An exactly solved model for mutation, recombination and selection

Michael Baake and Ellen Baake

Institut für Mathematik und Informatik,
Universität Greifswald, Jahnstr. 15a,
17487 Greifswald, Germany

Abstract

It is well known that rather general mutation-recombination models can be solved algorithmically (though not in closed form) by means of Haldane linearization. The price to be paid is that one has to work with a multiple tensor product of the state space one started from.

Here, we present a relevant subclass of such models, in continuous time, with independent mutation events at the sites, and crossover events between them. It admits a closed solution of the corresponding differential equation on the basis of the original state space, and also closed expressions for the linkage disequilibria, derived by means of Möbius inversion. As an extra benefit, the approach can be extended to a model with selection of additive type across sites. We also derive a necessary and sufficient criterion for the mean fitness to be a Lyapunov function and determine the asymptotic behaviour of the solutions.

Key Words: population genetics, recombination, nonlinear ODEs, measure-valued dynamical systems, Möbius inversion

MSC 2000: 92D10, 34L30 (primary); 37N30, 06A07, 60J25 (secondary)

Introduction

The basic mechanisms which create genetic variation in biological evolution are *mutation* and *recombination*. They are counteracted by *selection*, which removes variation. Genetic information may be quite generally described in terms of a collection of linearly ordered sites (i.e. a sequence of sites), each of which is occupied by an element of a given (finite or infinite) set which we denote as *site space*; if this set is finite, it is often termed *alphabet*. A specific sequence is also called *type*.

Mutation is treated as a random state change of a site variable, which occurs independently at every site. *Recombination* occurs on the occasion of sexual reproduction, and refers to the creation of 'offspring' sequences from two (randomly chosen) 'parental' ones, where a subset of the 'maternal' sites is combined with the complementary set of the 'paternal' sites, and the linear ordering along the sequence is maintained. This process is realized through one, or a number of, *crossover events*, where the two parental strands are interlaced between a pair of neighbouring sites. An important feature of recombination is that it removes dependencies between sites, known as *linkage disequilibria* in genetics. Finally, *selection* is caused by the flourishing of fit individuals at the expense of less fit ones.

We consider an infinite population of sequences which evolves under the joint action of mutation, selection or recombination, or of any combination thereof. This is to be considered as the infinite population limit (IPL) of the stochastic process alluded to, and defines a deterministic dynamical system for probability measures (in discrete or in continuous time). It describes the time evolution of the measure with probabilistic certainty, see [21, Ch. 11], and Thm. 2.1 of it in particular. Although there are many interesting and important questions connected with finite populations, we focus on the differential equation of the deterministic limit here, which we will call *IPL equation* from now on. In particular, we will not employ the traditional discrete dynamical systems, but follow the continuous route along the lines of Kimura [32] and Akin [2], which happens to be much less developed than it ought to be, see also [7] for a recent review.

Mutation is a linear process and straightforward to deal with. Selection involves some nonlinearity, which is due to norm conservation under the dynamics, but this nonlinearity may be removed through a simple transformation. Recombination contains a very different source of nonlinearity, which is due to the fact that pairs of objects are involved in the process, and is much harder to treat. Nevertheless, if both state space and time are discrete, a procedure (known as *Haldane linearization*, see [38, 16, 17] and [36, Ch. 6]) is available which transforms the dynamical system (exactly) into a linear one. It involves a multilinear transformation of the probabilities to a new set of variables, namely certain linkage disequilibria, which describe the deviation from statistical independence of sites. These variables decay independently and geometrically, i.e. they decouple and diagonalize the dynamics. Unfortunately, however, the procedure is cumbersome since it relies on recursions, and no closed form is available for the transformation in the general case.

In a previous paper [6], the special case of single crossover events was considered, where offspring sequences are composed from one maternal and one paternal segment. This scenario is particularly relevant in molecular evolution, where crossover events are rare, and it is most consistently described in continuous time. For discrete site spaces, and with the help of the corresponding vector space structure, the linearizing transform could be given in closed form with the help of elementary methods from multilinear algebra.

The aim of this article is to further develop this approach in a systematic measure-theoretic setting which also incorporates more general site spaces and does not require an explicit coordinatization. We will essentially start from the deterministic IPL equation and construct its solution explicitly, first for recombination only. The so-called Möbius inversion principle will then give a simple approach to the calculation of a suitable (and, in particular, complete) set of linkage disequilibria. It will then turn out that mutation and even selection may be included in the framework, provided fitness is additive, meaning that the fitness of any type may be decomposed into a sum of independent contributions of its individual sites, i.e. if there is no interaction between sites. Such results may be helpful for the solution of the corresponding inverse problem, i.e. the determination of recombination rates from experimental data, e.g. observed patterns of linkage disequilibria along sequences [15, 43].

The exposition will be more explicit than needed for a purely mathematical audience, and we also try to give rather precise references to background material we use. We hope that the article will become more self-contained this way and that it is also accessible for readers with a more biological background.

The structure of the paper is as follows. After some preliminaries in Section 1, we will briefly summarize the description of mutation through an IPL equation on the space of positive measures in Section 2, followed by some general remarks on measure-valued IPL equations. The core of the article is Section 3, where we solve, step by step, the IPL equation for recombination and construct an explicit solution of the abstract Cauchy problem, together with a closed form of the corresponding linkage disequilibria. The latter is based on an application of the inclusion-exclusion principle via Möbius inversion (a supplement is given in the Appendix). Section 4 combines mutation and recombination. Section 5 deals with selection and recombination, with some emphasis on the role of mean fitness as a Lyapunov function. Finally, Section 6 ties together all three evolutionary forces — still giving an explicit solution, expressions for the linkage disequilibria, and asymptotic properties. We close with some afterthoughts mainly aimed at the relationship to models in discrete time.

1 Preliminaries

If X is a locally compact space (by which we always mean to include the Hausdorff property), we use $\mathcal{M}_+(X)$ to denote the set of *finite* positive regular Borel measures on X, with $0 \in \mathcal{M}_+(X)$. Likewise, $\mathcal{M}(X)$ is the vector space of real (or signed) finite regular Borel measures. It is a Banach space under the norm $\|\omega\| = |\omega|(X)$ where $|\omega|$ denotes total variation measure. Due to the Riesz-Markov representation theorem, $\mathcal{M}(X)$ can also be viewed as the dual of $C_\infty(X, \mathbb{R})$, the Banach space of real-valued continuous functions which vanish at infinity, equipped with the usual supremum norm, see [39, Thm. IV.18], as well as [41, Ch. 6] and [39, Ch. IV.4] for general background material. Note that $\mathcal{M}(X)$ with the variation norm $\|.\|$ is actually a Banach lattice, and this gives access to the highly developed theory of positive operators [42, 5]. We will mainly be interested in the closed convex subsets $\mathcal{M}_+^m(X) := \{\omega \in \mathcal{M}_+(X) \mid \omega(X) = m\}$, and in $\mathcal{P}(X) = \mathcal{M}_+^1(X)$ in particular, the set of *probability measures* on X. Note that, for positive measures ω, we simply have $\|\omega\| = \omega(X)$.

If the Borel σ-algebra of X is generated by a family of sets that is closed under finite intersections, a regular Borel measure on X is already uniquely specified by its values on the elements of this generating family [9, 35]. This is a property that we will need several times, in particular if $X = X_1 \times X_2$ is a product space, equipped with the product topology.

Fact 1 *Let ν, ν' be two regular Borel measures on the locally compact product space $X = X_1 \times X_2$ which coincide on all "rectangles" $E_1 \times E_2$ where E_1 and E_2 each run through the Borel sets of X_1 and X_2. Then $\nu = \nu'$, i.e. $\nu(E) = \nu'(E)$ for all Borel sets E of X.*

PROOF: In view of the above remark, the only obstacle to cope with is the (non-vacuous!) situation when the σ-algebra generated by the rectangles $E_1 \times E_2$ is *not* the full Borel σ-algebra of X. However, the σ-algebra generated by the rectangles contains all Baire sets F of X, because the Baire sets of X possess the required Cartesian product property [9, Lemma 56.2], and the Borel sets of X_i contain the Baire sets of X_i. The equality of ν and ν' now follows from [9, Thm. 62.1] (this rests on the fact that every Baire measure has a unique extension to a regular Borel measure). $\square$

Standard examples of locally compact spaces include the compact ones, such as any finite set or the closed interval $[0, 1]$, but also $\mathbb{R}^k$ and $\mathbb{Z}^\ell$ with $k, \ell \geq 0$, and arbitrary combinations thereof. These certainly cover all meaningful parameter spaces to be expected in biological applications.

If X is a finite set (which is an important case in population genetics), $\mathcal{P}(X)$ is a *simplex*. If the cardinality of X is M, this simplex has dimension $M - 1$, i.e. any probability measure is a unique convex linear combination of the M *extremal* measures that constitute the vertices of the simplex. If $X = \{1, \ldots, M\}$, they are denoted by e_i, $i = 1, \ldots, M$, and fixed by their values on singleton sets, $e_i(\{j\}) = \delta_{i,j}$. In other words, any $\omega \in \mathcal{P}(X)$ is of the form $\omega = \sum_{i=1}^{M} a_i e_i$ with all $a_i \geq 0$ and $a_1 + \ldots + a_M = 1$. This provides the canonical coordinatization of this situation.

The set (or state space) X that we need will have a product structure, described on the basis of *sites*. For later convenience, we use $N = \{0, 1, \ldots, n\}$ for the set of sites, i.e. we start counting with 0 here. To site i, we attach the locally compact space X_i, and our state space is then

$$X \; = \; X_0 \times X_1 \times \ldots \times X_n \, , \tag{1}$$

which is still locally compact. One Banach space of measures to show up is the space $\mathcal{M}(X)$ with the corresponding variation norm $\|.\|$. Note that

$\mathcal{M}(X)$ contains the (algebraic) tensor product space $\mathcal{M}^{\otimes} := \bigotimes_{i=0}^{n} \mathcal{M}(X_i)$, and also its completion (here, the closure in the given $\|.\|$-norm of $\mathcal{M}(X)$). To simplify notation, the latter will also be denoted by $\mathcal{M}^{\otimes}$, because we shall only deal with Banach spaces here. Recall that $\mathcal{M}^{\otimes}$ contains the product measures $\omega = \omega_0 \otimes \cdots \otimes \omega_n$ with $\omega_i \in \mathcal{M}(X_i)$, but also all (finite) linear combinations of measures of this kind. Because we consider the completion, also all measures are contained which can be approximated with such linear combinations in the norm. All probability measures of product form are in this space, but note that the single measures in the product need not be probability measures themselves.

If $X_i = \{1, \ldots, M_i\}$ is finite, X is still a finite set, with $M = \prod_{i=0}^{n} M_i$ elements. Then, $\mathcal{M}(X) = \mathcal{M}^{\otimes}$, and this is simply a real vector space of dimension M. $\mathcal{M}(X) = \mathcal{M}^{\otimes}$ is also true for X discrete. In this case, the action of operators in tensor product form is well defined. In general, if $\mathcal{M}^{\otimes} \subsetneq \mathcal{M}(X)$, one can still go beyond $\mathcal{M}^{\otimes}$ under certain circumstances, e.g. by including integrals (rather than finite sums) of product measures. However, we do not want to enter this rather technical discussion, and refer to [35, Ch. IX.6] and [42, Ch. IV.7] for some background material, and to [18, Ch. 13] for some of the problems that are related to these difficulties.

X finite is the case most frequently studied in the theory of sequence evolution, and it was the motivation for this work, see [6] and references given there. However, many results hold in greater generality, which we want to cover in view of potential applications in quantitative genetics. There, the space X_i often is a state space such as $\mathbb{R}$, or a compact subset thereof. In this case, $\mathcal{M}^{\otimes}$ is a true subspace of $\mathcal{M}(X)$, which has to be taken care of later on (occasional restrictions of X to a finite set will be mentioned explicitly).

The main reason for using the above set N of sites is that we will need *ordered partitions* of N, which are uniquely specified by a set of cuts or *crossovers*. The possible cut positions are at the *links* between sites, which we denote by half-integers, i.e. by elements of the set $L = \{\frac{1}{2}, \frac{3}{2}, \ldots, \frac{2n-1}{2}\}$. We will use Latin indices for sites and Greek indices for links, and the implicit rule will always be that $\alpha = \frac{2i+1}{2}$ is the link between site i and $i+1$.

With this notation, the ordered partitions of N are in one-to-one correspondence with the subsets of L as follows. If $A = \{\alpha_1, \ldots, \alpha_p\} \subset L$, let N_A denote the ordered partition

$$\{0, \ldots, \lfloor \alpha_1 \rfloor\}, \{\lceil \alpha_1 \rceil, \ldots, \lfloor \alpha_2 \rfloor\}, \ \ldots \ , \{\lceil \alpha_p \rceil, \ldots, n\}$$

where $\lfloor\alpha\rfloor$ ($\lceil\alpha\rceil$) denotes the largest integer below α (the smallest above α). In particular, we have $N_{\varnothing} = N$ and $N_L = \{\{0\},\dots,\{n\}\}$. With this definition, it is clear that N_B is a *refinement* of N_A if and only if $A \subset B$. Consequently, the lattice of ordered partitions of N corresponds to the Boolean algebra of the finite set L, denoted by $\mathcal{B}(L)$, cf. [1, Ch. I.2]. We prefer this notation to that with partitions, as it is easier to deal with. If $A \subset B$, we will write $B-A$ for $B \setminus A$, and $\bar{A}$ for the set $L-A$.

This setup allows us to use the powerful tool of *Möbius inversion* from combinatorial theory [1, Ch. IV.2], which is a systematic way to employ the inclusion-exclusion principle. If f and g are mappings from $\mathcal{B}(L)$ to $\mathbb{R}$ which are, for all $A \subset L$, related by

$$g(A) \;=\; \sum_{B\subset A} f(B)\,, \qquad (2)$$

then this can be solved for f via the inversion formula [1, Thm. 4.18]

$$f(A) \;=\; \sum_{B\subset A} g(B)\,\mu(B,A) \qquad (3)$$

with the Möbius function $\mu(B,A) = (-1)^{|A-B|}$, where $|A-B|$ stands for the cardinality of the set $A-B$. For B not a subset of A, we set $\mu(B,A) = 0$ which makes the Möbius function into an element of the so-called incidence algebra, see [1, Ch. IV.1] for details. It is important to note that Möbius inversion is not restricted to functions, it also applies to bounded operators.

2 Mutation and Markov generator

The description of mutation is rather straight-forward. Let us start from a *finite* population. Since we are working in continuous time, we assume an independent Poisson clock for each individual member of a (finite) population, and a mutation occurs for an individual whenever its clock rings, according to prescribed mutation rates between (finitely many) types or states. Since the individuals are independent, this is a simple Markov process for each of them. If we now go to the infinite population limit, the time evolution of the probability measure for the types is, almost surely, described by a (deterministic) ordinary differential equation (ODE). This is the so-called *IPL* *equation*, compare [21, Thm. 11.2.1] for a general justification, which we will also rely on below. For the simple mutation case, this ODE is linear. It clearly coincides with the ODE for the probability measure of the individual Markov process, usually obtained from multiple realizations through the law of large numbers.

Let us consider the case that X is a finite state space of cardinality $|X| = M$ in more detail, where $\dim_{\mathbb{R}}(\mathcal{M}(X)) = M$. The mutation rate from state ℓ to state k is given by $Q_{k,\ell} = Q_{k\leftarrow\ell}$, where we already consider Q as a mapping acting on the corresponding probabilities, resp. measures. The rate matrix Q is a Markov generator, i.e. it has non-negative entries everywhere except on its diagonal, and vanishing column[1] sums. The time evolution is then fully described by the Markov semigroup $\{\exp(tQ) \mid t \geq 0\}$, see [21, Ch. 1.1 and Ch. 4.2]. We shall usually assume that Q is *irreducible*, i.e. it is possible to reach every state from any other one. In this case, the equilibrium state is unique and given by the properly normalized 0-eigenvector of the generator Q. It can actually be given in closed form, see [24, Lemma 6.3.1].

If X has the product structure introduced above, our mutation process is supposed to be of a more special form, for biological reasons. We assume that mutation happens at all sites in parallel and independently from one another, so that our generator has the form

$$Q \; = \; \sum_{i=0}^{n} Q_i \tag{4}$$

where each Q_i is, in a properly coordinatized way, the tensor product of a rate matrix at site i and unit matrices of matching dimension everywhere else, i.e.

$$Q_i \; = \; \mathbf{1}_{M_0} \otimes \cdots \otimes \mathbf{1}_{M_{i-1}} \otimes q_i \otimes \mathbf{1}_{M_{i+1}} \otimes \cdots \otimes \mathbf{1}_{M_n} \tag{5}$$

where q_i is a local rate matrix (of dimension M_i) for the state space X_i, acting on $\mathcal{M}(X_i)$. The rate matrices Q_i clearly commute with one another. Note also that Q of (4) is irreducible if and only if all the q_i are. The Markov semigroup inherits the tensor product structure, i.e. we have

$$\exp(tQ) \; = \; \prod_{i=0}^{n} \exp(tQ_i) \; = \; \bigotimes_{i=0}^{n} \exp(tq_i)\,. \tag{6}$$

[1]In contrast to the standard probability literature, we adopt the transposed version here since we are considering the situation from the (linear) operator point of view.

In view of our following description of recombination, we prefer to avoid an explicit coordinatization here, so we will not use matrix notation. This simply means that we have to reinterpret the generator Q as a linear operator on $\mathcal{M}(X)$. Nothing of the above actually changes, we only have to read Q (or q_i) as a linear mapping on $\mathcal{M}(X)$ (or on $\mathcal{M}(X_i)$). The two conditions for Q to be a Markov generator now read as follows (the analogous conditions apply to q_i in relation to $\mathcal{M}(X_i)$).

1. If ν is a positive measure and E any Borel set such that $\nu(E) = 0$, then $(Q\nu)(E) \geq 0$.

2. If ν is a positive measure, then $(Q\nu)(X) = 0$.

The first condition ensures that the semigroup generated by Q maps $\mathcal{M}_+(X)$ into itself. Under the present circumstances, where Q is bounded and $\mathcal{M}(X)$ is a reflexive Banach space, this condition is necessary and sufficient for the positivity of $\exp(tQ)$, see [5, Thm. 1.11]. It is sometimes also called the *positive minimum principle*. The second condition means that the semigroup is Markov, i.e. it preserves the norm of positive measures, and, in particular, maps $\mathcal{P}(X)$ into itself. In this setting, irreducibility implies that the kernel of the Markov generator Q is one-dimensional.

The IPL equation for our simple mutation process[2] now reads

$$\dot{\omega} \; = \; \Phi_{\mathrm{mut}}(\omega) \; := \; \Big(\sum_{i=0}^{n} Q_i\Big)\, \omega \tag{7}$$

which we will take, in generalization of the discrete situation, as the starting point for the analysis of mutation, without tracing it back to an explicit stochastic process. We then obtain, by employing standard results [4] from the theory of ordinary linear differential equations in (finite-dimensional) Banach spaces (see also Theorem 1 below):

Proposition 1 *The abstract Cauchy problem of the IPL equation* (7) *with initial condition* $\omega_0 \in \mathcal{P}(X)$ *has the unique solution*

$$\omega_t \; = \; \exp\Big(t \sum_{i=0}^{n} Q_i\Big)\, \omega_0$$

which is, for $t \geq 0$, *a one-parameter family of probability measures.* $\square$

[2] In this linear case, the IPL equation is closely related to the master equation commonly used in the physics literature, see [29, Ch. 5] for details.

To formulate a generalization of Prop. 1, let us forget about the product structure for a moment and consider the linear ODE

$$\dot{\omega} \;=\; Q\omega$$

with Q the generator of a uniformly (or norm) continuous Markov semigroup on $\mathcal{M}(X)$, compare [19, Ch. I.3]. This is the case if and only if the linear operator Q, in addition to satisfying assumptions 1. and 2. from above, is bounded, and hence defined on all of $\mathcal{M}(X)$, see [19, Cor. II.1.5]. In particular, we can then write the semigroup in exponential form [19, Thm. I.3.7], i.e. as $\exp(tQ)$, and the solution as $\omega_t = \exp(tQ)\,\omega_0$. In what follows, we will (non-constructively) assume that a process is given that leads to a bounded generator Q which is a linear operator on $\mathcal{M}(X)$, i.e. maps regular Borel measures to regular Borel measures. As long as this is the case, it is sufficient to work with assumptions 1. and 2., even if the space of measures considered is no longer reflexive. The analogue of Prop. 1 then holds on the Banach subspace $\mathcal{M}^{\otimes}$, to which we shall restrict our attention whenever Q is of the form specified in Eqs. (4) and (5). This makes no difference at all as long as X is discrete.

Many results can still be generalized to densely defined generators of strongly continuous semigroups, see [19, Ch. I.5], but already the well-posedness of the Cauchy problem needs some thought, compare [19, Ch. II.6] for a discussion. Also, the characterization of generators for positive semigroups becomes more involved, see [5, Ch. 3]. Usually, one would then rather describe the entire process by means of semigroups on function spaces, compare [21, Ch. 1.4]. Since all explicit mutation schemes we have in mind lead to uniformly continuous semigroups, we will not expand on the more general situation.

Let us instead add a few remarks on the general type of IPL equation that arises when recombination and selection are also included. This will also better explain our formulation of mutation, from the point of view of measure-valued differential equations. In what follows, it is sufficient to investigate the first order ODE

$$\dot{\omega} \;=\; \Phi(\omega) \tag{8}$$

on the Banach space $\mathcal{M}(X)$, where Φ is a mapping from $\mathcal{M}(X)$ into itself (alternatively, we can study (8) on any closed subspace of $\mathcal{M}(X)$ that is invariant under Φ). Unlike Φ_{mut} from (7), Φ need not be linear, and it is the nonlinear cases below that we are most interested in. The three properties we will meet below are:

A1 The mapping Φ is (globally) Lipschitz.

A2 If $\nu \in \mathcal{M}_+(X)$, i.e. ν is a positive measure, and E any Borel set such that $\nu(E) = 0$, then we have $\big(\Phi(\nu)\big)(E) \geq 0$.

A3 For any $\nu \in \mathcal{M}_+(X)$, we have $\big(\Phi(\nu)\big)(X) = 0$.

It is clear that our formulation of mutation constitutes a linear example of such a mapping.

Theorem 1 *If* $\Phi : \mathcal{M}(X) \to \mathcal{M}(X)$ *satisfies* (A1)*, the abstract Cauchy problem of the* ODE (8)*, with initial condition* $\omega_0 \in \mathcal{M}(X)$*, has a unique solution. If* Φ *also satisfies* (A2)*, the cone* $\mathcal{M}_+(X)$ *of positive measures is invariant under the semiflow for* $t \geq 0$ *(in other words,* $\mathcal{M}_+(X)$ *is positive invariant). Finally, if* Φ *also satisfies* (A3)*, the norm of positive measures is preserved in forward time. In particular, the convex set* $\mathcal{P}(X)$ *of probability measures is then positive invariant.*

Proof: If Φ is Lipschitz, we can invoke the Picard-Lindelöf Theorem for ODEs on Banach spaces, see [4, Thm. 7.6], so existence and uniqueness of the solution of the abstract Cauchy problem are clear.

If Φ also satisfies (A2), positive invariance of $\mathcal{M}_+(X)$ follows from a continuity argument, see p. 235 and Thm. 16.5 together with Remark 16.6 of [4] for a proof. If Φ is linear, (A2) is the so-called positive minimum principle, and our assertion also follows from [5, Thm. 1.11], which uses a functional analytic proof.

Finally, assume Φ satisfies (A1) – (A3). Let $\omega_0 \in \mathcal{M}_+^m(X)$ be the initial condition and denote the corresponding unique solution of (8) by ω_t. Then, $\omega_t \in \mathcal{M}_+(X)$ for all $t \geq 0$ by the previous argument, so $\|\omega_t\| = \omega_t(X)$. This implies $\frac{\mathrm{d}}{\mathrm{d}t}\|\omega_t\| = \big(\Phi(\omega_t)\big)(X) = 0$ by assumption (A3), so $\|\omega_t\| \equiv \|\omega_0\| = m$. This proves the assertion. $\square$

3 Recombination

This section deals with the nonlinear IPL equation for recombination, and is the core of our article. We develop the results step by step here. The combination with mutation will then be rather painless, and an addition of selection will be discussed after that.

3.1 Recombination on measures

Let X, Y be two locally compact spaces with attached measure spaces $\mathcal{M}(X)$ and $\mathcal{M}(Y)$. If $f : X \to Y$ is a continuous function and $\omega \in \mathcal{M}(X)$, then $f.\omega := \omega \circ f^{-1}$ is an element of $\mathcal{M}(Y)$, where $f^{-1}(y) := \{x \in X \mid f(x) = y\}$ means the preimage of $y \in Y$ in X, with obvious extension to $f^{-1}(B)$, the preimage of a subset $B \subset Y$ in X. Due to the continuity of f, $f^{-1}(B)$ is a Borel set in X if B is a Borel set in Y.

Let $X = X_0 \times \ldots \times X_n$ be as in Section 1, and let, from now on, N and L always denote the set of sites and links as introduced there. In this section, we can entirely work with the Banach space $\mathcal{M}(X)$, equipped with the variation norm $\|.\|$. Let $\pi_i : X \to X_i$ be the canonical *projection* which is continuous. It induces a mapping from $\mathcal{M}(X)$ to $\mathcal{M}(X_i)$ by $\omega \mapsto \pi_i.\omega$, where $(\pi_i.\omega)(E) = \omega(\pi_i^{-1}(E))$, for any Borel set $E \subset X_i$. By (slight) abuse of notation, we will use the symbol π_i also for this induced mapping. It is clear that π_i is linear and maps positive measures to positive measures of the same norm. As such, it is bounded and hence also continuous. In particular, it maps $\mathcal{P}(X)$ to $\mathcal{P}(X_i)$ and may then be understood as marginalization. Likewise, we can start from any (ordered) index set $I \subset N$ and define a projector $\pi_I : \mathcal{M}(X) \to \mathcal{M}(X_I)$ with $X_I := \bigtimes_{i \in I} X_i$. With this notation, $X_N = X$. We will frequently also use the abbreviation $\pi_{<\alpha}$ for the projector $\pi_{\{1,\ldots,\lfloor\alpha\rfloor\}}$, and $\pi_{>\alpha}$ for $\pi_{\{\lceil\alpha\rceil,\ldots,n\}}$. These objects may be understood as 'cut and forget' operators, since they give the distribution of what is left after a cut is made at α, and the trailing resp. leading segment is discarded.

This now enables us to introduce the elementary *recombination operator*, or *recombinator* as we will call it from now on, $R_\alpha : \mathcal{M}(X) \to \mathcal{M}(X)$, for $\alpha \in L$. If $\omega = 0$, $R_\alpha(\omega) := 0$, and otherwise

$$R_\alpha(\omega) \;:=\; \frac{1}{\|\omega\|}\big((\pi_{<\alpha}.\omega) \otimes (\pi_{>\alpha}.\omega)\big) \tag{9}$$

which is a (partial) product measure. Here and in what follows, we tacitly identify (if necessary) a product measure with its unique extension to a regular Borel measure on X, which is justified by Fact 1. The following property is now an immediate consequence of the definition.

Fact 2 *The recombinator R_α maps $\mathcal{M}_+(X)$ into itself and preserves the norm of positive measures. In particular, it maps $\mathcal{P}(X)$ into itself.* $\square$

Let us comment on the choice of (9). Being composed of the cut-and-forget operators for the leading and the trailing ends, $R_\alpha(\omega)$ has the interpetation of a 'cut-and-relink operator', which describes a cut at α, followed by (random) reunion of the resulting segments.

At first sight, it might appear more natural to drop the prefactor $1/\|\omega\|$. However, the norm of a positive measure ω would then not be preserved unless $\|\omega\| = 1$. In view of later extensions, it is more desirable not to be restricted to probability measures, and that is why we prefer (9) which makes R_α positive homogeneous of degree 1,

$$R_\alpha(a\,\omega) \;=\; |a|\, R_\alpha(\omega)\,, \tag{10}$$

for arbitrary $a \in \mathbb{R}$. Note, however, that R_α is *not* a linear operator, not even when restricted to $\mathcal{M}_+(X)$.

Fact 3 *Let $\alpha \in L$. The recombinator R_α satisfies $\|R_\alpha(\omega)\| \leq \|\omega\|$, for all $\omega \in \mathcal{M}(X)$, and is (globally) Lipschitz on $\mathcal{M}(X)$.*

Proof: Let us first observe that, for arbitrary $\omega, \omega' \in \mathcal{M}(X)$ and $\alpha \in L$, we obtain the inequality

$$\|(\pi_{<\alpha}.\omega) \otimes (\pi_{>\alpha}.\omega')\| \;\leq\; \|\omega\|\,\|\omega'\|\,,$$

which is a simple consequence of Hahn's decomposition for real measures, see [41, Thm. 6.14], applied separately to the factors of the product measure. For $0 \neq \omega \in \mathcal{M}(X)$, we then have

$$\|R_\alpha(\omega)\| \;=\; \frac{\|(\pi_{<\alpha}.\omega) \otimes (\pi_{>\alpha}.\omega)\|}{\|\omega\|} \;\leq\; \|\omega\|\,,$$

with equality for positive measures, as stated in Fact 2. Clearly, we also have $R_\alpha(0) = 0$, so that the first assertion follows.

Let $\omega, \omega' \in \mathcal{M}(X)$. If one of them is the 0-measure, say $\omega' = 0$, we have $\|R_\alpha(\omega) - R_\alpha(\omega')\| = \|R_\alpha(\omega)\| \leq \|\omega\| = \|\omega - \omega'\|$. So we may assume both

ω and ω' to be different from 0 and hence to have positive norm. With the above inequalities, we can now employ the following 3ε-type argument

$$
\begin{aligned}
&\|R_\alpha(\omega) - R_\alpha(\omega')\| \\
&\quad = \; \left\| \frac{\big(\pi_{<\alpha}.\omega\big) \otimes \big(\pi_{>\alpha}.(\omega - \omega')\big)}{\|\omega\|} + \frac{\big(\pi_{<\alpha}.(\omega - \omega')\big) \otimes \big(\pi_{>\alpha}.\omega'\big)}{\|\omega'\|} \right. \\
&\qquad\quad \left. + \left(\frac{1}{\|\omega\|} - \frac{1}{\|\omega'\|}\right) \Big[\big(\pi_{<\alpha}.\omega\big) \otimes \big(\pi_{>\alpha}.\omega'\big)\Big] \right\| \\
&\quad \leq \; \frac{\big\|\big(\pi_{<\alpha}.\omega\big) \otimes \big(\pi_{>\alpha}.(\omega - \omega')\big)\big\|}{\|\omega\|} + \frac{\big\|\big(\pi_{<\alpha}.(\omega - \omega')\big) \otimes \big(\pi_{>\alpha}.\omega'\big)\big\|}{\|\omega'\|} \\
&\qquad\quad + \left|\frac{1}{\|\omega\|} - \frac{1}{\|\omega'\|}\right| \cdot \big\|\big(\pi_{<\alpha}.\omega\big) \otimes \big(\pi_{>\alpha}.\omega'\big)\big\| \\
&\quad \leq \; 2\,\|\omega - \omega'\| + \big|\|\omega\| - \|\omega'\|\big| \;\; \leq \;\; 3\,\|\omega - \omega'\| \, .
\end{aligned}
$$

Together, this gives the second assertion, with Lipschitz constant ≤ 3. □

In view of Fact 2, it makes sense to investigate the properties of the recombinators restricted to the positive cone $\mathcal{M}_+(X)$. The crucial property which underlies our later analysis is the following.

Proposition 2 *The elementary recombinators, when restricted to $\mathcal{M}_+(X)$, are idempotents and commute with one another. In other words, we then have $R_\alpha^2 = R_\alpha$ and $R_\alpha R_\beta = R_\beta R_\alpha$ for arbitrary $\alpha, \beta \in L$.*

PROOF: The statement is trivial for the action on $\nu = 0$. So, let $\nu > 0$ be a (strictly) positive measure. We then have $\nu(X) = \|\nu\|$ and obtain

$$
\begin{aligned}
\pi_{<\alpha}.\big((\pi_{<\alpha}.\nu) \otimes (\pi_{>\alpha}.\nu)\big) \;\; &= \;\; \|\nu\|\,(\pi_{<\alpha}.\nu) \\
\pi_{>\alpha}.\big((\pi_{<\alpha}.\nu) \otimes (\pi_{>\alpha}.\nu)\big) \;\; &= \;\; \|\nu\|\,(\pi_{>\alpha}.\nu)
\end{aligned}
$$

in $\mathcal{M}_+(X_{<\alpha})$ resp. $\mathcal{M}_+(X_{>\alpha})$ where we adopt the same index convention for sets as we did for projectors. Using $\|R_\alpha(\nu)\| = \|\nu\|$ from Fact 2 and the linearity of the mappings $\nu \mapsto (\pi.\nu)$, one can now apply the definition of the elementary recombinators to check explicitly that

$$R_\alpha\big(R_\alpha(\nu)\big) \;=\; R_\alpha(\nu)\,.$$

For commutativity, we may again assume $\nu > 0$ and also $\alpha < \beta$. Then

$$\begin{aligned}
\pi_{<\alpha}.\big((\pi_{<\beta}.\nu) \otimes (\pi_{>\beta}.\nu)\big) &= \|\nu\| \, (\pi_{<\alpha}.\nu) \\
\pi_{>\alpha}.\big((\pi_{<\beta}.\nu) \otimes (\pi_{>\beta}.\nu)\big) &= \big((\pi_{\{\lceil\alpha\rceil,\dots,\lfloor\beta\rfloor\}}.\nu) \otimes (\pi_{>\beta}.\nu)\big).
\end{aligned}$$

The first equation can be verified directly, as in the previous case. The second can easily be checked on Borel sets of the product form $E = E_{\{\lceil\alpha\rceil,\dots,\lfloor\beta\rfloor\}} \times E_{>\beta}$, followed by an application of Fact 1. Combining these intermediate results, one obtains

$$R_\alpha\big(R_\beta(\nu)\big) = \frac{1}{\|\nu\|^2} \big((\pi_{<\alpha}.\nu) \otimes (\pi_{\{\lceil\alpha\rceil,\dots,\lfloor\beta\rfloor\}}.\nu) \otimes (\pi_{>\beta}.\nu)\big) = R_\beta\big(R_\alpha(\nu)\big) \, ,$$

which proves our assertion. □

Remark: In view of positive homogeneity of the recombinators, see Eq. (10), it would have been sufficient to prove our assertions on $\mathcal{P}(X)$. The above version, however, shows quite clearly where, and how many, normalization factors $\|\nu\|$ appear in the tensor products. If we restrict ourselves to probability measures below, one should keep this in mind for extending arguments to the full cone, $\mathcal{M}_+(X)$.

A close inspection of the proof of Proposition 2 shows that we have simultaneously proved the following useful property.

Lemma 1 *Let $\nu \in \mathcal{P}(X)$ and $\alpha \in L$. For all $\beta \in L$ with $\beta \geq \alpha$, we have $\pi_{<\alpha}.\big(R_\beta(\nu)\big) = \pi_{<\alpha}.\nu$. Similarly, $\pi_{>\alpha}.\big(R_\beta(\nu)\big) = \pi_{>\alpha}.\nu$, for all $\beta \leq \alpha$.* □

3.2 The IPL equation and its solution

Let us start with a brief description of the recombination process for finite X, and a population of m individuals, each of the form $x = (x_0, x_1, \dots, x_n)$ with $x_i \in X_i$. Every individual carries a Poisson clock at each link $\alpha \in L$, with parameters $\varrho_\alpha > 0$, which do not depend on the individual. If the clock at link α of the individual x rings, a random partner y is picked from the population for recombination at that link. The recombined pair is then $(x_0, \dots, x_{\lfloor\alpha\rfloor}, y_{\lceil\alpha\rceil}, \dots, y_n)$ and $(y_0, \dots, y_{\lfloor\alpha\rfloor}, x_{\lceil\alpha\rceil}, \dots, x_n)$.

To describe the entire population, let $Z_x(t)$ be the random variable that gives the number of x-individuals at time t, and $Z(t)$ the combined random vector with components $Z_x(t)$. Hence, if $Z(t) = z$, and $x \neq y$, we can have transitions from z to $z - u_x - u_y + u_{(x_{<\alpha}, y_{>\alpha})} + u_{(y_{<\alpha}, x_{>\alpha})}$, where we use our short hand notation for indices, and u_x to denote the unit vector corresponding to x. Such a transition occurs at rate $\varrho_\alpha z_x z_y/(m - z_x)$.

Note that this process implies instant mixing of all (geno-)types in the population. This is an idealization which neglects that maternal and paternal genes stay together for the lifetime of an individual. Nevertheless, this is a good and realistic model if recombination events are rare on the time scale of the individual life span. This is certainly true if our sites belong to the DNA sequence of a single gene, or a few adjacent genes. It is then well justified to describe recombination in terms of these first order effects only.

Let us look at the influence of increasing m, whence we write $Z^{(m)}(t)$ to indicate dependence on system size. As $m \to \infty$, the sequence of random processes $Z^{(m)}(t)/m$ converges almost surely to the solution of a differential equation with initial condition $Z^{(m)}(0)/m$ (resp. its limit as $m \to \infty$), see [21, Thm. 11.2.1]. The corresponding IPL equation [6, Eq. 2.5], reformulated in our measure-theoretic setting, reads

$$\dot{\omega} \; = \; \Phi_{\mathrm{rec}}(\omega) \; := \; \sum_{\alpha \in L} \varrho_\alpha \big(R_\alpha - \mathbf{1} \big)(\omega) \, . \tag{11}$$

In line with our strategy for the mutation processes, we take this nonlinear ODE as the general starting point for the recombination analysis on product spaces X built from arbitrary locally compact spaces X_i. We will assume that $\varrho_\alpha > 0$, for all $\alpha \in L$, without loss of generality (if $\varrho_\alpha = 0$, remove the link at α, absorb the pair $(\lfloor \alpha \rfloor, \lceil \alpha \rceil)$ into a single site, and identify $X_{\lfloor \alpha \rfloor} \times X_{\lceil \alpha \rceil}$ with the state space at that site, thus reducing the number of sites (and links) by one).

Proposition 3 *The abstract Cauchy problem of the IPL equation* (11) *has a unique solution. Furthermore, $\mathcal{M}_+(X)$ is positive invariant under the flow, with the norm of positive measures preserved. In particular, $\mathcal{P}(X)$ is positive invariant.*

Proof: Consider $\dot{\omega} = \Phi_{\mathrm{rec}}(\omega)$, which is a special case of (8), so we want to apply Theorem 1. By Fact 3, Φ_{rec} is Lipschitz, so assumption (A1) is satisfied.

Let $\nu \in \mathcal{M}_+(X)$, i.e. $\nu(E) \geq 0$ for all Borel sets $E \subset X$. Let E be any Borel subset of X such that $\nu(E) = 0$. Then

$$\Phi_{\mathrm{rec}}(\nu)(E) \; = \; \sum_{\alpha \in L} \varrho_\alpha R_\alpha(\nu)(E) \; \geq \; 0$$

because each $R_\alpha(\nu)$ is a positive measure and all $\varrho_\alpha > 0$ by assumption, so (A2) is satisfied.

Finally, with Fact 2, it is easy to check that $\Phi_{\text{rec}}(\nu)(X) = 0$ for any positive measure ν, which shows that assumption (A3) is satisfied, too. Theorem 1 then establishes our claims. $\square$

The difficulty in solving (11) stems from the nonlinearity of the right-hand side, so Φ_{rec} cannot be considered as the generator of an exponential semigroup. It is, however, rather natural to expect that the solution should still have a rather similar structure, as the R_α are at least positive homogeneous of degree one and commute with one another. Let us therefore, for any $G \subset L$, introduce the *composite* recombinators

$$R_G \; := \; \prod_{\alpha \in G} R_\alpha \, . \tag{12}$$

They are well-defined on $\mathcal{M}_+(X)$ due to Proposition 2, while an order of the product has to be specified otherwise. In any case, $\|R_G(\omega)\| \leq \|\omega\|$ for all $\omega \in \mathcal{M}(X)$. Note that $R_\varnothing = 1$ and $R_{\{\alpha\}} = R_\alpha$ in this notation. The composite recombinators are again positive homogeneous of degree one. A simple induction argument based on Proposition 2 gives the following result.

Corollary 1 *On $\mathcal{M}_+(X)$, the composite recombinators satisfy*

$$R_G R_H \; = \; R_{G \cup H} \, ,$$

for arbitrary $G, H \subset L$. Furthermore, each R_G maps $\mathcal{M}_+(X)$ into itself and preserves the norm of positive measures. $\square$

Let us pretend for a moment that the idempotents R_α were actually linear operators. In such a case, we would get

$$\exp\bigl(\varrho_\alpha t (R_\alpha - \mathbf{1})\bigr) \; = \; \exp(-\varrho_\alpha t)\mathbf{1} + \bigl(1 - \exp(-\varrho_\alpha t)\bigr) R_\alpha \, .$$

Taking the product over such terms for all $\alpha \in L$ and expanding it would formally lead to the sum

$$\sum_{G \subset L} a_G(t) \, R_G$$

with the coefficient functions

$$a_G(t) \; = \; \Bigl(\prod_{\alpha \in \overline{G}} \exp(-\varrho_\alpha t) \Bigr) \cdot \Bigl(\prod_{\beta \in G} \bigl(1 - \exp(-\varrho_\beta t)\bigr) \Bigr). \tag{13}$$

It will have a touch of magic below when we prove that this little "derivation" actually gives the correct answer! After we have established our main result in Theorem 2, we will come back to these coefficients and give them a probabilistic interpretation. This will also motivate why they are a very reasonable guess to start with.

As mentioned before, the elementary recombinators are *not* linear. Nevertheless, they have a related property on convex combinations. If $\omega = \sum_{i=1}^{k} a_i \nu_i$ is a convex linear combination of positive measures ν_i of equal norm, we get

$$R_\alpha(\omega) \; = \; \sum_{i=1}^{k} a_i \, R_\alpha(\nu_i) \; + \; \mathfrak{R} \tag{14}$$

where one can show, by a rather straight-forward calculation which we omit here, that the remainder $\mathfrak{R}$ is given by

$$\mathfrak{R} \; = \; - \sum_{i<j} a_i a_j \, \frac{\|\nu_i - \nu_j\|}{\|\omega\|} \, R_\alpha(\nu_i - \nu_j) \, .$$

This shows that the recombinators are indeed inherently nonlinear, but also that they might act like linear operators on *special* convex combinations, namely those for which the remainder vanishes. This is precisely what we need to solve our problem.

Proposition 4 *Let ν be a positive measure, $\alpha \in L$ and $a_G(t)$ the coefficient functions of* (13)*. Then, for any fixed $t \geq 0$, we have*

$$R_\alpha\Big(\sum_{G \subset L} a_G(t) \, R_G(\nu) \Big) \; = \; \sum_{G \subset L} a_G(t) \, R_{G \cup \{\alpha\}}(\nu) \, .$$

Before we prove this result, we formulate a special property of the coefficient functions first. Observe that, for fixed $t \geq 0$, $q_\alpha := \exp(-\varrho_\alpha t)$ is a number between 0 and 1. It can be interpreted as a probability (namely that link α has not been hit until time t). With this, the coefficients read

$$a_G^L \; = \; \prod_{\alpha \in L - G} q_\alpha \, \prod_{\beta \in G} (1 - q_\beta) \; \geq \; 0 \tag{15}$$

where we have suppressed the (fixed) time, but added the set of links, L, as an upper index. We can now formulate a crucial factorization property.

Lemma 2 *Let $L = L_1 \dot{\cup} L_2$ be a partition of L, and set $G_i = G \cap L_i$ for an arbitrary $G \subset L$. Then, the coefficients of* (15) *satisfy $a_G^L = a_{G_1}^{L_1} \cdot a_{G_2}^{L_2}$. Furthermore, for any $L^{'} \subset L$, we have*

$$\sum_{H \subset L'} a_H^{L'} \;=\; 1\,.$$

Proof: Since $L_1 \cap L_2 = \varnothing$, the first statement is a direct consequence of the product form of a_G^L in Eq. (15). The normalization property can be verified from the probabilistic interpretation mentioned above. If $1 - q_\alpha$ (resp. q_α) is the probability that link α has (resp. has not) been hit, $a_H^{L'}$ is the probability that, of the links in $L^{'}$, precisely H is spared. Consequently, $\sum_{H \subset L'} a_H^{L'}$ is the sum over the probabilities of all possible events, hence equal to 1. Alternatively, this identity can be derived from a simple Möbius inversion argument, as we show below in Fact 5. □

Proof of Proposition 4: Since the recombinators are positive homogeneous of degree one, it suffices to prove the statement for ν a probability measure. Let $\alpha \in L$ be fixed.

Set $\omega = \sum_{G \subset L} a_G^L \, R_G(\nu)$. Since $\nu \in \mathcal{P}(X)$ implies $\omega \in \mathcal{P}(X)$, we obtain

$$\begin{aligned}
R_\alpha(\omega) \;&=\; R_\alpha\Big(\sum_{G \subset L} a_G^L \, R_G(\nu)\Big) \\
&=\; \Big(\pi_{<\alpha} . \Big(\sum_{G \subset L} a_G^L \, R_G(\nu)\Big)\Big) \otimes \Big(\pi_{>\alpha} . \Big(\sum_{H \subset L} a_H^L \, R_H(\nu)\Big)\Big) \\
&=\; \sum_{G, H \subset L} a_G^L \, a_H^L \, \Big(\big(\pi_{<\alpha} . R_G(\nu)\big) \otimes \big(\pi_{>\alpha} . R_H(\nu)\big)\Big)
\end{aligned}$$

where we have used the linearity of the mappings $\pi_{<\alpha}$ and $\pi_{>\alpha}$.

Let us define $L_1 = \{\frac{1}{2}, \frac{3}{2}, \dots, \alpha\}$ and $L_2 = L - L_1$, so that $L = L_1 \dot{\cup} L_2$ is a partition of L. Also, let $G_i = G \cap L_i$ and $H_i = H \cap L_i$, for $G, H \subset L$. Lemma 1 then tells us that

$$\begin{aligned}
\big(\pi_{<\alpha} . R_G(\nu)\big) \otimes \big(\pi_{>\alpha} . R_H(\nu)\big) \;&=\; \big(\pi_{<\alpha} . R_{G_1}(\nu)\big) \otimes \big(\pi_{>\alpha} . R_{H_2}(\nu)\big) \\
&=\; \big(\pi_{<\alpha} . R_{G_1 \cup H_2}(\nu)\big) \otimes \big(\pi_{>\alpha} . R_{G_1 \cup H_2}(\nu)\big) \\
&=\; R_\alpha\big(R_{G_1 \cup H_2}(\nu)\big).
\end{aligned}$$

Inserting this into the previous equation and invoking Lemma 2 repeatedly gives

$$\begin{aligned}
R_\alpha(\omega) &= \sum_{G,H\subset L} a^L_G\, a^L_H\, R_\alpha\big(R_{G_1\cup H_2}(\nu)\big) \\
&= \sum_{G_1\subset L_1}\sum_{G_2\subset L_2}\sum_{H_1\subset L_1}\sum_{H_2\subset L_2} a^{L_1}_{G_1}\, a^{L_2}_{G_2}\, a^{L_1}_{H_1}\, a^{L_2}_{H_2}\, R_\alpha\big(R_{G_1\cup H_2}(\nu)\big) \\
&= \sum_{G_1\subset L_1}\sum_{H_2\subset L_2} a^{L_1}_{G_1}\, a^{L_2}_{H_2}\, R_\alpha\big(R_{G_1\cup H_2}(\nu)\big) \\
&= \sum_{K\subset L} a^L_K\, R_\alpha\big(R_K(\nu)\big)\,,
\end{aligned}$$

which proves our assertion. □

REMARK: Proposition 4 admits the following interpretation. Let ν be a positive measure, with $\|\nu\| = m > 0$. Then, the $2^{|L|}$ measures $R_G(\nu)$ with $G \subset L$ form the vertices of a $\|.\|$-closed simplex in $\mathcal{M}^m_+(X)$. On some of their convex combinations (in particular along solutions, as we will see shortly), the elementary recombinators R_α act linearly. It is this simplex, foliated into solution curves, to which the entire time evolution is constrained, with ν as the initial condition.

The positive measure in Proposition 4 was arbitrary. This means that, when restricting the action of the R_α's to $\mathcal{M}_+(X)$, we can formulate the rule on the level of operators. Observe that $R_\alpha\big(R_K(\nu)\big) = R_{K\cup\{\alpha\}}(\nu) = R_K(\nu')$ where $\nu' = R_\alpha(\nu)$. By a simple induction argument, we thus arrive at

Corollary 2 *Let $a_G(t)$ be the coefficient function of* (13)*, and let $t \geq 0$ be fixed. On $\mathcal{M}_+(X)$, the recombinators satisfy the equation*

$$R_H\Big(\sum_{G\subset L} a_G(t)\, R_G\Big) \;=\; \sum_{G\subset L} a_G(t)\, R_{G\cup H}$$

for arbitrary $H \subset L$. □

We now assume that the initial condition, ω_0, is a positive measure and make the following *ansatz* for the solution of (11):

$$\omega_t \;=\; \sum_{G\subset L} a_G(t)\, R_G(\omega_0) \tag{16}$$

with the coefficient functions $a_G(t)$ of (13). Note that they do *not* depend on ω_0. The initial values are $a_\varnothing(0) = 1$ and $a_G(0) = 0$ for all $\varnothing \neq G \subset L$.

By Corollary 1, each $R_G(\omega_0)$ is a positive measure with the same norm as ω_0. This implies that, as long as $a_G(t) \geq 0$ for $t \geq 0$, the ansatz for ω_t must form a convex linear combination of positive measures of equal norm if it is a solution of (7). This follows from Eq. (15) together with Lemma 2, or from Fact 5 below.

The time derivative of ω_t of (16) is $\dot{\omega}_t = \sum_{G\subset L} \dot{a}_G(t)\, R_G(\omega_0)$. On the other hand, Proposition 4 means that the R_α act linearly on the convex combination (16), and we obtain

$$
\begin{aligned}
\Phi_{\mathrm{rec}}(\omega_t) &= \sum_{\alpha\in L} \varrho_\alpha \, (R_\alpha - \mathbf{1})(\omega_t) \\
&= \sum_{\alpha\in L} \varrho_\alpha \sum_{G\subset L} a_G(t) \, \big(R_{G\cup\{\alpha\}}(\omega_0) - R_G(\omega_0)\big) \\
&= \sum_{\alpha\in L} \varrho_\alpha \Big[\sum_{\alpha\in \d{G}\subset L} a_{G\setminus\{\alpha\}}(t)\, R_G(\omega_0) \; - \sum_{\alpha\notin \d{G}\subset L} a_G(t)\, R_G(\omega_0) \Big] \\
&= \sum_{G\subset L} \Big[\sum_{\alpha\in G} \varrho_\alpha \, a_{G\setminus\{\alpha\}}(t) \; - \sum_{\beta\in\overline{G}} \varrho_\beta \, a_G(t) \Big] R_G(\omega_0)\,,
\end{aligned}
$$

where we use the notation $\d{G}$ in the third step to indicate the summation variable. It is now a straight-forward calculation to check that the coefficients $a_G(t)$ of (13) indeed satisfy the equations

$$
\dot{a}_G(t) \; = \; \sum_{\alpha\in G} \varrho_\alpha \, a_{G\setminus\{\alpha\}}(t) \; - \sum_{\beta\in\overline{G}} \varrho_\beta \, a_G(t)
$$

and that they constitute a convex combination in (16). Consequently, our ansatz solves the IPL equation (11), and, by Proposition 3, this is the unique solution we are after. We have thus established the following main result.

Theorem 2 *The ansatz* (16) *solves the IPL equation* (11) *with initial condition $\omega_0 \in \mathcal{M}_+(X)$ if and only if the coefficient functions are given by* (13), *i.e. by*

$$
a_G(t) \; = \; \exp\Big(-\sum_{\alpha\in\overline{G}} \varrho_\alpha t\Big) \cdot \prod_{\beta\in G} \big(1 - \exp(-\varrho_\beta t)\big)
$$

for all $G \subset L$. □

Remark: To interpret the coefficient $a_G(t)$, let us consider a single individual. Since $\exp(-\varrho_\alpha t)$ is the probability that link α has experienced no crossover event until time t (recall that we have assumed a Poisson process of rate ϱ_α at link α), $a_G(t)$ may be interpreted as the probability that the set of all links that have, up to time t, experienced at least one crossover event, is precisely G.

Note that the above result relies on the assumption of single, independent crossover events, which is described by recombinators that commute. In more general models, with multiple, dependent events, the algebraic structure is rather involved. This makes solutions much more cumbersome, or considerably less explicit in nature (for review, see [36, Ch. 6]).

Let us come back to the meaning of Eq. (16) in combination with Theorem 2. If φ_t denotes the flow of the IPL equation (11), we obtain, for all $t \geq 0$, the identity

$$\varphi_t \;=\; \sum_{G\subset L} a_G(t)\, R_G \tag{17}$$

which is valid on the cone $\mathcal{M}_+(X)$. As usual, $\varphi_0 = \mathbf{1}$ and $\varphi_t \circ \varphi_s = \varphi_{t+s}$, for all $t, s \geq 0$. This implies the identity

$$a_G(t+s) \;=\; \sum_{\substack{H,K\subset L\\ H\cup K=G}} a_H(t)\, a_K(s)\,,$$

which can be verified by direct computation. More interestingly, we also have

Fact 4 *On $\mathcal{M}_+(X)$, the forward flow of* (11) *commutes with the recombinators, i.e. $R_G \circ \varphi_t = \varphi_t \circ R_G$, for all $t \geq 0$ and $G \subset L$.*

PROOF: Let $\nu \in \mathcal{M}_+(X)$ and fix $G \subset L$. Then

$$\begin{aligned} R_G\big(\varphi_t(\nu)\big) \;&=\; R_G\Big(\sum_{H\subset L} a_H(t) R_H(\nu)\Big) \;=\; \sum_{H\subset L} a_H(t) R_{G\cup H}(\nu) \\ &=\; \sum_{H\subset L} a_H(t) R_H\big(R_G(\nu)\big) \;=\; \varphi_t\big(R_G(\nu)\big) \end{aligned}$$

by an application of Corollary 2. □

Once the solution is known, the remaining task is to identify linear combinations of the $R_H(\omega)$ that decouple from each other and decay exponentially. To this end, we employ combinatorial techniques to regroup the terms of the solution according to their exponential damping factors. Let us first expand the expression for $a_G(t)$,

$$a_G(t) \;=\; \sum_{K\subset G} (-1)^{|G-K|} \exp\Big(-\sum_{\alpha\in\overline{K}} \varrho_\alpha t\Big)\,.$$

This suggests to define new functions $b_K(t)$ via

$$b_K(t) \; = \; \exp\bigl(-\sum_{\alpha\in\overline{K}} \varrho_\alpha t\bigr)\,, \tag{18}$$

with the usual convention that the empty sum is 0. In particular, we have $b_{\varnothing}(t) = a_{\varnothing}(t) = \exp\bigl(-\sum_{\alpha\in L} \varrho_\alpha t\bigr)$ and $b_K(0) = 1$ for all $K \subset L$. Now, the Möbius inversion of (2) and (3), used backwards, gives us the relation

$$b_K(t) \; = \; \sum_{G\subset K} a_G(t)\,.$$

One immediate consequence is

$$\sum_{G\subset L} a_G(t) \; = \; b_L(t) \; = \; 1\,. \tag{19}$$

So, together with the observation that the functions $a_G(t)$ of Theorem 2 are always non-negative, we have independently confirmed

Fact 5 *If $\omega_0 \in \mathcal{M}_+(X)$, the coefficient functions $a_G(t)$ of Theorem 2 constitute a convex linear combination of positive measures in Eq. (16).* $\square$

The significance of the new functions becomes clear by realizing that there is an analogue on the level of operators. To this end, we rewrite the composite recombinators in terms of new operators via $R_H = \sum_{G \dot\supset H} T_G$ and obtain, by an obvious variant of Möbius inversion,

$$T_G \; := \; \sum_{H \dot\supset G} (-1)^{|H-G|}\, R_H\,. \tag{20}$$

A straight-forward calculation then reveals that

$$\omega_t \; = \; \sum_{G\subset L} a_G(t)\, R_G(\omega_0) \; = \; \sum_{K\subset L} b_K(t)\, T_K(\omega_0)\,. \tag{21}$$

Note that, as a consequence of Eqs. (10), (12) and (20), the operators T_G are positive homogeneous of degree one, i.e.

$$T_G(\alpha\nu) \; = \; |\alpha| \cdot T_G(\nu)\,. \tag{22}$$

Let us now introduce new measures $\nu_G(t) := b_G(t)\, T_G(\omega_0)$, which are elements of $\mathcal{M}(X)$, but no longer positive in general.

Proposition 5 *The signed measures $\nu_G(t)$ solve the Cauchy problem*

$$\dot{\nu}_G(t) \;=\; -\Big(\textstyle\sum_{\alpha\in\overline{G}}\varrho_\alpha\Big)\,\nu_G(t)$$

with initial condition $\nu_G(0) = T_G(\omega_0)$, for all $G \subset L$.

PROOF: The result is a direct consequence of the fact that the coefficient functions $b_G(t)$ solve the ordinary initial value problems

$$\dot{b}_G(t) \;=\; -\Big(\textstyle\sum_{\alpha\in\overline{G}}\varrho_\alpha\Big)\,b_G(t)$$

with initial conditions $b_G(0) = 1$, see above. □

So, the transformation (20) resulted in regrouping the terms of the solution to the IPL equation (11) according to their exponential decay factors in time. In particular,

$$\nu_L(t) \;\equiv\; T_L(\omega_0) \;=\; R_L(\omega_0) \;=\; \bigotimes_{i=0}^{n}(\pi_i.\omega_0)$$

is the unique *limit measure* of the process starting from ω_0. Due to the action of R_L, it is a complete product measure and reflects total independence, and we obtain $\omega_t \to R_L(\omega_0)$ as $t \to \infty$ in the $\|.\|$-topology. This is so because

$$\|\omega_t - \nu_L(t)\| \;=\; \Big\|\sum_{K\subsetneq L} b_K(t)\,T_K(\omega_0)\Big\| \;\le\; \sum_{K\subsetneq L} b_K(t)\,\|T_K(\omega_0)\|$$

where all remaining coefficient functions $b_K(t)$, i.e. those with $K \subsetneq L$, decay exponentially (recall that $\varrho_\alpha > 0$ for all $\alpha \in L$).

3.3 Linkage disequilibria

Starting from the measures $\nu_G(t)$, we will now identify a minimal, complete set of variables by evaluating certain k-point cylinder functions (called k-point functions from now on) or correlation functions known as linkage disequilibria in genetics. They are important for data analysis because they allow to evaluate associations between sites up to a given order from measured type frequencies, and average over all others by marginalization. This way, a certain amount of stochasticity, which is present in all real (finite) populations, is smoothed out.

Various different definitions of linkage disequilibria are available in the literature (see [12, p. 183–186] for an overview). But only special choices decouple (see [16, 17]), and these are the linkage disequilibria we are after. In view of the applications, we will now restrict ourselves to the case that X is a *finite* set, although the results hold, with only minor modifications, also more generally. Eq. (21) and Proposition 5 suggest to employ the signed measures $T_G(\omega_0)$. The corresponding functions $b_G(t)$ will then describe their evolution in time.

Let $\langle j_1, \ldots, j_k \rangle$, with $j_1 < \cdots < j_k$, symbolically denote a cylinder set in $X = X_N$ which is specified at sites j_i, for $1 \leq i \leq k$. More specifically, these are sets of the product form

$$\langle j_1, \ldots, j_k \rangle \;=\; X_{\{0,\ldots,j_1-1\}} \times \{x_{j_1}\} \times [\ldots] \times \{x_{j_k}\} \times X_{\{j_k+1,\ldots,n\}}$$

where [...] contains factors $\{x_i\}$ or X_i depending on whether i appears in $\langle j_1, \ldots, j_k \rangle$ or not. For $\nu \in \mathcal{P}(X)$ and arbitrary $\alpha \in L$, we then have

$$\begin{aligned} R_\alpha(\nu)\big(\langle j_1, \ldots, j_k \rangle\big) &= \big((\pi_{<\alpha}.\nu) \otimes (\pi_{>\alpha}.\nu)\big)\big(\langle j_1, \ldots, j_k \rangle\big) \\ &= \begin{cases} \nu\big(\langle j_1, \ldots, j_k \rangle\big) & \text{if } \alpha < j_1 \text{ or } \alpha > j_k \\ \nu\big(\langle j_1, \ldots, j_s \rangle\big)\, \nu\big(\langle j_{s+1}, \ldots, j_k \rangle\big) & \text{if } j_s < \alpha < j_{s+1}\,. \end{cases} \end{aligned}$$

For later convenience, we also define $\langle \varnothing \rangle = X$ so that $R_\alpha(\nu)\big(\langle \varnothing \rangle\big) = 1$.

Lemma 3 *If $\nu \in \mathcal{P}(X)$, we have $T_G(\nu)\big(\langle j_1, \ldots, j_k \rangle\big) = 0$ whenever the set $\overline{G}$ contains an element that is less than j_1 or larger than j_k.*

Proof: Let $I = \{\beta \mid j_1 < \beta < j_k\}$. Assume there is an $\alpha \in \overline{G \cup I}$. Then

$$\begin{aligned} & T_G(\nu)\big(\langle j_1, \ldots, j_k \rangle\big) \\ &\quad = \sum_{\substack{H \supset G \\ \cdot}} (-1)^{|H-G|} R_H(\nu)\big(\langle j_1, \ldots, j_k \rangle\big) \\ &\quad = \sum_{\substack{H \supset G \\ \cdot}} (-1)^{|H-G|} R_{H \cap I}(\nu)\big(\langle j_1, \ldots, j_k \rangle\big) \\ &\quad = \sum_{\substack{\alpha \notin H \supset G \\ \cdot}} \big((-1)^{|H-G|} + (-1)^{|(H \cup \{\alpha\})-G|}\big) R_{H \cap I}(\nu)\big(\langle j_1, \ldots, j_k \rangle\big) \end{aligned}$$

where the previous calculation was used in the second step, and summation is over H. Clearly, the last expression vanishes. $\square$

Let us now define the time-dependent k-point functions as

$$F_G^{(t)}(j_1, \ldots, j_k) \;=\; T_G(\omega_t)\big(\langle j_1, \ldots, j_k\rangle\big) \tag{23}$$

for arbitrary $G \subset L$, where the notation is again symbolic in that we only specify the positions j_i, but not the corresponding values. To relate this to Eq. (21), we show

Proposition 6 *If $\omega_0 \in \mathcal{M}_+(X)$, we have $T_G(\omega_t) = b_G(t)\, T_G(\omega_0)$, for all $G \subset L$ and $t \geq 0$.*

PROOF: Since $b_G(0) = 1$, equality holds for $t = 0$, and the claim follows if we show that $T_G(\omega_t)$ and $b_G(t)\, T_G(\omega_0)$ satisfy the same differential equation. With $\omega_t = \varphi_t(\omega_0)$, compare (17), we obtain

$$\begin{aligned}
\frac{d}{dt} T_G(\omega_t) \;&=\; \frac{d}{dt} \sum_{H \supseteq G} (-1)^{|H-G|} R_H\big(\varphi_t(\omega_0)\big) \\
&=\; \sum_{H \supseteq G} (-1)^{|H-G|}\, \frac{d}{dt} \varphi_t\big(R_H(\omega_0)\big) && \text{(by Fact 4)} \\
&=\; \sum_{H \supseteq G} (-1)^{|H-G|}\, \Phi_{\mathrm{rec}}\big(\varphi_t(R_H(\omega_0))\big) && \text{(by Eq. (11))} \\
&=\; \sum_{H \supseteq G} (-1)^{|H-G|} \sum_{\alpha \in \overline{G}} \varrho_\alpha \big(R_{H \cup \{\alpha\}} - R_H\big)(\omega_t) && \text{(by Fact 4)} \\
&=\; -\Big(\sum_{\alpha \in \overline{G}} \varrho_\alpha\Big)\, T_G(\omega_t)\,.
\end{aligned}$$

The last step is correct because

$$\sum_{\alpha \in \overline{G}} \varrho_\alpha \sum_{H \supseteq G} (-1)^{|H-G|} R_{H \cup \{\alpha\}} \;=\; 0$$

by an argument analogous to the one used in the last step of the proof of Lemma 3. Now, a comparison with Proposition 5 establishes the claim. $\square$

Even after Lemma 3, there are still too many functions around. It is thus reasonable to select an independent set from them. To see how to do this, assume that we have an index $\alpha \in G \cap I$, with $I = \{\beta \mid j_1 < \beta < j_k\}$ for a cylinder set of type $\langle j_1, \ldots, j_k\rangle$ as above. Let H be a subset of L that contains G, so $\alpha \in H$ in particular, and $\nu \in \mathcal{P}(X)$. Then $R_H(\nu) = R_\alpha(\nu_H)$

with $\nu_H = R_{H \setminus \{\alpha\}}(\nu)$. The little calculation before Lemma 3 now tells us that

$$
\begin{aligned}
R_H(\nu)\big(\langle j_1,\ldots,j_k\rangle\big) &= R_\alpha(\nu_H)\big(\langle j_1,\ldots,j_k\rangle\big) \\
&= \big[\nu_H\big(\langle j_1,\ldots,j_s\rangle\big)\big]\cdot\big[\nu_H\big(\langle j_{s+1},\ldots,j_k\rangle\big)\big] \\
&= \big[R_\alpha(\nu_H)\big(\langle j_1,\ldots,j_s\rangle\big)\big]\cdot\big[R_\alpha(\nu_H)\big(\langle j_{s+1},\ldots,j_k\rangle\big)\big] \\
&= \big[R_H(\nu)\big(\langle j_1,\ldots,j_s\rangle\big)\big]\cdot\big[R_H(\nu)\big(\langle j_{s+1},\ldots,j_k\rangle\big)\big]
\end{aligned}
$$

where $j_s < \alpha < j_{s+1}$. Consequently, defining $I_1 = \{\beta \mid j_1 < \beta < j_s\}$ and $I_2 = \{\beta \mid j_{s+1} < \beta < j_k\}$, and referring back to (20), we also get

$$
\begin{aligned}
&T_G(\nu)\big(\langle j_1,\ldots,j_k\rangle\big) \\
&= \sum_{H \supset G} (-1)^{|H-G|}\, \big[R_H(\nu)\big(\langle j_1,\ldots,j_s\rangle\big)\big]\cdot\big[R_H(\nu)\big(\langle j_{s+1},\ldots,j_k\rangle\big)\big] \\
&= \sum_{H \supset G} (-1)^{|H-G|} \sum_{\substack{K_1 \supset H \\ K_2 \supset H}} \big[T_{K_1}(\nu)\big(\langle j_1,\ldots,j_s\rangle\big)\big]\cdot\big[T_{K_2}(\nu)\big(\langle j_{s+1},\ldots,j_k\rangle\big)\big] \\
&= \sum_{H \supset G} (-1)^{|H-G|} \sum_{\substack{K_1 \supset (H \cup \overline{I}_1) \\ K_2 \supset (H \cup \overline{I}_2)}} \big[T_{K_1}(\nu)\big(\langle j_1,\ldots,j_s\rangle\big)\big]\cdot\big[T_{K_2}(\nu)\big(\langle j_{s+1},\ldots,j_k\rangle\big)\big]
\end{aligned}
$$

where Lemma 3 was used in the last step to remove terms that vanish. This equation means that $T_G(\nu)\big(\langle j_1,\ldots,j_k\rangle\big)$, whenever an $\alpha \in G \cap I$ exists, either vanishes (if Lemma 3 applies) or is a polynomial expression in ℓ-point functions with $\ell < k$.

In the above calculation, ν is an arbitrary probability measure, wherefore the equations apply to ω_t, for an arbitrary $t \geq 0$. Whenever $G \cap I \neq \varnothing$, the time-dependent k-point functions are polynomially dependent of ℓ-point functions with $\ell < k$. Consequently, they do not contain new information. So far, we have:

Proposition 7 *The k-point function $F_G^{(t)}(j_1,\ldots,j_k) = T_G(\omega_t)\big(\langle j_1,\ldots,j_k\rangle\big)$ can only be non-vanishing and (polynomially) independent from ℓ-point functions with $\ell < k$ if $G = \overline{I} = \{\beta < j_1\} \cup \{\beta > j_k\}$.* □

We choose this collection of k-point functions as our linkage equilibria.

Let us finally observe that the summation of a k-point function over all possible values x_i of one of the specified X_i (i.e. marginalization) reduces it to a $(k-1)$-point function, so we have one extra (linear) relation. This means that only $M_i - 1$ possible values can be prescribed independently at site i. On the other hand, given $\langle j_1, \ldots, j_k \rangle$, there is only one way to choose G due to Proposition 7, and then there are $(M_{j_1} - 1) \cdot \ldots \cdot (M_{j_k} - 1)$ different and independent choices to specify the actual values at the sites. Summing up all these possibilities results in

$$\sum_{D \subset N} \prod_{i \in D} (M_i - 1) \;=\; \prod_{i=0}^{n} \bigl(1 + (M_i - 1)\bigr) \;=\; \prod_{i=0}^{n} M_i \;=\; |X| \, .$$

This means that we have singled out the right number of functions. In view of Proposition 7, for t arbitrary but fixed, they completely determine the value of the signed measures $T_G(\omega_t)$ on all cylinder sets. These, in turn, are closed under finite intersections and generate the full σ-algebra of the (finite) space X, so all measures $T_G(\omega_t)$, and hence also ω_t, are uniquely specified, and we have achieved our goal. An explicit example has been worked out in Section 4 of [6], where the k-point functions $F_G^{(t)}(j_1, \ldots, j_k)$ appear as the components of the vector z of linkage disequilibria, up to a change of basis in the local site spaces.

If X is not a finite set, one has to use a generating family of Borel cylinder sets instead of just singleton sets, and invoke Fact 1. Although there is no simple counting argument, the general structure is still similar.

At this point, one could still argue that k-point functions w.r.t. the selection of *sites*, as our $F_G^{(t)}(j_1, \ldots, j_k)$, should be replaced by proper k-point *correlation* functions because these separate off all contributions of functions of lower order, i.e. of ℓ-point functions with $\ell < k$. This is just another application of the Möbius inversion principle, but one where *all* partitions (rather than only ordered ones) are needed. We provide the corresponding formulas in the Appendix. If one performs the necessary calculations, one quickly realizes that our previous inclusion-exclusion process w.r.t. ordered partitions of the *links* has far reaching consequences: most of the potential correction terms simply vanish, as a result of Lemma 3. In particular, we obtain

Theorem 3 *Let $S = \{j_1, \ldots, j_k\}$ be a set of site indices, in increasing order and without gaps, and let $G = \{\alpha < j_1\} \cup \{\alpha > j_k\}$. Then, the k-point function $F_G^{(t)}(j_1, \ldots, j_k) = T_G(\omega_t)\bigl(\langle j_1, \ldots, j_k \rangle\bigr)$ coincides with the corresponding k-point correlation function as given in Eq.* (43) *of the Appendix.*

These functions, for all possible choices of the set S, form a polynomially independent set of linkage disequilibria.

PROOF: We apply Lemma 3 with $\nu = \omega_t$. Due to the assumption on S versus G, the right-hand side of Eq. (43) boils down to the one term we already have, because all other terms vanish. Propositions 7 and 6 ensure the polynomial independence of these objects, which are our linkage disequilibria. $\square$

This result does not extend to all k-point functions. If, for a given k-point function, a non-vanishing correction term occurs in the corresponding correlation function, this will, in general, not decay with the same exponential rate as the original k-point function. So, grouping according to decay rates and according to correlation structures simultaneously is not possible in general. It is a rather remarkable fact that the set of linkage disequlibria is a set of exceptions, and one (as we demonstrated above for the case of discrete state spaces) that completely determines the probability measure.

4 Mutation and recombination

In this section, we will just combine the results of the previous two sections. This is possible because, as we will see, mutation and recombination are independent in our approach, i.e. the corresponding operators in the IPL equation commute. This is to be expected given the fact that mutation acts on the *sites* while recombination works via the *links*. However, to be able to formulate this in a more general situation than X finite or discrete, we now restrict ourselves to the Banach space $\mathcal{M}^{\otimes} = \bigotimes_{i\in N} \mathcal{M}(X_i)$ which, as explained earlier, is meant as the completion of the algebraic tensor product. In general, it is a (true) Banach subspace of $\mathcal{M}(X)$. Our IPL equation now reads

$$\dot{\omega} \; = \; \Big(\sum_{i\in N} \mu_i \, Q_i \, + \sum_{\alpha\in L} \varrho_\alpha \big(R_\alpha - \mathbf{1} \big) \Big) \, (\omega) \tag{24}$$

where we have taken the liberty to introduce mutation rates μ_i, all of which are assumed to be strictly positive. The idea behind this is to use some standardized version for the mutation operators Q_i of (5) so that the μ_i serve as relative coefficients, in line with the usual practice in the biological literature. The linear operators Q_i are supposed to be bounded, hence continuous, and thus possess a unique extension to $\mathcal{M}^{\otimes}$, compare [45, Thms. II.1.2 and II.1.5]. To show consistency, we observe

Lemma 4 *The Banach space $\mathcal{M}^{\otimes}$ is invariant under R_α, for all $\alpha \in L$, and hence positive invariant under the flow of* (24).

Proof: It is clear that R_α maps a finite linear combination of product measures onto another linear combination of this kind, compare the proof of Prop. 4. Since such linear combinations are dense in $\mathcal{M}^{\otimes}$ and R_α is Lipschitz on $\mathcal{M}(X)$, it maps the closed subspace $\mathcal{M}^{\otimes}$ of $\mathcal{M}(X)$ into itself. The statement on positive invariance is a direct consequence of [4, Thm. 16.5 and Remark 16.6]. □

Let $\mathcal{P}^{\otimes}$ be the subspace of probability measures in $\mathcal{M}^{\otimes}$. Referring back to Propositions 1 and 3 and to Theorem 1, the following result is immediate.

Proposition 8 *The abstract Cauchy problem of the IPL equation* (24)*, with initial condition $\omega_0 \in \mathcal{M}^{\otimes}$, has a unique solution. The cone $\mathcal{M}^{\otimes}_{+}$ is positive invariant, and the norm of a positive measure is preserved in forward time. In particular, the convex set $\mathcal{P}^{\otimes}$ is positive invariant.* □

To continue, let us call a positive linear operator W on $\mathcal{M}^{\otimes}$ *strictly positive* if $\omega \in \mathcal{M}^{\otimes}_{+}$ with $\omega > 0$ implies $W\omega > 0$. The key observation is now

Lemma 5 *Let W be a strictly positive bounded linear operator on $\mathcal{M}^{\otimes}$ which has a complete tensor product structure, i.e. $W = w_0 \otimes \cdots \otimes w_n$. On $\mathcal{M}^{\otimes}_{+}$, the elementary recombinator R_α then commutes with W, i.e. $WR_\alpha = R_\alpha W$. In particular, this is true if $W = \exp(tQ_i)$ is an element of a Markov semigroup, as in Section 2, for any $t \geq 0$, $i \in N$ and $\alpha \in L$.*

Proof: Let us first consider the case that W preserves the norm of a positive measure ν, i.e. $\|W\nu\| = \|\nu\|$. This is also true of R_α, $\alpha \in L$. Since W is linear and R_α positive homogeneous of degree 1, it is sufficient to prove the claim on $\mathcal{P}^{\otimes}$. So, let $\nu \in \mathcal{P}^{\otimes}$. W has a complete tensor product structure, so $W = W_{<\alpha} \otimes W_{>\alpha}$ in particular. Observe first that $W_{<\alpha} \circ \pi_{<\alpha} = \pi_{<\alpha} \circ W$ and $W_{>\alpha} \circ \pi_{>\alpha} = \pi_{>\alpha} \circ W$. These relations certainly hold when applied to a product measure $\nu = \nu_{<\alpha} \otimes \nu_{>\alpha}$, but, due to linearity of all mappings involved here, also on arbitrary (finite) linear combinations of measures of this kind. The latter are dense in $\mathcal{M}^{\otimes}$, so that continuity of the mappings establishes the relations, compare [45, Thm. II.1.5].

As a consequence, we obtain

$$W\big((\pi_{<\alpha}.\nu) \otimes (\pi_{>\alpha}.\nu)\big) \;=\; \big(\pi_{<\alpha}.(W\nu)\big) \otimes \big(\pi_{>\alpha}.(W\nu)\big)\,,$$

which proves the assertion for the case that W preserves the norm of ν.

Let us now consider the general case. The proof so far only required that W preserved the norm of the single ν under consideration. We employ again positive homogeneity of R_α. If $\nu > 0$, we have $W\nu > 0$ by assumption, so that $a := \frac{\|\nu\|}{\|W\nu\|} > 0$ is well defined. So we obtain $\|\nu\| = \|aW\nu\|$ and

$$\begin{aligned}
W R_\alpha(\nu) &= \tfrac{1}{a}\,(aW) R_\alpha(\nu) \\
&= \tfrac{1}{a}\, R_\alpha(aW\nu) \qquad \text{(by above argument)} \\
&= R_\alpha(W\nu) \qquad \text{(by Eq. (10))}
\end{aligned}$$

which proves the first assertion.

The second claim is obvious because elements of a Markov semigroup are strictly positive and because the generators Q_i, compare Eq. (5), have the required product structure. □

We can now put together our previous efforts. The obvious form of the solution of (24) is now

$$\omega_t \;=\; \exp\bigl(tQ\bigr) \sum_{G \subset L} a_G(t)\, R_G\,(\omega_0) \tag{25}$$

with $Q = \sum_{i=0}^{n} \mu_i Q_i$ and the coefficient functions $a_G(t)$ of Theorem 2. The verification that this indeed solves the IPL equation is a simple application of the product rule. Let $\nu_t := \sum_{G\subset L} a_G(t)\, R_G(\omega_0)$, so that $\omega_t = \exp(tQ)\, \nu_t$. Then we have

$$\begin{aligned}
\dot{\omega}_t &= Q\,\omega_t + \exp\bigl(tQ\bigr)\dot{\nu}_t \\
&= Q\,\omega_t + \exp\bigl(tQ\bigr) \sum_{\alpha\in L} \varrho_\alpha \bigl(R_\alpha - \mathbf{1}\bigr)(\nu_t) \qquad \text{(by Theorem 2)} \\
&= Q\,\omega_t + \Bigl(\sum_{\alpha\in L} \varrho_\alpha \bigl(R_\alpha - \mathbf{1}\bigr)\Bigr)\bigl(\exp(tQ)\nu_t\bigr) \qquad \text{(by Lemma 5)} \\
&= \Bigl(\sum_{i\in N} \mu_i\, Q_i \,+ \sum_{\alpha\in L} \varrho_\alpha \bigl(R_\alpha - \mathbf{1}\bigr)\Bigr)\,(\omega_t)\,.
\end{aligned}$$

So, together with Proposition 8, we have established:

Theorem 4 *The unique solution of the IPL equation* (24)*, with initial condition* $\omega_0 \in \mathcal{M}_+^{\otimes}$*, is given by* ω_t *of* (25)*, with the coefficient functions* $a_G(t)$ *of Theorem* 2. □

Let us take a closer look at the asymptotic behaviour. Since $a_G(t)$ decreases, as $t \to \infty$, exponentially to 0 unless $G = L$, we obtain

$$\begin{aligned} \omega_t &\sim \exp(tQ)\big(a_L(t) R_L(\omega_0)\big) \sim \exp(tQ) \bigotimes_{i=0}^{n} (\pi_i . \omega_0) \\ &= \bigotimes_{i=0}^{n} \big(\exp(tq_i)(\pi_i . \omega_0)\big) \end{aligned}$$

where all neglected terms are of lower order in that they vanish exponentially (recall that $\exp(tQ)$ and $\exp(tq_i)$ are Markov). This shows that the stationary measure, for any initial measure ω_0, is again a complete product measure[3]. Whether or not there is a unique global equilibrium measure then depends on the properties of the local mutation operators q_i. In the case that X is finite, uniqueness follows if all these generators are irreducible.

What remains to be done is to extend the Möbius trick and to evaluate the linkage disequilibria also for this case. Due to Lemma 5, we can equivalently write ω_t of (25) as

$$\omega_t = \sum_{G \subset L} a_G(t)\, R_G\big(\exp(tQ)\, \omega_0\big) . \tag{26}$$

At any fixed instant of time, $\exp(tQ)\, \omega_0$ is a positive measure, and we can employ Eq. (21) to obtain

$$\omega_t = \sum_{K \subset L} b_K(t)\, T_K\big(\exp(tQ)\, \omega_0\big) \tag{27}$$

with the functions $b_K(t)$ introduced in (18).

If we now assume again that X is finite, we can use the k-point cylinder functions as before to select a finite set of linkage disequilibria that completely determine the solution ω_t. They are the functions

$$F_G^t(j_1, \dots, j_k) = T_G(\omega_t)\big(\langle j_1, \dots, j_k \rangle\big) \tag{28}$$

for $G \subset L$ and selected cylinder sets $\langle j_1, \dots, j_k \rangle$ exactly as before.

Since mutation and recombination are independent of each other and the time evolutions commute, we can separate the time decay due to the two processes. The effect is as follows. Recombination is sensitive to sites

[3] Convergence to product measures is also known from various interacting particle systems, compare [8].

selected in the cylinder sets, but not to the actual values prescribed there. Mutation, in turn, has a tensor product structure with respect to the sites (which expresses the independence of individual events).

If $\exp(tQ)$ is Markov (so that Lemma 5 applies), it is easy to derive (in analogy with the proof of Proposition 6) that

$$\frac{d}{dt} T_G(\omega_t) \;=\; \Bigl(Q - \sum_{\alpha\in\overline{G}} \varrho_\alpha\Bigr)\, T_G(\omega_t)\,. \tag{29}$$

This shows how the recombination rates and the eigenvalues of Q together determine the fine structure of exponential decay. Note that diagonalizing Q (if at all possible) now corresponds to taking appropriate linear combinations of $T_G(\omega_t)\bigl(\langle j_1,\dots,j_k\rangle\bigr)$ for fixed G and $j_1,\dots,j_k$, but different values prescribed at the sites. For finite X, this has been worked out in [6], along with explicit examples.

5 Selection

Let us first look at selection in a slightly more general way, i.e. via an IPL equation on $\mathcal{M}(X)$ without explicit reference to its tensor product structure. Let $P\colon \mathcal{M}(X) \to \mathcal{M}(X)$ be a bounded linear operator which generates a positive semigroup. According to [5, Thm. 1.11], the latter is true if and only if P satisfies our assumption (A2), the positive minimum principle. Consider now the ODE

$$\dot{\omega} \;=\; \Phi_{\rm sel}(\omega) \;:=\; P\omega - \frac{P\omega(X)}{\|\omega\|}\,\omega \tag{30}$$

where $\Phi_{\rm sel}(0) = 0$ is the proper extension of $\Phi_{\rm sel}(\omega)$ to $\omega = 0$. This is motivated by the standard selection model (cf. [27]), where, in properly coordinatized form as indicated in Section 1, P is a diagonal matrix which keeps track of the 'fitness' of the various states, and $\frac{P\omega(X)}{\|\omega\|}$ is the 'mean fitness' of the population. This model also arises in the infinite population limit of the well-known Moran model, see [33] or [22, Ch. 3]. Here, in a population of m individuals with finite state space X as described in Section 3, every individual of type x reproduces at rate r_x, and the offspring replaces a randomly chosen individual in the population (possibly its own parent). Therefore, a transition from population state z to $z + u_x - u_y$ occurs at rate $r_x z_x z_y / m$. Along the lines of Section 3, the limit $m \to \infty$ yields a special case of the differential equation (30), where P is the diagonal matrix with elements r_x.

The more general form used here does not only cover more general X, but also interaction between mutation and reproduction (as opposed to the independent processes considered so far), e.g. the production of mutated offspring on the occasion of reproduction. In any case, the subtraction of the second term on the right hand side of (30) comes from the preservation of total mass, or, in more technical terms, is designed so that Φ_{sel} satisfies assumptions (A2) and (A3) from Section 2.

So far, our selection equation seems to imply that selection acts on haploids (i.e. individuals with only one copy of the genetic information per cell). If, however, individuals have two copies that are equivalent and do not interact (the diploid case without dominance), Eq. (30) is replaced by

$$\dot{\omega} \;=\; \frac{M(\omega \otimes P\omega + P\omega \otimes \omega)}{\|\omega\|} \;-\; \frac{\big(M(\omega \otimes P\omega + P\omega \otimes \omega)\big)(X)}{\|\omega\|^2}\,\omega\,, \qquad (31)$$

where $M(\mu \otimes \nu) := \nu(X) \cdot \mu$ denotes marginalization with respect to the second factor. In this formulation, the mean fitness is

$$\frac{1}{\|\omega\|^2}\big(M(\omega \otimes P\omega + P\omega \otimes \omega)\big)(X) \;=\; 2\,\frac{\omega(X) P\omega(X)}{\|\omega\|^2}\,.$$

For positive ω, the right-hand side of (31) becomes

$$\frac{P\omega(X)}{\|\omega\|}\,\omega \;+\; P\omega \;-\; 2\,\frac{P\omega(X)}{\|\omega\|}\,\omega \;=\; P\omega - \frac{P\omega(X)}{\|\omega\|}\,\omega\,,$$

that is, the diploid equation reduces to the haploid one in this case, in the sense that the flow is the same on $\mathcal{M}_+(X)$.

Let us now take a closer look at the differential equation (30).

Fact 6 *The mapping* $\Phi_{\text{sel}}\colon \mathcal{M}(X) \to \mathcal{M}(X)$ *is (globally) Lipschitz.*

Proof: Consider $\omega, \omega' \in \mathcal{M}(X)$. If one of them is the zero measure, ω' say, we get

$$\begin{aligned}\big\|\Phi_{\text{sel}}(\omega) - \Phi_{\text{sel}}(0)\big\| \;&=\; \big\|\Phi_{\text{sel}}(\omega)\big\| \;\le\; \|P\omega\| + \frac{|P\omega(X)|}{\|\omega\|}\,\|\omega\| \\ &\le\; 2\,\|P\omega\| \;\le\; 2\,\|P\|\,\|\omega\|\end{aligned}$$

where $\|P\| := \sup_{\|\omega\|\leq 1} \|P\omega\| < \infty$ because P is a bounded operator by assumption, and clearly $|P\omega(X)| \leq |P\omega|(X) = \|P\omega\|$.

Let now ω, ω' both be non-zero. Then

$$\left\|\Phi_{\text{sel}}(\omega) - \Phi_{\text{sel}}(\omega')\right\| \;\leq\; \|P\|\,\|\omega-\omega'\| + \left\|\frac{P\omega'(X)}{\|\omega'\|}\,\omega' - \frac{P\omega(X)}{\|\omega\|}\,\omega\right\|.$$

Observe that $\frac{P\omega(X)}{\|\omega\|} = P\big(\frac{\omega}{\|\omega\|}\big)(X)$. The second term on the right hand side of the above equation is then clearly majorized by $\|P\|\,\|\omega'-\omega\| + c\,\|\omega\|$ where

$$\begin{aligned} c \;&=\; \left| P\Big(\frac{\omega'}{\|\omega'\|}\Big)(X) - P\Big(\frac{\omega}{\|\omega\|}\Big)(X)\right| \;=\; \left| P\Big(\frac{\omega'}{\|\omega'\|} - \frac{\omega}{\|\omega\|}\Big)(X)\right| \\ &\leq\; \left\| P\Big(\frac{\omega'}{\|\omega'\|} - \frac{\omega}{\|\omega\|}\Big)\right\| \;\leq\; \frac{\|P\|}{\|\omega\|\,\|\omega'\|}\,\Big\|\, \|\omega\|\,\omega' - \|\omega'\|\,\omega\,\Big\|\,. \end{aligned}$$

Next, observe that

$$\Big\|\,\|\omega\|\,\omega' - \|\omega'\|\,\omega\,\Big\| \;\leq\; \|\omega'\|\,\big|\|\omega\| - \|\omega'\|\big| + \|\omega'\|\,\|\omega'-\omega\| \;\leq\; 2\,\|\omega'\|\,\|\omega-\omega'\|$$

so that we finally get

$$\left\|\Phi_{\text{sel}}(\omega) - \Phi_{\text{sel}}(\omega')\right\| \;\leq\; 4\,\|P\|\,\|\omega-\omega'\|\,.$$

Together with the previous calculation, we see that Φ_{sel} is globally Lipschitz, with Lipschitz constant ≤ 4. $\square$

So, we know that the IPL equation (30) defines a unique flow. As before, we have to check what happens with $\mathcal{M}_+(X)$ under the semiflow in forward time. Since $\omega_0 = 0$ trivially implies $\omega_t \equiv 0$ for all $t \geq 0$, we exclude this case from now on. Note that $\omega_0 \neq 0$ results in $\|\omega_t\| > 0$ for all $t \geq 0$, due to uniqueness. Let $\omega \in \mathcal{M}_+(X)$ be a positive measure and E a Borel set such that $\omega(E) = 0$. This implies $\Phi_{\text{sel}}(\omega)\,(E) = (P\omega)\,(E) \geq 0$ because P itself satisfies the positive minimum principle (A2) by assumption. Also, for any $\omega \in \mathcal{M}_+(X)$, we have

$$\Phi_{\text{sel}}(\omega)\,(X) \;=\; P\omega(X) - \frac{P\omega(X)}{\|\omega\|}\,\omega(X) \;=\; 0$$

because $\omega(X) = \|\omega\|$ for positive measures. Together with Fact 6, we see that assumptions (A1) – (A3) are satisfied, and we can invoke Theorem 1.

Proposition 9 *Assume that the linear operator P is bounded and satisfies* (A2). *Then the abstract Cauchy problem of the IPL equation* (30) *with initial condition ω_0 has a unique solution. The cone of positive measures is positive invariant under the flow, and the norm of positive measures is preserved. In particular, $\mathcal{P}(X)$ is positive invariant.* □

REMARK: We would like to mention that the assumption of bounded P is somewhat restricted. For non-compact X, many interesting selection models lead to unbounded P. For mutation and selection alone, the more general situation has been investigated in [20] and, more recently, in [11], in the framework of analytic semigroups, compare [19, Ch. II.4.a]. Our emphasis here is on the basic structure that emerges from the interaction with recombination; this will also carry over to more general cases.

Before we proceed, let us make the following observation.

Fact 7 *If the linear operator P is bounded and satisfies the positive minimum principle, the same is true of $P' = P + c\,\mathbf{1}$ for arbitrary $c \in \mathbb{R}$. Furthermore, the flow of the IPL equation* (30) *on $\mathcal{M}_+(X)$ remains unchanged if P is replaced by P'.*

PROOF: If ν is a positive measure and E a Borel set with $\nu(E) = 0$, then $P'\nu(E) = P\nu(E) + c\,\nu(E) = P\nu(E) \geq 0$ because P satisfies (A2) by assumption. Since P' is still bounded, the IPL equation (30) with P' in place of P conforms to Proposition 9. If $\omega \in \mathcal{M}_+(X)$, we obtain

$$P'\omega - \frac{P'\omega(X)}{\|\omega\|}\,\omega \;=\; P\omega + c\,\omega - \frac{P\omega(X)}{\|\omega\|}\,\omega - \frac{c\,\omega(X)}{\|\omega\|}\,\omega \;=\; P\omega - \frac{P\omega(X)}{\|\omega\|}\,\omega$$

from which the claim follows. □

Once again, although the ODE (30) is nonlinear, it can be solved in closed terms. This time, we employ Thompson's trick [44] through the substitution

$$\eta_t \;=\; \vartheta(t)\,\omega_t \tag{32}$$

where ω_t is a solution of (30). One then obtains

$$\dot{\eta}_t \;=\; \Big(\dot{\vartheta}(t) - \vartheta(t)\frac{P\omega_t(X)}{\|\omega_t\|}\Big)\omega_t + P\eta_t$$

and a significant simplification is reached if the term in brackets vanishes because the remaining ODE is then linear. This is achieved by the choice

$$\vartheta(t) \;=\; \exp\left(\int_0^t \frac{P\omega_\tau(X)}{\|\omega_\tau\|}\,\mathrm{d}\tau\right) \;=\; \exp\left(\frac{1}{\|\omega_0\|}\int_0^t P\omega_\tau(X)\,\mathrm{d}\tau\right) \tag{33}$$

where the second step follows from Proposition 9. Clearly, $\vartheta(t)$ is well defined (whenever $\omega_0 \neq 0$, which is all we need), and we have reduced the Cauchy problem of (30) to that of the simple linear evolution equation

$$\dot{\eta} \;=\; P\eta\,. \tag{34}$$

This ODE defines a uniformly continuous positive semigroup (since P was assumed to be bounded and to satisfy (A2), the positive minimum principle). The solution of (34) will no longer have fixed norm, but one can always get back to ω_t via

$$\omega_t \;=\; \frac{\|\eta_0\|}{\|\eta_t\|}\,\eta_t\,.$$

Note that $\eta_0 = \omega_0$ and $\|\omega_t\| \equiv \|\eta_0\|$.

Let us next consider the function

$$L(t) \;=\; \frac{P\omega_t(X)}{\|\omega_t\|} \tag{35}$$

which is defined on any orbit of the flow of (30). $L(t)$ is of particular interest on orbits of positive measures, where it admits the interpretation as *mean* (or *averaged*) *fitness*. Here, we know $\|\omega_t\| \equiv \|\omega_0\|$ by Proposition 9, so that we obtain

$$\begin{aligned}
\dot{L}(t) \;&=\; \frac{\frac{\mathrm{d}}{\mathrm{d}t}\big(P\omega_t(X)\big)}{\|\omega_t\|} \;=\; \frac{P\dot{\omega}_t(X)}{\|\omega_t\|} \;=\; \frac{1}{\|\omega_t\|}\,P\Big(P\omega_t - \frac{P\omega_t(X)}{\|\omega_t\|}\,\omega_t\Big)(X) \\
&=\; P^2\Big(\frac{\omega_t}{\|\omega_t\|}\Big)(X) - \Big(P\Big(\frac{\omega_t}{\|\omega_t\|}\Big)(X)\Big)^2
\end{aligned}$$

which has the form of a variance. So we can state

Proposition 10 *If, under the assumptions of Proposition* 9*, P satisfies the condition $(P\omega(X))^2 \leq P^2\omega(X)$ on $\mathcal{P}(X)$, the function $L(t)$ of* (35) *is a Lyapunov function for the flow of* (30) *on the positive cone $\mathcal{M}_+(X)$.*

Proof: From the above calculation, it is clear that $\dot{L}(t) \geq 0$ on all orbits in $\mathcal{M}_+(X)$ if P satisfies the inequality $(P\omega(X))^2 \leq P^2\omega(X)$ on $\mathcal{P}(X)$, so L cannot decrease along such an orbit. $\square$

Remark: Our definition of a Lyapunov function on $\mathcal{M}_+(X)$ is *global* and (up to a sign) that of [4, Ch. 18]. Note that the stricter version of [26], where $\dot{L}(t) = 0$ would correspond to a unique equilibrium on $\mathcal{M}_+(X)$, is not so useful here because the asymptotic state (as $t \to \infty$) of the selection equation depends on the initial condition, i.e. there is no unique equilibrium in general. However, one might profit from the use of *local* Lyapunov functions, compare [26, Thm. 1.0.2 (iii)], but we do not expand on this here.

The condition on P can be reformulated by noting that

$$P^2\omega(X) - (P\omega(X))^2 \;=\; (P - c\,\mathbf{1})^2\omega(X)$$

with $c = P\omega(X)$. A sufficient condition for Proposition 10 to hold is then that $(P - c\,\mathbf{1})^2$ is a positive operator for all (or sufficiently many) $c \in \mathbb{R}$. A particularly well studied case of this is when X is finite and P is a diagonal matrix in the canonical basis consisting of the extremal measures of $\mathcal{P}(X)$. In this case, Proposition 10 is known as Fisher's fundamental theorem, see [27] for details. In the more general case, Lyapunov functions may be considered even more important since they determine the 'direction' of the evolution process in a situation where little information is available otherwise, since the solution given by Eq. (36) is not very explicit then.

The results of this section can also be formulated for the (sub-)space $\mathcal{M}^{\otimes}$ of $\mathcal{M}(X)$, if it is invariant under the action of P. In view of the product structure of X, let us now assume that we have $P = \sum_{i=0}^{n} P_i$, with bounded P_i that are locally represented by p_i (as with q_i versus Q_i before). Clearly, P maps $\mathcal{M}^{\otimes}$ into itself. We call this situation additivity across sites, in complete analogy to our previous discussion of mutation. We can then rewrite our solution as

$$\eta_t \;=\; \exp(tP)\,\eta_0 \;=\; \Big(\bigotimes_{i=0}^{n} \exp(tp_i)\Big)\eta_0\,. \tag{36}$$

With some further restrictions on the linear operator P, an analogue of Proposition 10 remains true even in the presence of recombination. This rests on the applicability of Lemma 5. We thus consider the IPL equation

$$\dot{\omega} \;=\; P\omega - \frac{P\omega(X)}{\|\omega\|}\,\omega + \sum_{\alpha\in L} \varrho_\alpha (R_\alpha - \mathbf{1})(\omega) \;=\; \Phi_{\text{sel}}(\omega) + \Phi_{\text{rec}}(\omega) \tag{37}$$

whose Cauchy problem has all the nice properties we need, see Proposition 11 below in the special case $Q = 0$. We now assume:

1. P has complete product structure as a generator, i.e. $P = \sum_{i=0}^{n} P_i$ with $P_i = \mathbf{1} \otimes \cdots \otimes \mathbf{1} \otimes p_i \otimes \mathbf{1} \otimes \cdots \otimes \mathbf{1}$.

2. Each P_i is itself a bounded, strictly positive operator.

If ω_t is a solution of (37), we again define $L(t)$ as in (35) and obtain, by Lemma 5,

$$
\begin{aligned}
\dot{L}(t) &= \frac{1}{\|\omega_t\|} P\Big(P\omega_t - \frac{P\omega_t(X)}{\|\omega_t\|}\,\omega_t + \sum_{\alpha\in L} \varrho_\alpha (R_\alpha - \mathbf{1})(\omega_t)\Big)(X) \\
&= \left[P^2\Big(\frac{\omega_t}{\|\omega_t\|}\Big)(X) - \Big(P\Big(\frac{\omega_t}{\|\omega_t\|}\Big)(X)\Big)^2\right] + \frac{\sum_{i=0}^n \Phi_{\mathrm{rec}}(P_i\,\omega_t)(X)}{\|\omega_t\|}.
\end{aligned}
$$

The last term vanishes due to our general assumptions because $P_i\,\omega_t > 0$ and then $\Phi_{\mathrm{rec}}(P_i\,\omega_t)(X) = 0$ due to (A3). So, we are back to the condition already encountered above. To summarize:

Theorem 5 *Let $P = \sum_{i=0}^n P_i$ satisfy the assumptions of Proposition 9, and let each P_i be a bounded, strictly positive operator with complete product structure. If P also satisfies the condition $(P\omega(X))^2 \leq P^2\omega(X)$ on $\mathcal{P}^{\otimes}$, the function $L(t)$ of (35) is a Lyapunov function for the flow of (37) on the positive cone $\mathcal{M}_+^{\otimes}$.* □

In the absence of recombination, there are other Lyapunov functions known for certain combinations of selection with *mutation*. They rely on the spectral theorem applied to $P + Q$, see [28]. In selection-*recombination* equations where P violates the product structure, the mean fitness $L(t)$ need no longer be a Lyapunov function. Moreover, the possibility of periodic solutions [3] demonstrates that, in more general (diploid) models (e.g. with dominance), no meaningful Lyapunov function is to be expected.

6 All three

In this last step, we combine all three processes, with the general assumptions as before. In view of the inherent product structure, we only consider the dynamics on the Banach space $\mathcal{M}^{\otimes}$. The IPL equation now reads

$$
\begin{aligned}
\dot{\omega} &= \Phi_{\mathrm{mut}}(\omega) + \Phi_{\mathrm{rec}}(\omega) + \Phi_{\mathrm{sel}}(\omega) \qquad (38)\\
&= (Q + P)\omega - \frac{P\omega(X)}{\|\omega\|}\,\omega + \sum_{\alpha\in L} \varrho_\alpha \big(R_\alpha - \mathbf{1}\big)(\omega)
\end{aligned}
$$

and we immediately get the following result, again from Theorem 1, and Lemma 4.

Proposition 11 *Let Q be a bounded Markov generator and P a bounded generator of a positive semigroup, both of product form. Let R_α be the recombinators of Eq.* (9)*. Then, the abstract Cauchy problem of the IPL equation* (38) *has a unique solution. The cone $\mathcal{M}_+^\otimes$ is positive invariant and the norm of positive measures is preserved under the forward flow.* $\square$

Remark: Since Q is a Markov generator, we know from Section 2 that $Q\omega(X) = 0$ for all $\omega \in \mathcal{M}_+^\otimes$, and we could also start from an IPL equation where Q is absorbed into P — it would give the same flow on $\mathcal{M}_+^\otimes$. We retain the separation into mutation and selection because, in more general situations, it is often adequate from both the biological and the mathematical point of view (for example, the mutation operator is usually bounded, but the selection operator may be unbounded); for review, see [12, Ch. IV]. We will also combine Q and P, but only after Thompson's linearization transformation.

Let ω_t, $t \geq 0$, be the solution for initial condition ω_0. Define η_t as above in (32), with $\vartheta(t)$ of (33). Then, ω_t is a solution of (38) if and only if η_t solves the reduced IPL equation

$$\dot{\eta} \;=\; S\eta + \Phi_{\mathrm{rec}}(\eta) \tag{39}$$

where $S = Q + P$ is the bounded generator of a uniformly continuous semigroup of positive operators. Note that the right hand side of (39) still satisfies assumptions (A1) and (A2), but no longer (A3). So, the corresponding Cauchy problem still has a unique solution, with $\mathcal{M}_+^\otimes$ being positive invariant, but the norm of positive measures need no longer be preserved under the flow in forward time — and this is precisely the point of this exercise!

From now on, we generally assume that both mutation and selection are adapted to the special product form of our state space X, so $S = \sum_{i=0}^n S_i$ (with corresponding local operator s_i). Hence, $\exp(tS)$ is again a tensor product of local operators.

Lemma 6 *If $S = \sum_{i=0}^n S_i$ is the bounded generator of a uniformly continuous semigroup of positive operators, then we have $\exp(tS)R_\alpha = R_\alpha \exp(tS)$ on $\mathcal{M}_+^\otimes$, for all $t \geq 0$ and $\alpha \in L$.*

Proof: Fix $t \geq 0$ and set $W = \exp(tS)$. This is a positive operator by assumption. Also, since S is bounded, $\nu > 0$ implies $\exp(tS)\nu > 0$ and W is strictly positive. The result then follows from Lemma 5. $\square$

This result means that we can use all our above methods again and construct immediately the solution of (39). At this point, we particularly profit from our approach in that we can still solve the case with (additive) selection. In the context of Haldane linearization, any form of selection has, so far, appeared as a major obstacle, due to the fact that the flow induced by P fails to preserve the norm of positive measures [40].

Theorem 6 *If* $S = \sum_{i=0}^{n} S_i$ *satisfies the assumptions of Lemma* 6*, the solution of the reduced IPL equation* (39)*, with initial condition* $\eta_0 \in \mathcal{M}_+^{\otimes}$*, is given by*

$$\eta_t \;=\; \exp(tS) \sum_{G \subset L} a_G(t)\, R_G(\eta_0)$$

with the coefficients $a_G(t)$ *of* (13)*. The solution of the abstract Cauchy problem for the original IPL equation* (38) *emerges from here via*

$$\omega_t \;=\; \frac{\|\omega_0\|}{\|\eta_t\|}\, \eta_t$$

where $\omega_0 = \eta_0$*. If* $\omega_0 \in \mathcal{P}(X)$*, then* $\{\omega_t \mid t \geq 0\}$ *is a one-parameter family of probability measures.* $\square$

In line with our previous reasoning, we can determine the asymptotic behaviour,

$$\eta_t \;\sim\; \bigotimes_{i=0}^{n} \big(\exp(ts_i)(\pi_i.\eta_0)\big),$$

where we have used the product structure of $\exp(tS)$ and the fact that all neglected terms, as $t \to \infty$, are exponentially small in comparison. The meaning for ω_t is, once again, that stationary measures are complete product measures, and the properties of the linear operators s_i determine whether there is a unique global equilibrium measure. This is connected to the general Perron-Frobenius theory of positive operators which is rather involved in general, see [42, Ch. V.5] and [25]. If, however, X is finite (so that $\mathcal{M}^{\otimes}$ is finite-dimensional, and $\mathcal{M}^{\otimes} = \mathcal{M}(X)$) and all s_i are irreducible, there are unique $\nu_i \in \mathcal{P}(X_i)$ so that $\exp(ts_i)\nu_i = \exp(t\lambda_i)\nu_i$ with $\lambda_i \in \mathbb{R}$ being the largest eigenvalue of s_i. In this case, as a simple calculation shows, we obtain

$$\omega_t \;\longrightarrow\; \nu_0 \otimes \cdots \otimes \nu_n$$

in the $\|.\|$-topology, as $t \to \infty$, for any initial condition $\omega_0 \in \mathcal{P}(X)$.

Also, the following observation results immediately from Theorem 6.

Corollary 3 *If an initial condition $\omega_0 \in \mathcal{P}^{\otimes}$ is a product measure at link $\alpha \in L$, this is also true of the corresponding solution ω_t of* (38)*, for all $t \geq 0$. In particular, if ω_0 is a complete product measure, this remains the case under the forward flow, i.e. for all ω_t with $t \geq 0$.* $\square$

Let us return to the general discussion. The remainder is then a copy of what we did in Section 4, with Q replaced by S. In particular, we get

$$\eta_t \;=\; \sum_{K \subset L} b_K(t)\, T_K\big(\exp(tS)\eta_0\big)$$

from which one can, once again, determine the linkage disequilibria. Note, however, that the meaning has changed now, because the norm of η_t varies with time. In particular, one has to consider the quotient $\eta_t/\|\eta_t\|$, rather than η_t alone, to extract the correct behaviour for the linkage disequilibria $F_G^t(j_1, \ldots, j_k) = T_G(\omega_t)\big(\langle j_1, \ldots, j_k\rangle\big)$. To be concrete, observe first that

$$\frac{d}{dt} T_G(\eta_t) \;=\; \Big(S - \sum_{\alpha \in \overline{G}} \varrho_\alpha\Big)\, T_G(\eta_t)$$

in perfect analogy with (29). Since T_G is positive homogeneous of degree one (Eq. (22)), and $\|\eta_t\| = \eta_t(X)$ for positive measures, one obtains

$$\frac{d}{dt} T_G(\omega_t) \;=\; \Big(S - \frac{S\omega_t(X)}{\|\omega_t\|}\,\mathbf{1} - \sum_{\alpha \in \overline{G}} \varrho_\alpha\Big) T_G(\omega_t)\,.$$

Clearly, knowledge of the mean fitness, $S\omega_t(X)/\|\omega_t\|$, is now required to determine the dynamics of the linkage disequilibria.

7 Afterthoughts

In this article, we have constructed an explicit solution of the single-crossover recombination model in continuous time, with mutation and additive selection. It is quite astonishing that such a solution should be possible at all – after all, it is an explicit representation of a nonlinear semigroup. However, it is no coincidence that this works in continuous time, rather than in discrete time. Let us discuss this for recombination alone. The discrete-time analogue of our single-crossover model is the so-called model with *complete interference* [13]:

$$\omega_{n+1} \;=\; \sum_{\alpha\in L} \varrho_\alpha R_\alpha(\omega_n) + \Big(1 - \sum_{\alpha\in L} \varrho_\alpha\Big)\,\omega_n \,. \tag{40}$$

Similar as this may look to its continuous-time relative, the probabilistic structure is quite different. Single crossovers in continuous time imply *independence of links*, as expressed in the coefficient functions (13) and the resulting factorization property (Lemma 2). In contrast, a second crossover is inhibited for the duration of an entire generation in discrete time, due to interference of crossovers with each other (hence the name); see also [37]. As a result, independence is lost, which makes the discrete model inherently more difficult.

Of course, this also applies to the situation with selection. Models of recombination and selection based on independent sites and finite site spaces have been thoroughly investigated in the population genetics literature, see [22, 23, 30, 31, 34, 36] for some key references and [12, 14] for recent comprehensive reviews. Independence of sites with respect to selection is reflected by a tensor product structure of P, may be interpreted as lack of interaction between genes, and is known as absence of *epistasis* in genetics. More precisely, since the dynamical systems mostly considered so far were in discrete time, a comparison with our setting is more adequate at the level of the semigroup, rather than that of the generator.

Two notions of independence have been used, compare [30, 34, 12], which would translate into our setting as either $\exp(P) = \prod_i \exp(P_i) = \bigotimes_i \exp(p_i)$ ('multiplicative fitness') or as $\exp(P)$ replaced by $\sum_i \exp(P_i)$ ('additive fitness'). Previously, much emphasis has been on the effects of dominance (i.e. the interaction between the two alleles joined in a diploid genotype). This may lead to multiple equilibria, which need not all be of product type, and astonishing differences in the qualitative behaviour of the multiplicative and additive scenario are observed, as reviewed in [30, 12]. However, these effects are absent if there is no dominance (as in our model); in particular, all equilibria are then of product type. Thus, our simple continuous-time model might well serve as an exactly solved reference case which also captures the qualitative features of the corresponding models in discrete time, although no explicit solution is available there.

Now, the logical next step would be to extend the analysis to the inclusion of interactions between sites, which occur as soon as selection is no longer additive across sites. Alas, this is much more involved, and even the simplest cases go far beyond what we have outlined above. The reason is that selection now forces the introduction of further terms in the right hand side of the IPL equation so that the corresponding semigroups no longer commute with recombination. Nevertheless, several situations can be envisioned that admit at least a perturbative approach. In line with the single-crossover assumption, an expansion for small recombination rates would be appropriate, in contrast to the well-known quasi-linkage-equilibrium approach for large recombination rates (for review, see [12]). We hope to report on some progress in this direction soon.

Appendix: Moments versus correlations

As mentioned above, it is often desirable to separate effects that stem from mutual interactions of differently many "particles" or, as in the above discussion, from specification at a different number of sites. For two sites, correlation C and moments F are related by $C(\{i,j\}) = F(\{i,j\}) - F(\{i\})F(\{j\})$, where the arguments are meant as symbolic labels. Since this is a rather general structure, we briefly describe its systematic treatment by means of Möbius inversion, also known as inclusion-exclusion principle.

Let $S = \{1, 2, \ldots, k\}$ be a finite set which will serve as the index set of the particles or the specified sites, the latter through $\langle j_1, \ldots, j_\ell \rangle$. Let $\mathcal{A} = \{A_1, \ldots, A_p\}$ be a *partition* of S, i.e. S is the disjoint union of the members of $\mathcal{A}$. Unlike before, the partition need not be ordered. Let the partition $\mathcal{B} = \{B_1, \ldots, B_q\}$ be a *refinement* of $\mathcal{A}$, so that

$$A_1 = B_{j_{1,1}} \cup \cdots \cup B_{j_{1,n_1}} \, , \; \ldots \; , \, A_p = B_{j_{p,1}} \cup \cdots \cup B_{j_{p,n_p}}$$

where $\{\{j_{1,1}, \ldots, j_{1,n_1}\}, \ldots, \{j_{p,1}, \ldots, j_{p,n_p}\}\}$ is a partition of $\{1, \ldots, q\}$, hence $n_1 + \ldots + n_p = q$. We write $\mathcal{B} \preccurlyeq \mathcal{A}$ in this case, where $\preccurlyeq$ defines a partial order which makes S into a poset. The corresponding Möbius function, compare [10, p. 86], is given by

$$\begin{aligned} \mu(\mathcal{B}, \mathcal{A}) &= \prod_{i=1}^{p} (-1)^{n_i - 1} (n_i - 1)! \qquad (41) \\ &= (-1)^{p + n_1 + \ldots + n_p} \, (n_1 - 1)! \cdot \ldots \cdot (n_p - 1)! \end{aligned}$$

If $\mathcal{C}$ is any refinement of $\mathcal{A}$, μ satisfies the formula

$$\sum_{\mathcal{C}\preccurlyeq\mathcal{B}\preccurlyeq\mathcal{A}} \mu(\mathcal{B},\mathcal{A}) \;=\; \begin{cases} 1 & \text{if } \mathcal{A}=\mathcal{C} \\ 0 & \text{otherwise.} \end{cases}$$

Let us now, for a partition $\mathcal{A} = \{A_1,\ldots,A_p\}$, introduce the function $F(\mathcal{A}) = F(A_1)\cdot\ldots\cdot F(A_p)$, and similarly for the correlations, C. These quantities are related by

$$F(\mathcal{A}) \;:=\; \sum_{\mathcal{B}\preccurlyeq\mathcal{A}} C(\mathcal{B}) \;=\; \sum_{\mathcal{B}\preccurlyeq\mathcal{A}} \prod_{B\in\mathcal{B}} C(B) \tag{42}$$

because this precisely reflects the idea to separate off contributions from subsets of different cardinality. The Möbius inversion formula then gives the following formula for the special case that $\mathcal{A} = \{A\}$:

$$C(A) \;=\; \sum_{\mathcal{B}\preccurlyeq\mathcal{A}} F(\mathcal{B})\,\mu(\mathcal{B},\mathcal{A}) \;=\; \sum_{\mathcal{B}\preccurlyeq\mathcal{A}} (-1)^{|\mathcal{B}|-1}\big(|\mathcal{B}|-1\big)! \prod_{i=1}^{|\mathcal{B}|} F(B_i) \tag{43}$$

where $|\mathcal{B}|$ denotes the number of sets in the partition $\mathcal{B} = \{B_1,\ldots,B_{|\mathcal{B}|}\}$. The following example might illustrate this formula:

$$\begin{aligned} C(\{1,2,3\}) \;&=\; F(\{1,2,3\}) \;+\; 2\,F(\{1\})F(\{2\})F(\{3\}) \\ &\quad - F(\{1\})F(\{2,3\}) - F(\{2\})F(\{1,3\}) - F(\{3\})F(\{1,2\}) \end{aligned}$$

which is to be compared with

$$\begin{aligned} F(\{1,2,3\}) \;&=\; C(\{1,2,3\}) \;+\; C(\{1\})C(\{2\})C(\{3\}) \\ &\quad + C(\{1\})C(\{2,3\}) + C(\{2\})C(\{1,3\}) + C(\{3\})C(\{1,2\}) \end{aligned}$$

according to (42). Let us finally remark that formula (43) can be applied factorwise if $\mathcal{A} = \{A_1,\ldots,A_p\}$ because then $C(\mathcal{A}) = C(A_1)\cdot\ldots\cdot C(A_p)$ by definition.

Acknowledgements

It is our pleasure to thank Reinhard Bürger, Hans-Otto Georgii, Joachim Hermisson, Achim Klenke and Manfred Wolff for a number of clarifying discussions, and Ulrich Hermisson and Oliver Redner for carefully reading the manuscript. M.B. would like to thank Robert V. Moody and the Department of Mathematical Sciences of the University of Alberta (Edmonton, Canada) for hospitality, where part of this work was done.

References

[1] M. Aigner, *Combinatorial Theory*, Springer, Berlin (1979); reprint (1997).

[2] E. Akin, *The Geometry of Population Genetics*, LNB 31, Springer, Berlin (1979).

[3] E. Akin, *Cycling in simple genetic systems*, J. Math. Biol. **13** (1982) 305–324.

[4] H. Amann, *Gewöhnliche Differentialgleichungen*, 2nd ed., de Gruyter, Berlin (1995); (older) English ed.: *Ordinary Differential Equations*, de Gruyter, Berlin (1990).

[5] W. Arendt, *Characterization of positive semigroups on Banach lattices*, in: *One-parameter Semigroups of Positive Operators*, ed. R. Nagel, LNM 1184, Springer, Berlin (1986), pp. 247–291.

[6] E. Baake, *Mutation and recombination with tight linkage*, J. Math. Biol. **42** (2001) 455–488.

[7] E. Baake and W. Gabriel, *Biological evolution through mutation, selection and drift: An introductory review*, in: *Annual Review of Computational Physics*, vol. VII, ed. D. Stauffer, World Scientific, Singapore (2000), pp. 203–264; cond-mat/9907372.

[8] G. Ben Arous and O. Zeitouni, *Increasing propagation of chaos for mean field models*, Ann. Inst. H. Poincaré – Prob. Stat. **35** (1999) 85–102.

[9] S. K. Berberian, *Measure and Integration*, Macmillan, New York (1965).

[10] C. Berge, *Principles of Combinatorics*, Academic Press, New York (1971).

[11] R. Bürger and I. Bomze, *Stationary distributions under mutation-selection balance: structure and properties*, Adv. Appl. Prob. **28** (1996) 227–251.

[12] R. Bürger, *The Mathematical Theory of Selection, Recombination, and Mutation*, Wiley, Chichester (2000).

[13] F. B. Christiansen, *The effect of population subdivision on multiple loci without selection*, in: *Mathematical evolutionary theory*, ed. M. W. Feldman, Princeton University Press, Princeton (1989), pp. 71–85.

[14] F. B. Christiansen, *Population Genetics of Multiple Loci*, Wiley, Chichester (2000).

[15] A. Clark et al., *Haplotype structure and population genetic inferences from nucleotide-sequence variation in human lipoprotein lipase*, Am. J. Hum. Gen. **63** (1998), 595–612.

[16] K. J. Dawson, *The decay of linkage disequilibria under random union of gametes: how to calculate Bennett's principal components*, Theor. Pop. Biol. **58** (2000) 1–20.

[17] K. J. Dawson, *The evolution of a population under recombination: How to linearise the dynamics*, Lin. Alg. Appl. (2002), in press.

[18] R. M. Dudley, *Real Analysis and Probability*, Chapman and Hall, New York (1989).

[19] K.-J. Engel and R. Nagel, *One-Parameter Semigroups for Linear Evolution Equations*, GTM 194, Springer, New York (2000).

[20] I. Eshel, *Evolution processes with continuity of types*, Adv. Appl. Prob. **4** (1972) 475–507.

[21] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, Wiley, New York (1986).

[22] W. J. Ewens, *A generalized fundamental theorem of natural selection*, Genetics **63** (1969) 531–537.

[23] W. J. Ewens, *Mean fitness increases when fitnesses are additive*, Nature **221** (1969) 1076.

[24] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, 2nd ed., Springer, New York (1998).

[25] G. Greiner, *Spectral theory of positive semigroups on Banach lattices*, in: *One-parameter Semigroups of Positive Operators*, ed. R. Nagel, LNM 1184, Springer, Berlin (1986), pp. 292–332.

[26] J. Guckenheimer and Ph. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, corr. 3rd printing, Springer, New York (1990).

[27] J. Hofbauer, *The selection-mutation equation*, J. Math. Biol. **23** (1985) 41–53.

[28] B. L. Jones, *Some principles governing selection in self-reproducing macromolecular systems – an analog of Fisher's fundamental theorem*, J. Math. Biol. **6** (1978) 169–175.

[29] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, North-Holland, Amsterdam (1981).

[30] S. Karlin, *General two-locus selection models: Some objectives, results and interpretation*, Theor. Pop. Biol. **7** (1975) 364–398.

[31] S. Karlin and U. Liberman, *Central equilibria in multilocus systems. I. Generalized nonepistatic selection regimes*, Genetics **91** (1979) 777–798.

[32] M. Kimura, *A stochastic model concerning the maintenance of genetic variability in quantitative characters*, Proc. Natl. Acad. Sci. **54** (1965) 731–736.

[33] J. F. C. Kingman, *Markov population processes*, J. Appl. Prob. **6** (1969) 1–18.

[34] V. Kirzhner and Yu. Lyubich, *Multilocus dynamics under haploid selection*, J. Math. Biol. **35** (1997) 391–408.

[35] S. Lang, *Real and Functional Analysis*, 3rd ed., Springer, New York (1993).

[36] Yu. I. Lyubich, *Mathematical Structures in Population Genetics*, Springer, Berlin (1992).

[37] H. Manos and U. Liberman, *Discrete chiasma formation models and their associated high order interference*, J. Math. Biol. **36** (1998) 448–468.

[38] D. McHale and G. A. Ringwood, *Haldane linearisation of baric algebras*, J. London Math. Soc. (2) **28** (1983) 17–26.

[39] M. Reed and B. Simon, *Functional Analysis*, 2nd ed., Academic Press, San Diego, CA (1980).

[40] G. A. Ringwood, *Hypergeometric algebras and Mendelian genetics*, Niew Archief voor Wiskunde (4) **3** (1985) 69–83.

[41] W. Rudin, *Real and Complex Analysis*, 3rd ed., McGraw-Hill, New York (1987).

[42] H. H. Schaefer, *Banach Lattices and Positive Operators*, Springer, Berlin (1974).

[43] S. Schaeffer and E. L. Miller, *Estimates of linkage disequilibrium and the recombination parameter determined from segregating nucleotide sites in the alcohol dehydrogenase region of Drosophila pseudoobscura*, Genetics **135** (1993), 541–552.

[44] C. J. Thompson and J. L. McBride, *On Eigen's theory of the self-organization of matter and the evolution of biological macromolecules*, Math. Biosci. **21** (1974) 127–142.

[45] D. Werner, *Funktionalanalysis*, 3rd ed., Springer, Berlin (2000).