1}, \underline{1}) = 1$, and one easily finds

$$\theta^S(\mathcal{A}, \mathcal{A}) = \frac{\varrho^S(\mathcal{A})}{\psi^S(\underline{1}) - \psi^S(\mathcal{A})}$$

for any $\mathcal{A} \in \mathbb{P}(S)$ with $|\mathcal{A}| = 2$. One may now proceed inductively in the number of parts. If there were some $\mathcal{A} \neq \underline{0}$ with $\theta^S(\mathcal{A}, \mathcal{A}) = 0$, where we may assume \mathcal{A} to be the coarsest partition with this property, we would get from Eq. (34) an additional linear relations among the solution functions a_t^S , which is impossible. Since $\theta^S(0, 0) = 1$ by the second assertion, the claim is true on the top level (defined by S). Repeating the argument for any non-empty $U \subset S$ completes the argument.

Finally, the invertibility of θ^U as an element of the incidence algebra is a standard consequence of claim (3); compare [1]. \square

Consequently, under the assumptions of part (3) of Proposition 8, θ^U has a unique (left and right) inverse, η^U say. This means that

$$\sum_{\mathcal{B} \in [\mathcal{A}, \mathcal{C}]} \theta^U(\mathcal{A}, \mathcal{B}) \eta^U(\mathcal{B}, \mathcal{C}) = \delta^U(\mathcal{A}, \mathcal{C}) = \sum_{\mathcal{B} \in [\mathcal{A}, \mathcal{C}]} \eta^U(\mathcal{A}, \mathcal{B}) \theta^U(\mathcal{B}, \mathcal{C})$$

holds for all $\mathcal{A} \preccurlyeq \mathcal{C}$. The coefficients η^U are thus determined by the relation $\eta^U(\mathcal{A}, \mathcal{A}) = 1/\theta^U(\mathcal{A}, \mathcal{A})$ together with the recursion

$$\eta^U(\mathcal{A}, \mathcal{C}) = -\frac{1}{\theta^U(\mathcal{A}, \mathcal{A})} \sum_{\mathcal{A} \prec \mathcal{B} \preccurlyeq \mathcal{C}} \theta^U(\mathcal{A}, \mathcal{B}) \eta^U(\mathcal{B}, \mathcal{C})$$

for $\mathcal{A} \prec \mathcal{C}$, which derives from the left equality above. Alternatively, one may use the corresponding formula that derives from the other identity.

Either from direct calculations, or by invoking the results from Section 4, in particular Theorem 2 and Corollary 2, the following result is obvious (see the text after Remark 3 for the definitions of ζ and μ).

Corollary 5. *For any U with $1 \leq |U| \leq 3$, one has $\eta^U = \zeta^U$, and hence also $\theta^U = \mu^U$. For larger sets U , one has $\eta^U(\mathcal{A}, \mathcal{A}) = 1$ for $\mathcal{A} = \underline{1}$ and for all $\mathcal{A} \in \mathbb{P}(U)$ with two parts, if one of them is a singleton set. In the latter case, also $\theta^U(\mathcal{A}, \mathcal{A}) = 1$.*

Proof. Corollary 2 implies the first claim, while Theorem 2 gives the second one for $\mathcal{A} = \underline{1}$ as well as for any \mathcal{A} with two parts, provided one of them is a singleton set. The last claim follows from $\theta^U(\mathcal{A}, \mathcal{A}) \eta^U(\mathcal{A}, \mathcal{A}) = 1$. \square

Let us now define

$$b_t^U(\mathcal{A}) = \sum_{\mathcal{B} \succ \mathcal{A}} \eta^U(\mathcal{A}, \mathcal{B}) a_t^U(\mathcal{B}), \quad (43)$$

which is an analogue of the summatory function from Section 4 (see the proof of Proposition 5). Using $\eta^U * \theta^U = \delta^U$, we now obtain

$$b_t^U(\mathcal{A}) = e^{-\psi^U(\mathcal{A})t}, \quad (44)$$

hence $b_t^U(\emptyset) \equiv 1$ for all $\emptyset \neq U \subseteq S$ as a consequence of $\psi^U(\emptyset) = 0$; compare Lemma 4. One difference to the linear case is that the summation weights in Eq. (43) generally depend on the recombination rates, while they were constant (in fact, given by the ζ -function of the incidence algebra) in Section 4. Due to the properties of the decay rates ψ^U , one inherits the corresponding relations among the coefficient functions b_t^U for $\emptyset \neq U \subseteq S$. In particular,

$$b_t^U(\underline{1}) = e^{-\chi^U(\underline{1})t} = \prod_{\mathcal{A} \in \mathbb{P}(U)} e^{-\varrho^U(\mathcal{A})t}$$

together with

$$b_t^U(\mathcal{A}) = \prod_{i=1}^{|\mathcal{A}|} b_t^{A_i}(\underline{1})$$

determines all coefficients, while Lemma 4 implies the additional relation

$$b_t^U(\mathcal{A}) = \prod_{i=1}^{|\mathcal{B}|} b_t^{B_i}(\mathcal{A}|_{B_i})$$

for any $\mathcal{A}, \mathcal{B} \in \mathbb{P}(U)$ with $\mathcal{A} \preccurlyeq \mathcal{B}$.

Note that the function b_t^U generally does *not* emerge from b_t^S via marginalisation in the sense of Eq. (22), which is another important difference to the special situation of Section 4 and Remark 10.

Lemma 7. *Let θ^S denote the coefficients from Theorem 4 in the case of distinct rates, and assume that $\varrho^S(\mathcal{A}) > 0$ for all $\mathcal{A} \in \mathbb{P}(S)$ with two parts, so that θ^S is an invertible element of the incidence algebra. Then, for any non-empty $U \subseteq S$, one has the following properties.*

1. $\eta^U(\underline{0}, \underline{0}) = 1$;
2. $\eta^U(\underline{0}, \mathcal{A}) = 1$ holds for all $\mathcal{A} \in \mathbb{P}(U)$;
3. $\eta^U(\mathcal{A}, \underline{1}) = 1$ holds for all $\mathcal{A} \in \mathbb{P}(U)$ and all $\emptyset \neq U \subseteq S$.

Proof. The invertibility of θ^U is clear from part (3) of Proposition 8, while $\eta^U(\underline{0}, \underline{0}) = 1$ is a consequence of the first assertion of the same proposition.

Next, we already know that $\eta^U(\underline{0}, \underline{0}) = 1$. For $\mathcal{A} \succ \underline{0}$, we can proceed inductively via the standard inversion formula [1] for $\eta^U = (\theta^U)^{-1}$, which gives

$$\begin{aligned} \eta^U(\underline{0}, \mathcal{A}) &= -\eta^U(\mathcal{A}, \mathcal{A}) \sum_{\mathcal{C} \prec \mathcal{A}} \eta^U(\underline{0}, \mathcal{C}) \theta^U(\mathcal{C}, \mathcal{A}) = -\eta^U(\mathcal{A}, \mathcal{A}) \sum_{\mathcal{C} \prec \mathcal{A}} \theta^U(\mathcal{C}, \mathcal{A}) \\ &= \eta^U(\mathcal{A}, \mathcal{A}) \left(\theta^U(\mathcal{A}, \mathcal{A}) - \sum_{\mathcal{C} \prec \mathcal{A}} \theta^U(\mathcal{C}, \mathcal{A}) \right) = \eta^U(\mathcal{A}, \mathcal{A}) \theta^U(\mathcal{A}, \mathcal{A}) = 1, \end{aligned}$$

where the second step employs the induction hypothesis while the second assertion of Proposition 8 was used in the penultimate step.

The third claim is a consequence of the initial conditions together with Eqs. (43) and (44), because

$$1 = b_0^U(\mathcal{A}) = \sum_{\mathcal{B} \succ \mathcal{A}} \eta^U(\mathcal{A}, \mathcal{B}) a_0^U(\mathcal{B}) = \sum_{\mathcal{B} \succ \mathcal{A}} \eta^U(\mathcal{A}, \mathcal{B}) \delta^U(\mathcal{B}, \underline{1}) = \eta^U(\mathcal{A}, \underline{1})$$

holds for any $\mathcal{A} \in \mathbb{P}(U)$. \square

It is quite clear that the structure of η is somewhat simpler than that of θ . At present, we have η defined as the inverse function to θ in the incidence algebra, but it would be nice to also have a direct way to calculate it, for instance via another recursion.

9. Some comments on the singular cases. Our focus in the previous section was on the generic case that allowed for a recursively defined, general solution. Let us now briefly look into what happens when certain degeneracies among the decay rates occur.

Recall that our ansatz (34) requires, a priori, the linear independence of the exponential functions $e^{-\psi^U(\mathcal{A})t}$, and thus the distinctness of the decay rates $\psi^U(\mathcal{A})$, for $\mathcal{A} \in \mathbb{P}(U)$, separately for all non-empty $U \subseteq S$. A posteriori, we need to understand what happens when two decay rates, as functions of the recombination rates ϱ^S of the system on the top level, become equal. Once again, this is best looked at inductively. If $\emptyset \neq U \subseteq S$ satisfies $|U| \leq 3$, we are in the realm of the ‘linear’ solution of Section 4, and no consequences emerge from degeneracies. In other words, the solution formula from Proposition 5 is valid for *all* values of the recombination rates, irrespective of possible degeneracies; compare Corollaries 5 and 2.

When we step up in system size, there can be degeneracies that are still ‘harmless’ in the sense that the θ -coefficients extend continuously to these situations. In this case, the (now reduced) set of exponentials in Eq. (34) is still sufficient, and the solution valid. However, as is evident from Eq. (42), the situation changes when $\psi^U(\mathcal{B}) = \psi^U(\underline{1})$ for some $\mathcal{B} \in \mathbb{P}(U)$, as we then hit a singularity. To understand the underlying phenomenon, let us assume that U corresponds to the smallest subsystem where this type of ‘bad’ degeneracy occurs. Now, rewrite Eq. (41) as

$$\dot{a}_t^U(\mathcal{A}) = -\psi^U(\underline{1}) a_t^U(\mathcal{A}) + \sum_{\mathcal{A} \prec \mathcal{B} \prec \underline{1}} \varepsilon(\mathcal{A}, \mathcal{B}) e^{-\psi^U(\mathcal{B})t}$$

with $\varepsilon(\mathcal{A}, \mathcal{B}) = \sum_{\mathcal{C} \prec \mathcal{A} \prec \underline{1}} \varrho^U(\mathcal{C}) \prod_{i=1}^{|\mathcal{C}|} \theta^{C_i}(\mathcal{A}|_{C_i}, \mathcal{B}|_{C_i})$. Note that the $\varepsilon(\mathcal{A}, \mathcal{B})$ are well-defined for all $\mathcal{A}, \mathcal{B} \in \mathbb{P}(U)$ and *all* values of the recombination rates, because only θ -coefficients of subsystems smaller than U occur in the product. We are now in the standard situation of ODE theory that we have summarised in Lemma 8 of the Appendix: Precisely when $\psi^U(\mathcal{B}) = \psi^U(\underline{1})$ for some $\mathcal{B} \in \mathbb{P}(U)$, we need an additional function for our ansatz, namely $t \cdot e^{-\psi^U(\mathcal{B})t}$. So, our original ansatz (34) no longer suffices, and has to be modified accordingly. At the same time, in line with our previous observation, additional degeneracies of the form $\psi^U(\mathcal{B}) = \psi^U(\mathcal{B}')$ with $\mathcal{B}, \mathcal{B}' \neq \underline{1}$ are harmless.

Now, this dichotomic structure continues on each new level: There are ‘harmless’ degeneracies that do not require additional functions for our ansatz, while degeneracies of the type $\psi^U(\mathcal{B}) = \psi^U(\underline{1})$ once again render the function set of the ansatz incomplete, even one that was augmented in the previous step. Lemma 9 of the Appendix reviews a typical case, while the remarks following it show how the

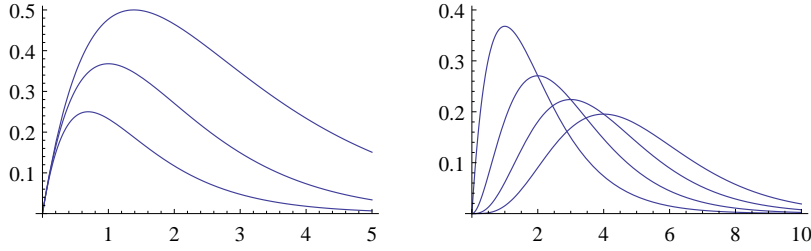


FIGURE 3. Left panel: Illustration of the function $E_0(\alpha, \beta; t)$ of Lemma 8 for $\alpha = 1$ and three values of β , namely $\beta = 0.5$ (top curve), $\beta = 1$ (middle) and $\beta = 2$ (bottom). Right panel: Illustration of $E_m(1, 1; t)$ for parameters $m \in \{0, 1, 2, 3\}$ (from left to right).

degenerate case produces extra monomial factors of increasing exponents for each degeneracy of the ‘bad’ type. At this point, we hope that the general structure is sufficiently clear, so that we can leave further details to the reader.

Let us summarise our discussion as follows.

Corollary 6. *If S is a finite set as before, the generic recursive solution from Theorem 4 extends to all recombination rates $\varrho^S(\mathcal{A})$, with $\mathcal{A} \in \mathbb{P}(S)$, such that $\psi^U(\mathbf{1}) \neq \psi^U(\mathcal{B})$ holds for all non-empty $U \subseteq S$ and all $\mathcal{B} \in \mathbb{P}(U)$. \square*

Appendix. Here, we state some useful results from classical ODE theory, whose proofs are straightforward exercises that are left to the reader.

Lemma 8. *Let ϱ and $\sigma_1, \dots, \sigma_m$ be non-negative real numbers. Then, the Cauchy problem defined by the ODE*

$$\dot{g} = -\varrho g + \sum_{i=1}^m \varepsilon_i e^{-\sigma_i t}$$

together with the initial condition $g(0) = g_0$ has the unique solution

$$g(t) = g_0 e^{-\varrho t} + \sum_{i=1}^m \varepsilon_i E_0(\varrho, \sigma_i; t),$$

where

$$\begin{aligned} E_0(\alpha, \beta; t) &:= \sum_{\ell=0}^{\infty} (-1)^\ell (\alpha^\ell + \alpha^{\ell-1}\beta + \dots + \alpha\beta^{\ell-1} + \beta^\ell) \frac{t^{\ell+1}}{(\ell+1)!} \\ &= \begin{cases} t e^{-\alpha t}, & \text{if } \alpha = \beta, \\ \frac{1}{\alpha - \beta} (e^{-\beta t} - e^{-\alpha t}), & \text{otherwise,} \end{cases} \end{aligned}$$

is a smooth function that is non-negative for all $t \geq 0$ and symmetric under the exchange of the two parameters α and β . \square

An illustration of the function E_0 is shown in the left panel of Figure 3. A similar type of result emerges for mixtures of exponentials with higher order monomials.

Lemma 9. *Let ϱ and σ be non-negative real numbers, and $m \in \mathbb{N}_0$. Then, the Cauchy problem*

$$\dot{g} = -\varrho g + \varepsilon \frac{t^m}{m!} e^{-\sigma t}$$

with initial condition $g(0) = g_0$ has the unique solution $g(t) = g_0 e^{-\varrho t} + \varepsilon E_m(\varrho, \sigma; t)$, where

$$\begin{aligned} E_m(\varrho, \sigma; t) &:= e^{-\varrho t} \int_0^t \frac{\tau^m}{m!} e^{(\varrho-\sigma)\tau} d\tau \\ &= \begin{cases} e^{-\varrho t} \frac{(-1)^{m+1}}{(\varrho-\sigma)^{m+1}} + e^{-\sigma t} \sum_{\ell=0}^m \frac{(-1)^\ell t^{m-\ell}}{(m-\ell)! (\varrho-\sigma)^{\ell+1}}, & \text{if } \sigma \neq \varrho, \\ \frac{t^{m+1}}{(m+1)!} e^{-\varrho t}, & \text{if } \sigma = \varrho, \end{cases} \end{aligned}$$

is a smooth non-negative function for all $t \geq 0$. \square

Some examples are illustrated in the right panel of Figure 3. By standard results from ODE theory, it is clear how to combine the results of Lemmas 8 and 9 to cover further situations. Let us only add that the ODE

$$\dot{g} = -\varrho g + \sum_{\ell=0}^m \varepsilon_\ell \frac{t^\ell}{\ell!} e^{-\varrho t}$$

with initial condition $g(0) = g_0$ has the unique solution

$$g(t) = g_0 e^{-\varrho t} + \sum_{\ell=0}^m \varepsilon_\ell \frac{t^{\ell+1}}{(\ell+1)!} e^{-\varrho t},$$

which explains the appearance of monomial factors with increasing powers in the solution of such equations with degenerate rates.

Acknowledgments. MB would like to thank Roland Speicher for valuable discussions. We thank an anonymous reviewer for his insightful and constructive comments. This work was supported by the German Research Foundation (DFG), within the SPP 1590.

REFERENCES

- [1] M. Aigner, *Combinatorial Theory*, reprint, Springer, Berlin, 1997.
- [2] H. Amann, *Gewöhnliche Differentialgleichungen*, 2nd ed., de Gruyter, Berlin, 1995.
- [3] E. Baake, Deterministic and stochastic aspects of single-crossover recombination, in: *Proceedings of the International Congress of Mathematicians, Hyderabad, India, 2010*, Vol. VI, ed. J. Bhatia, Hindustan Book Agency, New Delhi (2010), 3037–3053.
- [4] E. Baake and I. Herms, [Single-crossover dynamics: Finite versus infinite populations](#), *Bull. Math. Biol.*, **70** (2008), 603–624; [arXiv:q-bio/0612024](#).
- [5] M. Baake, [Recombination semigroups on measure spaces](#), *Monatsh. Math.*, **146** (2005), 267–278 and **150** (2007), 83–84 (Addendum); [arXiv:math.CA/0506099](#).
- [6] M. Baake and E. Baake, [An exactly solved model for mutation, recombination and selection](#), *Can. J. Math.*, **55** (2003), 3–41 and **60** (2008), 264–265 (Erratum); [arXiv:math.CA/0210422](#).
- [7] E. Baake and T. Hustedt, Moment closure in a Moran model with recombination, *Markov Proc. Rel. Fields*, **17** (2011), 429–446; [arXiv:1105.0793](#).
- [8] E. Baake and U. von Wangenheim, [Single-crossover recombination and ancestral recombination trees](#), *J. Math. Biol.*, **68** (2014), 1371–1402; [arXiv:1206.0950](#).
- [9] M. Baake and R. Speicher, in preparation.
- [10] J. H. Bennett, On the theory of random mating, *Ann. Human Gen.*, **18** (1954), 311–317.
- [11] R. Bürger, *The Mathematical Theory of Selection, Recombination and Mutation*, Wiley, Chichester, 2000.

- [12] K. J. Dawson, [The decay of linkage disequilibrium under random union of gametes: How to calculate Bennett's principal components](#), *Theor. Popul. Biol.*, **58** (2000), 1–20.
- [13] K. J. Dawson, [The evolution of a population under recombination: How to linearise the dynamics](#), *Lin. Alg. Appl.*, **348** (2002), 115–137.
- [14] R. Durrett, *Probability Models for DNA Sequence Evolution*, 2nd ed., Springer, New York, 2008.
- [15] M. Esser, S. Probst and E. Baake, Partitioning, duality, and linkage disequilibria in the Moran model with recombination, submitted, preprint [arXiv:1502.05194](#).
- [16] W. J. Ewens and G. Thomson, Properties of equilibria in multi-locus genetic systems, *Genetics*, **87** (1977), 807–819.
- [17] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. I, 3rd ed., Wiley, New York, 1986.
- [18] H. Geiringer, [On the probability theory of linkage in Mendelian heredity](#), *Ann. Math. Stat.*, **15** (1944), 25–57.
- [19] Y. Lyubich, *Mathematical Structures in Population Genetics*, Springer, Berlin, 1992.
- [20] T. Nagylaki, J. Hofbauer and P. Brunovski, [Convergence of multilocus systems under weak epistasis or weak selection](#), *J. Math. Biol.*, **38** (1999), 103–133.
- [21] J. R. Norris, *Markov Chains*, Cambridge University Press, Cambridge, 1998, reprint, 2005.
- [22] O. Redner and M. Baake, [Unequal crossover dynamics in discrete and continuous time](#), *J. Math. Biol.*, **49** (2004), 201–226; [arXiv:math.DS/0402351](#).
- [23] N. J. A. Sloane, [The On-line encyclopedia of integer sequences](#), *Lecture Notes in Computer Science*, **4573** (2007), p 130, <https://oeis.org/>
- [24] E. D. Sontag, [Structure and stability of certain chemical networks and applications to the kinetic proofreading model of T-cell receptor signal transduction](#), *IEEE Trans. Automatic Control*, **46** (2001), 1028–1047; [arXiv:math.DS/0002113](#).
- [25] E. Spiegel and C. J. O'Donnell, *Incidence Algebras*, Marcel Dekker, New York, 1997.
- [26] U. von Wangenheim, E. Baake and M. Baake, [Single-crossover recombination in discrete time](#), *J. Math. Biol.*, **60** (2010), 727–760; [arXiv:0906.1678](#).

Received August 2014; revised March 2015.

E-mail address: mbaake@math.uni-bielefeld.de

E-mail address: ebaake@techfak.uni-bielefeld.de

E-mail address: msalamat@techfak.uni-bielefeld.de