Oliver Redner · Michael Baake

Unequal Crossover Dynamics in Discrete and Continuous Time

Abstract. We analyze a class of models for unequal crossover (UC) of sequences containing sections with repeated units that may differ in length. In these, the probability of an 'imperfect' alignment, in which the shorter sequence has d units without a partner in the longer one, scales like q^d as compared to 'perfect' alignments where all these copies are paired. The class is parameterized by this penalty factor q. An effectively infinite population size and thus deterministic dynamics is assumed. For the extreme cases $q = 0$ and $q = 1$, and any initial distribution whose moments satisfy certain conditions, we prove the convergence to one of the known fixed points, uniquely determined by the mean copy number, in both discrete and continuous time. For the intermediate parameter values, the existence of fixed points is shown.

1. Introduction

Recombination is a by-product of (sexual) reproduction that leads to the mixing of parental genes by exchanging genes (or sequence parts) between homologous chromosomes (or DNA strands). This is achieved through an alignment of the corresponding sequences, along with crossover events which lead to reciprocal exchange of the induced segments. In this process, imperfect alignment may result in sequences that differ in length form the parental ones; this is known as *unequal crossover* (UC). Imperfect alignment is facilitated by the presence of repeated elements (as is observed in some rDNA sequences, compare [5]), and is believed to be an important driving mechanism for their evolution. The repeated elements may follow an evolutionary course independent of each other and thus give rise to evolutionary innovation. For a detailed discussion of these topics, see [22,23] and references therein.

This article is concerned with a class of models for UC, originally investigated by Shpak and Atteson [22] for discrete time, which is built on preceding work

Institut für Mathematik, Univ. Greifswald, Jahnstr. 15a, 17487 Greifswald, Germany
Fakultät für Mathematik. Univ. Bielefeld, Pf. 100131, 33501 Bielefeld, Germany
e-mail: redner@uni-greifswald.de
e-mail: mbaake@mathematik.uni-bielefeld.de

Key words: Unequal crossover – Recombination – Quadratic operators – Probability generating functions – Lyapunov functions

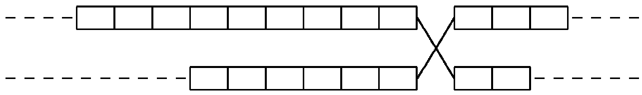

Fig. 1. Snapshot of an unequal crossover event as described in the text. Rectangles denote the relevant blocks, while the dashed lines indicate possible extensions with other elements that are disregarded here.

by Ohta [15] and Walsh [25] (see [22] for further references). Starting from their partly heuristic results, we prove various existence and uniqueness theorems and analyze the convergence properties, both in discrete and in continuous time. This will require a rather careful mathematical development because the dynamical systems are infinite dimensional.

In this model class, one considers individuals whose genetic sequences contain a section with repeated units. These may vary in number, $i \in \mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$, where $i = 0$ is explicitly allowed, corresponding to no unit being present (yet). The composition of these sections (with respect to mutations that might have occurred) and the rest of the sequence are ignored here.

In the course of time, recombination events happen, each of which basically consists of three steps. First, independent pairs are formed at random (in equidistant time steps, or at a fixed rate). Then, their respective sections are randomly aligned, possibly with imperfections in form of 'overhangs', according to some probability distribution for the various possibilities. Finally, both sequences are cut at an arbitrary common position between two adjacent building blocks, with uniform distribution for the cut positions, and their right (or left) fragments are interchanged. This so-called unequal crossover is schematically depicted in Figure 1. Obviously, the total number of relevant units is conserved in each event.

We assume the population size to be (effectively) infinite. (Concerning finite populations, see the remarks in Section 7.) Then, almost surely in the probabilistic sense, compare [4, Sec. 11.2], the population is described by the deterministic time evolution of a probability measure $\boldsymbol{p} \in \mathcal{M}_1^+(\mathbb{N}_0)$, which we identify with an element $\boldsymbol{p} = (p_k)_{k \in \mathbb{N}_0}$ in the appropriate subset of $\ell^1(\mathbb{N}_0)$. Since we will not consider any genotype space other than $\mathbb{N}_0$ in this article, reference to it will be omitted in what follows. These spaces are complete in the metric derived from the usual ℓ^1 norm, which is the same as the total variation norm here. The metric is denoted by

$$d(\boldsymbol{p}, \boldsymbol{q}) = \|\boldsymbol{p} - \boldsymbol{q}\|_1 = \sum_{k \geqslant 0} |p_k - q_k| . \tag{1}$$

With this notation, the above process is described by the recombinator

$$\mathcal{R}(\boldsymbol{p})_i = \frac{1}{\|\boldsymbol{p}\|_1} \sum_{j,k,\ell \geqslant 0} T_{ij,k\ell}\, p_k\, p_\ell . \tag{2}$$

Here, $T_{ij,k\ell} \geqslant 0$ denotes the probability that, given a pair (k, ℓ), this pair turns into (i, j). Consequently, for normalization, we require

$$\sum_{i,j \geqslant 0} T_{ij,k\ell} = 1 \qquad \text{for all } k, \ell \in \mathbb{N}_0. \tag{3}$$

The factor $p_k \, p_\ell$ in (2) describes the probability that a pair (k, ℓ) is formed, i.e., we assume that two individuals are chosen independently from the population. We assume further that $T_{ij,k\ell} = T_{ji,k\ell} = T_{ij,\ell k}$, i.e., that $T_{ij,k\ell}$ is symmetric with respect to both index pairs, which is reasonable. Then, the summation over j in (2) represents the breaking-up of the pairs after the recombination event. These two ingredients of the dynamics constitute what is known as (*instant*) *mixing* and are responsible for the quadratic nature of the iteration process.

As mentioned above, we will only consider processes that conserve the total copy number in each event, i.e., $T^{(q)}_{ij,k\ell} > 0$ for $i + j = k + \ell$ only. Together with the normalization (3) and the symmetry condition from above, this yields the (otherwise weaker) condition

$$\sum_{i \geqslant 0} i \, T_{ij,k\ell} = \sum_{i,j \geqslant 0} \frac{i+j}{2} \, T_{ij,k\ell} = \frac{k+\ell}{2}, \tag{4}$$

which implies conservation of the *mean copy number* in the population,

$$\sum_{i \geqslant 0} i \, \mathcal{R}(\boldsymbol{p})_i = \sum_{i,j,k,\ell \geqslant 0} i \, T^{(q)}_{ij,k\ell} \, p_k \, p_\ell = \sum_{k,\ell \geqslant 0} \frac{k+\ell}{2} \, p_k \, p_\ell = \sum_{k \geqslant 0} k \, p_k .$$

Condition (3) and the presence of the prefactor $1/\|\boldsymbol{p}\|_1$ in (2) make $\mathcal{R}$ norm non-increasing, i.e., $\|\mathcal{R}(\boldsymbol{x})\|_1 \leqslant \|\boldsymbol{x}\|_1$, and positive homogeneous of degree 1, i.e., $\mathcal{R}(a\boldsymbol{x}) = |a| \mathcal{R}(\boldsymbol{x})$, for all $\boldsymbol{x} \in \ell^1$ and $a \in \mathbb{R}$. Furthermore, $\mathcal{R}$ is a positive operator with $\|\mathcal{R}(\boldsymbol{x})\|_1 = \|\boldsymbol{x}\|_1$ for all positive elements $\boldsymbol{x} \in \ell^1$. Thus, it is guaranteed that $\mathcal{R}$ maps $\mathcal{M}^+_r$, the space of positive measures of total mass r, into itself. This space is complete in the topology induced by the norm $\|.\|_1$, i.e., by the metric d from (1). (For $r = 1$, of course, the prefactor is redundant but ensures numerical stability of an iteration with $\mathcal{R}$.)

Given an initial configuration $\boldsymbol{p}_0 = \boldsymbol{p}(t = 0)$, the dynamics may be taken in discrete time steps, with subsequent generations,

$$\boldsymbol{p}(t+1) = \mathcal{R}(\boldsymbol{p}(t)), \qquad t \in \mathbb{N}_0 . \tag{5}$$

Our treatment of this case will be set up in a way that also allows for a generalization of the results to the analogous process in continuous time, where generations are overlapping,

$$\tfrac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{p}(t) = \varrho \, (\mathcal{R} - \mathbb{1})(\boldsymbol{p}(t)), \qquad t \in \mathbb{R}_{\geqslant 0} . \tag{6}$$

Obviously, the (positive) parameter ϱ in (6) only leads to a rescaling of the time t. We therefore choose $\varrho = 1$ without loss of generality. Furthermore, the formulation of discrete versus continuous time dynamics in (5) and (6) is chosen so that the fixed points of (5) are identical to the equilibria of (6), regardless of ϱ. This can easily be verified by a direct calculation. In what follows, we will thus use the term *fixed point* for both discrete and continuous dynamics.

In the UC model, one distinguishes 'perfect' alignments, in which each unit in the shorter sequence has a partner in the longer sequence, and 'imperfect' alignments, with 'overhangs' of the shorter sequence relative to the longer one. To come to a reasonable probability distribution for the various possiblities, the first are taken to be equally probable among each other, while the latter are penalized by a factor q^d relative to the first, where $q \in [0,1]$ is a model parameter and d is the length of the overhang (at most the entire length of the shorter sequence; in the example of Figure 1, interpreted as a snapshot right after the crossover event took place, we have $d = 1$). In the extreme case $q = 0$, only perfect alignments may occur, whereas for $q = 1$ overhangs are not penalized at all and one obtains the uniform distribution on the possibilities. For obvious reasons, the first case is dubbed *internal UC*, the second *random UC* [22].

It is now straightforward, though a bit tedious, to derive the transition probabilities $T^{(q)}_{ij,k\ell}$. To this end, one has to trace what happens in steps two and three of the recombination event only, while the random formation of pairs does not enter here. This has been done in [22] and need not be repeated. However, in view of our above remarks, it is desirable to rewrite the findings in a way that reflects the natural symmetry properties of the $T^{(q)}_{ij,k\ell}$. In compact notation, this leads to the transition probabilities

$$T^{(q)}_{ij,k\ell} = C^{(q)}_{k\ell}\, \delta_{i+j,k+\ell}\, (1 + \min\{k,\ell,i,j\})\, q^{0 \vee (k\wedge\ell - i\wedge j)}\,, \tag{7}$$

where $k \vee \ell := \max\{k,\ell\}$, $k \wedge \ell := \min\{k,\ell\}$, and $0^0 = 1$. The $C^{(q)}_{k\ell}$ are chosen such that (3) holds, i.e., $\sum_{i,j\geqslant 0} T^{(q)}_{ij,k\ell} = 1$, and are hence symmetric in k and ℓ. Explicitly, they read (see also [22, Sec. 2.1])

$$C^{(q)}_{k\ell} = \frac{(1-q)^2}{(k\wedge\ell+1)(|k-\ell|+1)(1-q)^2 + 2q(k\wedge\ell - (k\wedge\ell+1)q + q^{k\wedge\ell+1})}\,.$$

Note further that the total number of units is indeed conserved in each event and that the process is symmetric within both pairs. Hence (4) is satisfied.

Let us briefly come back to the question of 'discrete' versus 'continuous' time, which are considered simultaneously for good reasons. Common to both is the nonlinearity that stems from the probability that a certain (random) pair is formed in the first place. Then, for the discrete time dynamics (5), the $T^{(q)}_{ij,k\ell}$ have the direct meaning of the transition probability that, given a pair (k,ℓ), this turns into a pair (i,j). In contrast, for the continuous time dynamics (6), the number $T^{(q)}_{ij,k\ell}$ is to be understood as the probability to obtain a pair (i,j) *conditioned* on a recombination event with a pair of type (k,ℓ), of which each recombines at the same rate. In probabilistic terminology, the $\mathcal{R}$ of (5) is the discrete time skeleton of the process in continuous time, also called the embedded discrete time process.

The aim of this article is to find answers to the following questions:

1. Are there fixed points of the dynamics?

2. Given the mean copy number m, is there a unique fixed point?
3. If so, under which conditions and in which sense does an initial distribution converge to this fixed point under time evolution?

Of course, the trivial fixed point with $p_0 = 1$ and $p_k = 0$ for $k > 0$ always exists, which we generally exclude from our considerations. But even then, the answer to the first question is positive for general operators of the form (2) that satisfy (3) and some rather natural further condition. This is discussed in Section 2. For the extreme cases $q = 0$ (internal UC) and $q = 1$ (random UC), fixed points are known explicitly for every m and it has been conjectured [22] that, under mild conditions, also questions 2 and 3 can be answered positively for all values of $q \in [0, 1]$. Indeed, for both extreme cases, norm convergence of the population distribution to the fixed points can be shown, which is done in Sections 3 and 5, respectively. Since the dynamical systems involved are infinite dimensional, a careful analysis of compactness properties is needed for rigorous answers. The proofs for $q = 1$ are based on alternative representations of probability measures via generating functions, presented in Section 4. For the intermediate parameter regime, we can only show that there exists a fixed point for every m, but neither its uniqueness nor convergence to it, see Section 6. Some remarks in Section 7 conclude this article.

2. Existence of fixed points

Let us begin by stating the following general fact.

Proposition 1. *If the recombinator $\mathcal{R}$ of* (2) *satisfies* (3)*, then the global Lipschitz condition*

$$\|\mathcal{R}(\boldsymbol{x}) - \mathcal{R}(\boldsymbol{y})\|_1 \leqslant C\|\boldsymbol{x} - \boldsymbol{y}\|_1$$

is satisfied, with constant $C = 3$ on ℓ^1, respectively $C = 2$ if $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{M}_r$.

Proof. Let $\boldsymbol{x}, \boldsymbol{y} \in \ell^1$ be non-zero (otherwise the statement is trivial). Then,

$$\begin{aligned}
\|\mathcal{R}(\boldsymbol{x}) - \mathcal{R}(\boldsymbol{y})\|_1 &= \sum_{i\geqslant 0}\left|\sum_{j,k,\ell\geqslant 0} T_{ij,k\ell}\left(\frac{x_k\, x_\ell}{\|\boldsymbol{x}\|_1} - \frac{y_k\, y_\ell}{\|\boldsymbol{y}\|_1}\right)\right| \\
&\leqslant \sum_{k,\ell\geqslant 0}\left|\frac{x_k\, x_\ell}{\|\boldsymbol{x}\|_1} - \frac{y_k\, y_\ell}{\|\boldsymbol{y}\|_1}\right| \sum_{i,j\geqslant 0} T_{ij,k\ell} \\
&= \sum_{k,\ell\geqslant 0}\left|\frac{x_k\, x_\ell}{\|\boldsymbol{x}\|_1} - \frac{x_k\, y_\ell}{\|\boldsymbol{x}\|_1} + \frac{x_k\, y_\ell}{\|\boldsymbol{x}\|_1} - \frac{y_k\, y_\ell}{\|\boldsymbol{y}\|_1}\right| \\
&\leqslant \sum_{k,\ell\geqslant 0}\left(\frac{|x_k|}{\|\boldsymbol{x}\|_1}|x_\ell - y_\ell| + |y_\ell|\left|\frac{x_k}{\|\boldsymbol{x}\|_1} - \frac{y_k}{\|\boldsymbol{y}\|_1}\right|\right) \\
&= \|\boldsymbol{x} - \boldsymbol{y}\|_1 + \frac{1}{\|\boldsymbol{x}\|_1}\big\|\|\boldsymbol{y}\|_1\boldsymbol{x} - \|\boldsymbol{x}\|_1\boldsymbol{y}\big\|_1 .
\end{aligned}$$

The last term becomes

$$\frac{1}{\|\boldsymbol{x}\|_1}\big\|\|\boldsymbol{y}\|_1\boldsymbol{x}-\|\boldsymbol{x}\|_1\boldsymbol{y}\big\|_1$$
$$=\frac{1}{\|\boldsymbol{x}\|_1}\big\|(\|\boldsymbol{y}\|_1-\|\boldsymbol{x}\|_1)\boldsymbol{x}+\|\boldsymbol{x}\|_1(\boldsymbol{x}-\boldsymbol{y})\big\|_1\leqslant 2\|\boldsymbol{x}-\boldsymbol{y}\|_1,$$

from which $\|\mathcal{R}(\boldsymbol{x})-\mathcal{R}(\boldsymbol{y})\|_1\leqslant 3\|\boldsymbol{x}-\boldsymbol{y}\|_1$ follows for $\boldsymbol{x},\boldsymbol{y}\in\ell^1$. If $\boldsymbol{x},\boldsymbol{y}\in\mathcal{M}_r$, the above calculation simplifies to $\|\mathcal{R}(\boldsymbol{x})-\mathcal{R}(\boldsymbol{y})\|_1\leqslant 2\|\boldsymbol{x}-\boldsymbol{y}\|_1$. □

In continuous time, this is a sufficient condition for the existence and uniqueness of a solution of the initial value problem (6), cf. [1, Thms. 7.6 and 10.3]. Another useful notion in this respect is the following.

Definition 1. [1, Sec. 18] *Let Y be an open subset of a Banach space E and let $f\colon Y\to E$ satisfy a (local) Lipschitz condition. A continuous function L from $X\subset Y$ to $\mathbb{R}$ is called a* Lyapunov function *for the initial value problem*

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x}(t)=f(\boldsymbol{x}(t)),\qquad \boldsymbol{x}(0)=\boldsymbol{x}_0\in X,$$

if the orbital derivative $\dot{L}(\boldsymbol{x}_0):=\liminf_{t\to 0^+}\frac{1}{t}\big(L(\boldsymbol{x}(t))-L(\boldsymbol{x}_0)\big)$ *satisfies*

$$\dot{L}(\boldsymbol{x}_0)\leqslant 0 \tag{8}$$

for all initial conditions $\boldsymbol{x}_0\in X$.

If further $\dot{L}(\boldsymbol{x}_{\mathrm{F}})=0$ for a single fixed point $\boldsymbol{x}_{\mathrm{F}}$ only, then L is called a *strict* Lyapunov function. If a Lyapunov function exists, we have

Theorem 1. [1, Thm. 17.2 and Cor. 18.4] *With the notation of Definition 1, assume that there is a Lyapunov function L, that the set X is closed, and that, for an initial condition $\boldsymbol{x}_0\in X$, the set $\{\boldsymbol{x}(t):t\in\mathbb{R}_{\geqslant 0},\boldsymbol{x}(t)\text{ exists}\}$ is relatively compact in X. Then, $\boldsymbol{x}(t)$ exists for all $t\geqslant 0$ and*

$$\lim_{t\to\infty}\operatorname{dist}(\boldsymbol{x}(t),X_L)=0,$$

where $\operatorname{dist}(\boldsymbol{x},X_L)=\inf_{\boldsymbol{y}\in X_L}\|\boldsymbol{x}-\boldsymbol{y}\|$ and X_L denotes the largest invariant subset of $\{\boldsymbol{x}\in X:\dot{L}(\boldsymbol{x})=0\}$ (in forward and backward time). □

Obviously, if L is a strict Lyapunov function, we have $X_L=\{\boldsymbol{x}_{\mathrm{F}}\}$ and this theorem implies $d(\boldsymbol{x}(t),\boldsymbol{x}_{\mathrm{F}})\to 0$ as $t\to\infty$.

Returning to the original question of the existence of fixed points, we now recall the following facts, compare [3,21] for details.

Proposition 2. [27, Cor. to Thm. V.1.5] *Assume the sequence $\big(\boldsymbol{p}^{(n)}\big)$ in $\mathcal{M}_1^+$ to converge in the weak-$*$ topology (i.e., pointwise, or vaguely) to some $\boldsymbol{p}\in\mathcal{M}_1^+$, i.e.,*

$$\lim_{n\to\infty}p_k^{(n)}=p_k\quad\textit{for all }k\in\mathbb{N}_0,\quad\textit{with }p_k\geqslant 0\quad\textit{and }\textstyle\sum_{k\geqslant 0}p_k=1.$$

Then, it also converges weakly (in the probabilistic sense) and in total variation, i.e., $\lim_{n\to\infty}\|\boldsymbol{p}^{(n)}-\boldsymbol{p}\|_1=0$. □

Proposition 3. *Assume that the recombinator $\mathcal{R}$ from* (2) *satisfies* (3) *and has a convex, weak-$*$ closed invariant set $M \subset \mathcal{M}_1^+$, i.e., $\mathcal{R}(M) \subset M$, that is* tight*, i.e., for every $\varepsilon > 0$ there is an $m \in \mathbb{N}_0$ such that $\sum_{k \geqslant m} p_k < \varepsilon$ for every $\boldsymbol{p} \in M$. Then, $\mathcal{R}$ has a fixed point in M.*

Proof. Prohorov's theorem [21, Thm. III.2.1] states that tightness and relative compactness in the weak-$*$ topology are equivalent (see also [3, Chs. 1.1 and 1.5]). In our case, M is tight and weak-$*$ closed, therefore, due to Proposition 2, norm compact. Furthermore, M is convex and $\mathcal{R}$ is (norm) continuous by Proposition 1. Thus, the claim follows from the Leray–Schauder–Tychonov fixed point theorem [18, Thm. V.19]. □

With respect to the UC model, we will see that such compact invariant subsets indeed exist.

3. Internal unequal crossover

After these preliminaries, let us begin with the case of internal UC with perfect alignment only, i.e., $q = 0$ in (7). This case is the simplest because, in each recombination event, no sequences exceeding the longer of the participating sequences can be formed. Here, on $\mathcal{M}_1^+$, the recombinator (2) simplifies to

$$\mathcal{R}_0(\boldsymbol{p})_i = \sum_{\substack{k,\ell \geqslant 0 \\ k \wedge \ell \leqslant i \leqslant k \vee \ell}} \frac{p_k \, p_\ell}{1 + |k - \ell|} . \tag{9}$$

From now on, we write $\mathcal{R}_q$ rather than $\mathcal{R}$ whenever we look at a recombinator with (fixed) parameter q. It is instructive to generalize the notion of reversibility (or detailed balance, compare [22, (4.1)]).

Definition 2. *We call a probability measure $\boldsymbol{p} \in \mathcal{M}_1^+$* reversible *for a recombinator $\mathcal{R}$ of the form* (2) *if, for all $i, j, k, \ell \geqslant 0$,*

$$T_{ij,k\ell} \, p_k \, p_\ell = T_{k\ell,ij} \, p_i \, p_j . \tag{10}$$

The relevance of this concept is evident from the following property.

Lemma 1. *If $\boldsymbol{p} \in \mathcal{M}_1^+$ is reversible for $\mathcal{R}$, it is also a fixed point of $\mathcal{R}$.*

Proof. Assume $\boldsymbol{p}$ to be reversible. Then, by (3),

$$\mathcal{R}(\boldsymbol{p})_i = \sum_{j,k,\ell \geqslant 0} T_{ij,k\ell} \, p_k \, p_\ell = \sum_{j,k,\ell \geqslant 0} T_{k\ell,ij} \, p_i \, p_j = p_i \sum_{j \geqslant 0} p_j = p_i .$$

□

So, in our search for fixed points, we start by looking for solutions of (10). Since, for $q = 0$, forward and backward transition probabilities are simultaneously non-zero only if $\{i, j\} = \{k, \ell\} \subset \{n, n+1\}$ for some n, the components p_k may only be positive on this small set as well. By the following proposition, this indeed characterizes all fixed points for $q = 0$.

Proposition 4. *A probability measure $\boldsymbol{p} \in \mathcal{M}_1^+$ is a fixed point of $\mathcal{R}_0$ if and only if its mean copy number $m = \sum_{k\geqslant 0} k\, p_k$ is finite, $p_{\lfloor m\rfloor} = \lfloor m\rfloor + 1 - m$, $p_{\lceil m\rceil} = m + 1 - \lceil m\rceil$, and $p_k = 0$ for all other k. This includes the case that m is integer and $p_{\lfloor m\rfloor} = p_{\lceil m\rceil} = p_m = 1$.*

Proof. The ‘if’ part was stated in [22, Sec. 4.1] and follows easily by insertion into (9) or (10). For the ‘only if’ part, let i denote the smallest integer such that $p_i > 0$. Then,

$$\mathcal{R}(\boldsymbol{p})_i = p_i^2 + 2p_i \sum_{\ell\geqslant 1} \frac{p_{i+\ell}}{1+\ell} = p_i \left(p_i + p_{i+1} + \sum_{\ell\geqslant 2} \frac{2}{\ell+1} p_{i+\ell} \right) \leqslant p_i,$$

where the last step follows since $\frac{2}{\ell+1} < 1$ in the last sum, with equality if and only if $p_k = 0$ for all $k \geqslant i+2$. This implies $m < \infty$ and the uniqueness of $\boldsymbol{p}$ (given m) with the non-zero frequencies as claimed. □

It it possible to analyze the case of internal UC on the basis of the compact sets to be introduced below in Section 4. However, as J. Hofbauer pointed out to us [8], it is more natural to start with a larger compact set to be introduced in (11). Our main result in this section is thus

Theorem 2. *Assume that, for the initial condition $\boldsymbol{p}(0)$ and fixed $r > 1$, the r-th moment exists, $\sum_{k\geqslant 0} k^r p_k(0) < \infty$. Then, $m = \sum_{k\geqslant 0} k\, p_k(0)$ is finite and, both in discrete and in continuous time, $\lim_{t\to\infty} \|\boldsymbol{p}(t) - \boldsymbol{p}\|_1 = 0$ with the appropriate fixed point $\boldsymbol{p}$ from Proposition 4.*

The proof relies on the following lemma, which slightly modifies and completes the convergence arguments of [22, Sec. 4.1], puts them on rigorous grounds, and extends them to continuous time.

Lemma 2. *Let $r > 1$ be arbitrary, but fixed. Consider the set of probability measures with fixed mean $m < \infty$ and a centered r-th moment bounded by $C < \infty$,*

$$\mathcal{M}_{1,m,C}^+ = \{\boldsymbol{p} \in \mathcal{M}_1^+ : \sum_{k\geqslant 0} k\, p_k = m,\ M_r(\boldsymbol{p}) \leqslant C\}, \tag{11}$$

equipped with (the metric induced by) the total variation norm, where

$$M_s(\boldsymbol{p}) = \sum_{k\geqslant 0} |k-m|^s\, p_k \tag{12}$$

for $s \in \{1, r\}$. This is a compact and convex space. Both M_1 and M_r satisfy $M_s(\mathcal{R}_0(\boldsymbol{p})) \leqslant M_s(\boldsymbol{p})$, with equality if and only if $\boldsymbol{p}$ is a fixed point of $\mathcal{R}_0$. Furthermore, M_1 is a continuous mapping from $\mathcal{M}_{1,m,C}^+$ to $\mathbb{R}_{\geqslant 0}$ and a Lyapunov function for the dynamics in continuous time.

Proof. Let a sequence $(\boldsymbol{p}^{(n)}) \subset \mathcal{M}^+_{1,m,C}$ be given and consider the random variables $\boldsymbol{f}^{(n)} = (k)_{k\in\mathbb{N}_0}$ on the probability spaces $(\mathbb{N}_0, \boldsymbol{p}^{(n)})$. Their expectation values are equal to m, which, by Markov's inequality [21, p. 599], implies the tightness of the sequence $(\boldsymbol{p}^{(n)})$. Hence, by Prohorov's theorem [21, Thm. III.2.1] (see also [3, Chs. 1.1 and 1.5]), it contains a convergent subsequence $(\boldsymbol{p}^{(n_i)})$ (recall that, by Proposition 2, norm and pointwise convergence are equivalent in this case). Let $\widetilde{\boldsymbol{p}} \in \mathcal{M}^+_1$ denote its limit and $\widetilde{\boldsymbol{f}} = (k)_{k\in\mathbb{N}_0}$ a random variable on $(\mathbb{N}_0, \widetilde{\boldsymbol{p}})$, to which the $\boldsymbol{f}^{(n_i)}$ converge in distribution. Since $r > 1$, the $\boldsymbol{f}^{(n_i)}$ are uniformly integrable by Markov's and Hölder's inequalities. Hence, due to [9, Lemma 3.11], their expectation values, which all equal m, converge to the one of $\widetilde{\boldsymbol{f}}$, which is thus m as well. Consider now the random variables $\boldsymbol{g}^{(n_i)} = \widetilde{\boldsymbol{g}} = (|k-m|^r)_{k\in\mathbb{N}_0}$ on $(\mathbb{N}_0, \boldsymbol{p}^{(n)})$ and $(\mathbb{N}_0, \widetilde{\boldsymbol{p}})$, respectively. The expectation values of the $\boldsymbol{g}^{(n_i)}$ are bounded by C, which, again by [9, Lemma 3.11], is then also an upper bound for the expectation value of $\widetilde{\boldsymbol{g}}$ (to which the $\boldsymbol{g}^{(n_i)}$ converge in distribution). This proves the compactness of $\mathcal{M}^+_{1,m,C}$. The convexity is obvious.

With respect to the second statement, consider

$$\begin{aligned} M_s(\mathcal{R}_0(\boldsymbol{p})) &= \sum_{i\geqslant 0} \sum_{\substack{k,\ell\geqslant 0 \\ k\wedge\ell\leqslant i\leqslant k\vee\ell}} \frac{|i-m|^s}{1+|k-\ell|}\, p_k\, p_\ell \\ &= \sum_{k,\ell\geqslant 0} \frac{p_k\, p_\ell}{1+|k-\ell|} \, \frac{1}{2} \sum_{i=k\wedge\ell}^{k\vee\ell} \big(|i-m|^s + |k+\ell-i-m|^s\big). \end{aligned} \tag{13}$$

For notational convenience, let $j = k+\ell-i$. We now show

$$|i-m|^s + |k+\ell-i-m|^s \leqslant |k-m|^s + |\ell-m|^s. \tag{14}$$

If $\{k,\ell\} = \{i,j\}$, then (14) holds with equality. Otherwise, assume, without loss of generality, that $k < i \leqslant j < \ell$. If $m \leqslant k$ or $m \geqslant \ell$, we have equality for $s = 1$ but a strict inequality for $s = r$ due to the convexity of $x \mapsto x^r$. (For $s = 1$, this describes the fact that a recombination event between two sequences that are both longer or both shorter than the mean does not change their mean distance to the mean copy number.) In the remaining cases, the inequality is strict as well. Hence, $M_s(\mathcal{R}_0(\boldsymbol{p})) \leqslant M_s(\boldsymbol{p})$ with equality if and only if $\boldsymbol{p}$ is a fixed point of $\mathcal{R}_0$, since otherwise the sum in (13) contains at least one term for which (14) holds as a strict inequality.

To see that M_1 is continuous, select a converging sequence $(\boldsymbol{p}^{(n)})$ in $\mathcal{M}^+_{1,m,C}$ and consider the random variables $\boldsymbol{h}^{(n)} = (|k-m|)_{k\in\mathbb{N}_0}$ on $(\mathbb{N}_0, \boldsymbol{p}^{(n)})$. As above, the latter are uniformly integrable, from which the continuity of M_1 follows. Since $M_1(\boldsymbol{p})$ is linear in $\boldsymbol{p}$ and thus infinitely differentiable, so is the solution $\boldsymbol{p}(t)$ for every initial condition $\boldsymbol{p}_0 \in \mathcal{M}^+_{1,m,C}$, compare [1, Thm. 9.5 and Rem. 9.6(b)]. Therefore, we have

$$\dot{M}_1(\boldsymbol{p}_0) = \liminf_{t\to 0^+} \frac{M_1(\boldsymbol{p}(t)) - M_1(\boldsymbol{p}_0)}{t} = M_1(\mathcal{R}_0(\boldsymbol{p}_0)) - M_1(\boldsymbol{p}_0) \leqslant 0,$$

again with equality if and only if $\boldsymbol{p}_0$ is a fixed point. Thus, M_1 is a Lyapunov function. $\square$

Proof of Theorem 2. By assumption, the r-th moment of $\boldsymbol{p}(0)$ exists, which is equivalent to the existence of the centered r-th moment by Minkowski's inequality [21, Sec. II.6.6]. This obviously implies the existence of the mean m. By Lemma 2, $\boldsymbol{p}(t) \in \mathcal{M}^{+}_{1,m,C}$ follows for all $t \geqslant 0$, directly for discrete time and via a satisfied subtangent condition [14, Thm. VI.2.1] (see also [1, Thm. 16.5]) for continuous time. In the discrete case, due to the compactness of $\mathcal{M}^{+}_{1,m,C}$, there is a convergent subsequence $(\boldsymbol{p}(t_i))$ with some limit $\boldsymbol{p}$. Consider now the mean distance M_1 to the mean copy number from (12). If $\lim_{t\to\infty} \boldsymbol{p}(t) = \boldsymbol{p}$, we have, due to the continuity of M_1 and $\mathcal{R}_0$,

$$M_1(\mathcal{R}_0(\boldsymbol{p})) = \lim_{t\to\infty} M_1(\mathcal{R}_0(\boldsymbol{p}(t))) = \lim_{t\to\infty} M_1(\boldsymbol{p}(t+1)) = M_1(\boldsymbol{p}),$$

thus $\boldsymbol{p}$ is a fixed point by Lemma 2. Otherwise, there are two convergent subsequences $(\boldsymbol{p}(t_i))$, with limit $\boldsymbol{p}$, and $(\boldsymbol{p}(s_i))$, with limit $\boldsymbol{q}$, whose indices alternate, $t_i < s_i < t_{i+1}$. Then, we also have $M_1(\mathcal{R}_0(\boldsymbol{p}(t_i))) \geqslant M_1(\boldsymbol{p}(s_i))$ and $M_1(\mathcal{R}_0(\boldsymbol{p}(s_i))) \geqslant M_1(\boldsymbol{p}(t_{i+1}))$, and therefore

$$\begin{aligned} M_1(\boldsymbol{p}) \geqslant M_1(\mathcal{R}_0(\boldsymbol{p})) &= \lim_{i\to\infty} M_1(\mathcal{R}_0(\boldsymbol{p}(t_i))) \geqslant \lim_{i\to\infty} M_1(\boldsymbol{p}(s_i)) = M_1(\boldsymbol{q}) \\ \geqslant M_1(\mathcal{R}_0(\boldsymbol{q})) &= \lim_{i\to\infty} M_1(\mathcal{R}_0(\boldsymbol{p}(s_i))) \geqslant \lim_{i\to\infty} M_1(\boldsymbol{p}(t_{i+1})) = M_1(\boldsymbol{p}). \end{aligned}$$

Thus, both $\boldsymbol{p}$ and $\boldsymbol{q}$ are fixed points by Lemma 2 and hence equal by Proposition 4. In continuous time, the claim follows from Theorem 1 since M_1 is a Lyapunov function by Lemma 2. $\square$

Note that, for $q = 0$, the recombinator can be expressed in terms of explicit frequencies $\pi_{k,\ell}$ of fragment pairs before concatenation (with copy numbers k and ℓ) as $\mathcal{R}_0(\boldsymbol{p})_i = \sum_{j=0}^{i} \pi_{j,i-j}$. However, we have, so far, not been able to use this for a simplification of the above treatment.

4. Alternative probability representations

Our next goal is to find the analogue of Theorem 2 for the case of $q = 1$ (random UC). Whereas the convergence arguments for the case $q = 0$ relied on a compact set of probability measure defined via the r-th moment, we are not (yet) able to extend this approach to $q > 0$. Instead, we will consider, as an alternative representation for a probability measure $\boldsymbol{p} \in \mathcal{M}^{+}_{1}$, the generating function

$$\psi(z) = \sum_{k\geqslant 0} p_k z^k, \tag{15}$$

for which $\psi(1) = \|\boldsymbol{p}\|_1 = 1$ and the radius of convergence is at least 1. We will restrict our discussion to such $\boldsymbol{p}$ for which $\limsup_{k\to\infty} \sqrt[k]{p_k} < 1$. Then, the radius of convergence, $\rho(\psi) = 1/\limsup_{k\to\infty} \sqrt[k]{p_k}$ by Hadamard's formula [20, 10.5], is larger than 1. This is, biologically, no restriction since for any 'realistic' system there are only finitely many non-zero p_k (and therefore $\rho(\psi) = \infty$). Mathematically, this condition ensures the existence of all moments and enables us to go back and forth between the probability measure and its generating function, even when looked at $\psi(z)$ only in the vicinity of $z = 1$ (see Proposition 6 below and [21, Sec. II.12]). By abuse of notation, we define the induced recombinator for these generating functions as

$$\mathcal{R}(\psi)(z) = \sum_{k \geqslant 0} \mathcal{R}(\boldsymbol{p})_k \, z^k . \tag{16}$$

In general, with the exception of the case $q = 1$, we do not know any simple expression for $\mathcal{R}(\psi)$ in terms of ψ. Nevertheless, (16) will be central to our further analysis.

It is advantageous to use the local expansion around $z = 1$, written in the form

$$\psi(z) = \sum_{k \geqslant 0} (k+1) a_k (z-1)^k , \tag{17}$$

whose coefficients are given by

$$a_k = \frac{1}{(k+1)!} \psi^{(k)}(1) = \frac{1}{k+1} \sum_{\ell \geqslant k} \binom{\ell}{k} p_\ell =: \boldsymbol{a}(\boldsymbol{p})_k \geqslant 0 . \tag{18}$$

In particular, $a_0 = 1$ and $a_1 = \frac{1}{2} \sum_{\ell \geqslant 0} \ell \, p_\ell$. This definition of a_k is size biased, and will become clear from the simplified dynamics for $q = 1$ that results from it. For the sake of compact notation, we use $\boldsymbol{a} = (a_k)_{k \in \mathbb{N}_0}$ both for the coefficients and for the mapping. The coefficients $\boldsymbol{a}$ are elements of the following compact, convex metric space.

Definition 3. *For fixed α and δ with $0 < \alpha \leqslant \delta < \infty$, let*

$$X_{\alpha,\delta} = \{ \boldsymbol{a} = (a_k)_{k \in \mathbb{N}_0} : a_0 = 1, \, a_1 = \alpha, \, 0 \leqslant a_k \leqslant \delta^k \text{ for } k \geqslant 2 \} .$$

On this space, define the metric

$$d(\boldsymbol{a}, \boldsymbol{b}) = \sum_{k \geqslant 0} d_k |a_k - b_k| \tag{19}$$

with $d_k = (\gamma/\delta)^k$ for some $0 < \gamma < \frac{1}{3}$.

It is obvious that d is indeed a metric and that $X_{\alpha,\delta}$ is a convex set, i.e., we have $\eta \, \boldsymbol{a} + (1 - \eta) \boldsymbol{b} \in X_{\alpha,\delta}$ for all $\boldsymbol{a}$, $\boldsymbol{b} \in X_{\alpha,\delta}$ and $\eta \in [0, 1]$. Note that we use the same symbol d as in (1) since it will always be clear which metric is meant. The space $X_{\alpha,\delta}$ is naturally embedded in the Banach space (cf. [26, Sec. 24.I])

$$H_{\gamma/\delta} = \{ \boldsymbol{x} \in \mathbb{R}^{\mathbb{N}_0} : \|\boldsymbol{x}\| < \infty \} \tag{20}$$

with the norm $\|\boldsymbol{x}\| = \sum_{k \geqslant 0} (\gamma/\delta)^k |x_k|$, for γ and δ as in Definition 3. In particular, $d(\boldsymbol{a}, \boldsymbol{b}) = \|\boldsymbol{a} - \boldsymbol{b}\|$. Furthermore, we have the following two propositions.

Proposition 5. *The space $X_{\alpha,\delta}$ is compact in the metric d of* (19).

Proof. In metric spaces, compactness and sequential compactness are equivalent, compare [12, Thm. II.3.8]. Hence, let $(\boldsymbol{a}^{(n)})$ be any sequence in $X_{\alpha,\delta}$. By assumption, $a_0^{(n)} \equiv 1$ and $a_1^{(n)} \equiv \alpha$. Furthermore, each element sequence $(a_k^{(n)}) \subset [0, \delta^k]$ has a convergent subsequence. We now inductively define, for every k, a convergent subsequence $(a_k^{(n_{k,i})})$, with limit a_k, such that the indices $\{n_{k,i} : i \in \mathbb{N}\}$ are a subset of the preceding indices $\{n_{k-1,i} : i \in \mathbb{N}\}$. This way, we can proceed to a 'diagonal' sequence $(\boldsymbol{a}^{(n_{i,i})})$. The latter is now shown to converge to $\boldsymbol{a} = (a_k)$, which is obviously an element of $X_{\alpha,\delta}$. To this end, let $\varepsilon > 0$ be given. Choose m large enough such that $\sum_{k>m}(2\gamma)^k < \varepsilon/2$, and then i such that $\sum_{k=0}^{m} d_k |a_k^{(n_{i,i})} - a_k| < \varepsilon/2$. Then

$$
\begin{aligned}
d(\boldsymbol{a}^{(n_{i,i})}, \boldsymbol{a}) &= \sum_{k=2}^{m} d_k |a_k^{(n_{i,i})} - a_k| + \sum_{k>m} d_k |a_k^{(n_{i,i})} - a_k| \\
&< \frac{\varepsilon}{2} + \sum_{k>m} (2\gamma)^k < \varepsilon,
\end{aligned}
\tag{21}
$$

which proves the claim. □

Proposition 6. *If* $\limsup_{k\to\infty} \sqrt[k]{p_k} < 1$, *the coefficients* a_k *from* (18) *exist and* $\boldsymbol{a}(\boldsymbol{p}) \in X_{\alpha,\delta}$ *with* $\alpha = \boldsymbol{a}(\boldsymbol{p})_1 = \frac{1}{2}m = \frac{1}{2}\sum_{k\geqslant 0} k\, p_k$ *and some* δ. *Conversely, if* $\boldsymbol{p}(\boldsymbol{a}) \in X_{\alpha,\delta}$ *for some* α, δ, *one has* $\limsup_{k\to\infty} \sqrt[k]{p_k} < 1$.

For a proof, we need the following

Lemma 3. *Let* $f_0(z) = \sum_{k\geqslant 0} c_k z^k$ *be a power series with non-negative coefficients* c_k *and* $f_x(z) = \sum_{k\geqslant 0} \frac{1}{k!} f_0^{(k)}(x)(z-x)^k$ *the expansion of* f_0 *around some* $x \in [0, \rho(f_0)[$. *Then,* $\rho(f_0) = x + \rho(f_x)$, *including the case that both radii of convergence are infinite.*

Proof. Since the open disc $B_x(\rho(f_0) - x)$ is entirely included in $B_0(\rho(f_0))$, the inequality $\rho(f_x) \geqslant \rho(f_0) - x$ immediately follows from the theorem of representability by power series [20, Thm. 10.16]. Consider now the power series $f_{xe^{i\varphi}}(z) = \sum_{k\geqslant 0} \frac{1}{k!} f_0^{(k)}(xe^{i\varphi})(z - xe^{i\varphi})^k$ with arbitrary $\varphi \in [0, 2\pi[$. Its coefficients satisfy $|f_0^{(k)}(xe^{i\varphi})| \leqslant \sum_{n\geqslant k} \frac{n!}{(n-k)!} c_k x^{n-k} = f_0^{(k)}(x)$ due to the non-negativity of the c_k. This implies $\rho(f_{xe^{i\varphi}}) \geqslant \rho(f_x)$ by Hadamard's formula. Therefore, f_0 admits an analytic continuation on $B_0(x + \rho(f_x))$, the uniqueness of which follows from the monodromy theorem [20, Thm. 16.16]. The theorem of representability by power series then implies the inequality $\rho(f_0) \geqslant x + \rho(f_x)$, which, together with the opposite inequality above, proves the claim. □

Proof of Proposition 6. The assumption implies $\rho(\psi) > 1$ for ψ from (15). Then, from Lemma 3, we know that $\limsup_{k\to\infty} \sqrt[k]{(k+1)a_k} < \infty$. Since furthermore $a_k \leqslant (k+1)a_k$, also $\limsup_{k\to\infty} \sqrt[k]{a_k} < \infty$, so there is an upper bound δ for $\sqrt[k]{a_k}$ and thus $\boldsymbol{a}(\boldsymbol{p}) \in X_{\alpha,\delta}$. The converse statement follows from (18) and Lemma 3. □

Therefore, any mapping from $X_{\alpha,\delta}$ into itself that is continuous with respect to the metric d from (19) has a fixed point by the Leray–Schauder–Tychonov theorem [18, Thm. V.19].

Note further that each $X_{\alpha,\delta}$ contains a maximal element with respect to the partial order $\boldsymbol{a} \leqslant \boldsymbol{b}$ defined by $a_k \leqslant b_k$ for all $k \in \mathbb{N}_0$, which is given by $(1, \alpha, \delta^2, \delta^3, \ldots)$. This property finally leads to

Proposition 7. *The space $P_{\alpha,\delta} := \{\boldsymbol{p} \in \mathcal{M}_1^+ : \boldsymbol{a}(\boldsymbol{p}) \in X_{\alpha,\delta}\}$, equipped with (the metric induced by) the total variation norm, is compact and convex.*

The proof is based on the following two lemmas.

Lemma 4. *For any subset of $P_{\alpha,\delta}$, the corresponding generating functions from* (15) *are locally bounded on $B_{1+1/\delta}(0)$.*

Proof. It is sufficient to show boundedness on every compact $K \subset B_{1+1/\delta}(0)$, see [19, Sec. 7.1]. Thus, let such a K be given and fix $r \in [0, \frac{1}{\delta}[$ so that K is contained in $\overline{B_{1+r}(0)}$. Then, for every $\boldsymbol{p} \in P_{\alpha,\delta}$ and every $z \in K$,

$$
\begin{aligned}
|\psi(z)| &= \Big|\sum_{k\geqslant 0} p_k z^k\Big| \leqslant \sum_{k\geqslant 0} p_k (1+r)^k = \psi(1+r) \\
&= \sum_{k\geqslant 0} (k+1)\boldsymbol{a}(\boldsymbol{p})_k \, r^k \leqslant 1 + 2\,\alpha\, r + \sum_{k\geqslant 2} (k+1)(r\delta)^k < \infty,
\end{aligned}
$$

where $r\delta < 1$ was used. This needed to be shown. □

Lemma 5. *If, for a sequence $(\boldsymbol{p}^{(n)}) \subset P_{\alpha,\delta}$, the coefficients $\boldsymbol{a}^{(n)} = \boldsymbol{a}(\boldsymbol{p}^{(n)})$ from* (18) *converge to some $\boldsymbol{a}$ with respect to the metric d from* (19)*, then the generating functions ψ_n from* (15) *converge compactly to some ψ with $\psi(z) = \sum_{k\geqslant 0} p_k z^k$ and the $\boldsymbol{p}^{(n)}$ thus converge in norm to $\boldsymbol{p} \in P_{\alpha,\delta}$.*

Proof. By to Lemma 4, the sequence (ψ_n) is locally bounded in $B_{1+1/\delta}(0)$. Due to the pointwise convergence $|a_k^{(n)} - a_k| \leqslant d_k^{-1} d(\boldsymbol{a}^{(n)}, \boldsymbol{a}) \to 0$, we have

$$
\psi_n^{(k)}(1) = (k+1)!\, a_k^{(n)} \xrightarrow{n\to\infty} (k+1)!\, a_k = \psi^{(k)}(1)
$$

for every $k \in \mathbb{N}_0$. Then, the compact convergence $\psi_n \to \psi$ follows from Vitali's theorem [19, Thm. 7.3.2]. In particular, this implies that $p_k^{(n)} \to p_k \geqslant 0$ and $1 = \sum_{k\geqslant 0} p_k^{(n)} = \psi_n(1) \to \psi(1) = \sum_{k\geqslant 0} p_k$, thus $\boldsymbol{p} \in \mathcal{M}_1^+$.

Now, choose $r \in \,]1, 1+\frac{1}{\delta}[$. Then there is, for every $\varepsilon > 0$, an n_ε such that $\sup_{|z|\leqslant r} |\psi(z) - \psi_n(z)| < \varepsilon$ for all $n \geqslant n_\varepsilon$. This implies

$$
|p_k^{(n)} - p_k| = \left| \frac{1}{2\pi i} \oint_{|z|=r} \frac{\psi_n(z) - \psi(z)}{z^{k+1}} \, \mathrm{d}z \right| < \frac{\varepsilon}{r^k}
$$

for all $n \geqslant n_\varepsilon$ by Cauchy's integral formula [13, Thm. 7.3]. Now, let $\varepsilon > 0$ be given. Then

$$
\|\boldsymbol{p}^{(n)} - \boldsymbol{p}\|_1 = \sum_{k\geqslant 0} |p_k^{(n)} - p_k| < \varepsilon \frac{1}{1 - \frac{1}{r}}
$$

for all $n \geqslant n_\varepsilon$, which proves the claim. □

Proof of Proposition 7. Let $(\boldsymbol{p}^{(n)})$ denote an arbitrary sequence in $P_{\alpha,\delta}$ and $(\boldsymbol{a}^{(n)}) = (\boldsymbol{a}(\boldsymbol{p}^{(n)}))$ the corresponding sequence in $X_{\alpha,\delta}$. Due to Proposition 5, there is a convergent subsequence $(\boldsymbol{a}^{(n_i)})_i$. Then, by Lemma 5, $(\boldsymbol{p}^{(n_i)})_i$ converges in norm to some $\boldsymbol{p} \in P_{\alpha,\delta}$. This proves the compactness property. The convexity of $P_{\alpha,\delta}$ is a simple consequence of the convexity of $\mathcal{M}_1^+$, the linearity of the mapping $\boldsymbol{a}$, and the convexity of $X_{\alpha,\delta}$. □

Another property of the mapping $\boldsymbol{a}\colon P_{\alpha,\delta} \to X_{\alpha,\delta}$ is stated in

Lemma 6. *For every α and δ, the mapping $\boldsymbol{a}\colon P_{\alpha,\delta} \to X_{\alpha,\delta}$ from* (18) *is continuous* (*with respect to the total variation norm and the metric d*) *and injective. Its inverse $\boldsymbol{p}\colon \boldsymbol{a}(P_{\alpha,\delta}) \to P_{\alpha,\delta}$ is continuous as well.*

Proof. Let $\boldsymbol{p}, \boldsymbol{q} \in P_{\alpha,\delta}$ and assume $\boldsymbol{a}(\boldsymbol{p}) = \boldsymbol{a}(\boldsymbol{q})$. Then, as in the proof of Lemma 3, the uniqueness of the generating function in $B_{1+1/\delta}(0)$ follows, and thus $\boldsymbol{p} = \boldsymbol{q}$, which proves the injectivity of $\boldsymbol{a}$. The other statements follow from Vitali's theorem [19, Thm. 7.3.2]: Norm convergence of a sequence $(\boldsymbol{p}^{(n)})$ in $P_{\alpha,\delta}$ to some $\boldsymbol{p}$ implies convergence of its element sequences and thus compact convergence of the corresponding generating functions ψ_n to ψ, which is given by $\psi(z) = \sum_{k\geqslant 0} p_k\, z^k$. This, in turn, implies convergence of each sequence $(\boldsymbol{a}(\boldsymbol{p}^{(n)})_k)$ to $\boldsymbol{a}(\boldsymbol{p})_k$, from which, as in (21), the convergence $(\boldsymbol{a}(\boldsymbol{p}^{(n)})) \to \boldsymbol{a}(\boldsymbol{p})$ (with respect to d) follows. The converse is the statement of Lemma 5 (see also [16, Prop. 1.6.8]). □

Note that, if $\rho(\psi) > 2$, the inverse of the mapping $\boldsymbol{a}$ is given by

$$\boldsymbol{p}(\boldsymbol{a})_k = \sum_{\ell \geqslant k} (-1)^{\ell-k} \binom{\ell}{k} (\ell+1)\, a_\ell .$$

5. Random unequal crossover

Let us now turn to the random UC model, described by $q = 1$ in (7). Here, the recombinator (2) simplifies to [22, (3.1)]

$$\mathcal{R}_1(\boldsymbol{p})_i = \sum_{\substack{k,\ell \geqslant 0 \\ k+\ell \geqslant i}} \frac{1 + \min\{k, \ell, i, k+\ell-i\}}{(k+1)(\ell+1)}\, p_k\, p_\ell . \tag{22}$$

As for internal UC, by Lemma 1, the reversibility condition (10) directly leads to an expression for fixed points,

$$\frac{p_k}{k+1}\frac{p_\ell}{\ell+1} = \frac{p_i}{i+1}\frac{p_j}{j+1} \quad \text{for all } k+\ell = i+j .$$

This has $p_k = C(k+1)x^k$ as a solution, with appropriate parameter x and normalization constant C. Again, it turns out that all fixed points are given this way.

Proposition 8. [22, Thm. A.2] *Every fixed point* $\boldsymbol{p} \in \mathcal{M}_1^+$ *of* $\mathcal{R}_1$ *is of the form*

$$p_k = \left(\frac{2}{m+2}\right)^2 (k+1) \left(\frac{m}{m+2}\right)^k, \tag{23}$$

where $m = \sum_{k \geqslant 0} k\, p_k \geqslant 0$. □

One can verify this in several ways, one being a direct inductive calculation.

The main result of this section is

Theorem 3. *Assume that* $\limsup_{k\to\infty} \sqrt[k]{p_k(0)} < 1$. *Then, both in discrete and in continuous time,* $\lim_{t\to\infty} \|\boldsymbol{p}(t) - \boldsymbol{p}\|_1 = 0$ *with the appropriate fixed point* $\boldsymbol{p}$ *from Proposition* 8.

For a proof, we consider the following alternative process, verbally described in [22, p. 720f]. It is a two-step stick breaking and glueing procedure which ultimately induces the same (deterministic) dynamics as random UC, even though the underlying process is rather different. This will lead to a simple expression for the induced recombinator of the coefficients $\boldsymbol{a}$ from (18), which allows for an explicit solution.

Proposition 9. *Let* $\boldsymbol{p} \in \mathcal{M}_1^+$. *Then,*

$$\pi_k = \sum_{\ell \geqslant k} \frac{1}{\ell+1} p_\ell \tag{24}$$

gives a probability measure $\boldsymbol{\pi} \in \mathcal{M}_1^+$, *and the recombinator can be written as*

$$\mathcal{R}_1(\boldsymbol{p})_i = \sum_{j=0}^{i} \pi_j\, \pi_{i-j} = (\boldsymbol{\pi} * \boldsymbol{\pi})_i, \tag{25}$$

where $*$ *denotes the convolution in* $\ell^1(\mathbb{N}_0)$.

Here, (24) describes a breaking process in which, without any pairing, each sequence is cut equally likely between any two of its building blocks. In a second step, described by (25), these fragments are paired randomly and joined (or 'glued').

Proof. It is easily verified that $\boldsymbol{\pi}$ is normalized to 1. With respect to (25), note the following identity for $k + \ell \geqslant i$,

$$|\{j : (i-\ell) \vee 0 \leqslant j \leqslant i \wedge k\}| = 1 + \min\{k, \ell, i, k+\ell-i\},$$

which can be shown by treating the four cases on the LHS separately. With this, inserting (24) into the RHS of (25) yields

$$\sum_{j=0}^{i} \pi_j\, \pi_{i-j} = \sum_{j=0}^{i} \sum_{k \geqslant j} \sum_{\ell \geqslant i-j} \frac{1}{(k+1)(\ell+1)}\, p_k\, p_\ell$$

$$= \sum_{\substack{k,\ell \geqslant 0 \\ k+\ell \geqslant i}} \frac{1}{(k+1)(\ell+1)}\, p_k\, p_\ell \sum_{j=(i-\ell)\vee 0}^{i \wedge k} 1$$
$$= \sum_{\substack{k,\ell \geqslant 0 \\ k+\ell \geqslant i}} \frac{1+\min\{k,\ell,i,k+\ell-i\}}{(k+1)(\ell+1)}\, p_k\, p_\ell \;=\; \mathcal{R}_1(\boldsymbol{p})_i\,.$$

□

This nice structure has an analogue on the level of the generating functions.

Proposition 10. *Under the assumptions of Theorem* 3*, let* $\phi(z) = \sum_{k\geqslant 0} \pi_k z^k$ *denote the generating function for* $\boldsymbol{\pi}$ *from* (24)*. Then,*

$$\phi(z) = \frac{1}{1-z} \int_z^1 \psi(\zeta)\,\mathrm{d}\zeta \qquad \textit{and} \qquad \mathcal{R}_1(\psi)(z) = \phi(z)^2.$$

Proof. Recall that $\psi(z) = \sum_{k\geqslant 0} p_k z^k$. Equations (24) and (25) lead to

$$\phi(z) = \sum_{k\geqslant 0}\sum_{\ell\geqslant k} \frac{1}{\ell+1} p_\ell z^k = \sum_{\ell\geqslant 0} \frac{1}{\ell+1} p_\ell \sum_{k\leqslant \ell} z^k = \sum_{\ell\geqslant 0} \frac{1}{\ell+1} p_\ell \frac{1-z^{\ell+1}}{1-z}$$
$$= \frac{1}{1-z} \sum_{\ell\geqslant 0} p_\ell \frac{1-z^{\ell+1}}{\ell+1} = \frac{1}{1-z} \sum_{\ell\geqslant 0} p_\ell \int_z^1 \zeta^\ell\,\mathrm{d}\zeta = \frac{1}{1-z}\int_z^1 \psi(\zeta)\,\mathrm{d}\zeta$$

and, due to absolute convergence of the series involved,

$$\mathcal{R}_1(\psi)(z) = \sum_{k\geqslant 0} \mathcal{R}_1(\boldsymbol{p})_k z^k = \sum_{k\geqslant 0} z^k \sum_{\ell=0}^{k} \pi_\ell \pi_{k-\ell}$$
$$= \sum_{\ell\geqslant 0} \pi_\ell z^\ell \sum_{k\geqslant \ell} \pi_{k-\ell} z^{k-\ell} = \phi(z)^2.$$

□

The following lemma states that the radius of convergence of ψ does not decrease under the random UC dynamics. Thus, it is ensured that, if $\rho(\psi) > 1$, also $\mathcal{R}_1(\psi)$ may be described by an expansion at $z = 1$, i.e., by coefficients $\boldsymbol{a}$.

Lemma 7. *The radius of convergence of* $\mathcal{R}_1(\psi)$ *is* $\rho(\mathcal{R}_1(\psi)) \geqslant \rho(\psi)$*.*

Proof. As $1/\rho(\psi) = \limsup_{k\to\infty} \sqrt[k]{p_k} =: x \leqslant 1$ and $\lim_{k\to\infty} \sqrt[k]{k+1} = 1$, there is a constant $C > 0$ with $p_k \leqslant C(k+1)x^k$ for all k. Note the identity

$$\sum_{j=0}^{n} (1+\min\{i,j,n-i,n-j\}) = (i+1)(n-i+1)$$

for $i \leqslant n$, which follows from an elementary calculation. Then, (22) implies

$$\mathcal{R}_1(\boldsymbol{p})_i \leqslant C^2 \sum_{\substack{k,\ell\geqslant 0 \\ k+\ell\geqslant i}} x^{k+\ell} (1+\min\{k,\ell,i,k+\ell-i\})$$

$$
\begin{aligned}
&= C^2 \sum_{n \geqslant i} x^n \sum_{j=0}^{n} (1 + \min\{i, j, n-i, n-j\}) \\
&= C^2 (i+1) x^i \sum_{\ell \geqslant 0} (\ell + 1)\, x^\ell = \left(\frac{C}{1-x}\right)^2 (i+1)\, x^i .
\end{aligned}
$$

Accordingly, $\limsup_{k\to\infty} \sqrt[k]{\mathcal{R}_1(\boldsymbol{p})_k} \leqslant x \leqslant 1$, which proves the claim. □

These results enable us to derive the following expression for the coefficients $\boldsymbol{a}$, using the expansion of (17):

$$
\begin{aligned}
\mathcal{R}_1(\psi)(z) &= \left[\frac{1}{z-1}\int_1^z \psi(\zeta)\,\mathrm{d}\zeta\right]^2 \\
&= \left[\sum_{k\geqslant 0} a_k (z-1)^k\right]^2 = \sum_{k\geqslant 0}\left(\sum_{n=0}^{k} a_n a_{k-n}\right)(z-1)^k .
\end{aligned}
\tag{26}
$$

So it is natural to define the induced recombinator

$$
\tilde{\mathcal{R}}_1(\boldsymbol{a})_k = \frac{1}{k+1}\sum_{n=0}^{k} a_n a_{k-n} \geqslant 0, \tag{27}
$$

for which we have

Lemma 8. *The recombinator $\tilde{\mathcal{R}}_1$ given by* (27) *maps each space $X_{\alpha,\delta}$ into itself and is continuous with respect to the metric d from* (19).

Proof. Let $\alpha, \delta > 0$ be given and $\boldsymbol{a}, \boldsymbol{b} \in X_{\alpha,\delta}$. Trivially, $\tilde{\mathcal{R}}_1(\boldsymbol{a})_0 = 1$ and $\tilde{\mathcal{R}}_1(\boldsymbol{a})_1 = \alpha$. For $k \geqslant 2$, $\tilde{\mathcal{R}}_1(\boldsymbol{a})_k = \frac{1}{k+1}\sum_{\ell\leqslant k} a_\ell\, a_{k-\ell} \leqslant \delta^k$. This proves the first statement. For the continuity, note first that every $\tilde{\mathcal{R}}_1(\boldsymbol{a})_k$ with $k \geqslant 2$ is continuous as a mapping from $X_{\alpha,\delta}$ to $[0, \delta^k]$. Now, let $\varepsilon > 0$ be given. Choose n large enough so that $\sum_{k>n}(2\gamma)^k < \varepsilon/2$, where γ is the parameter introduced in Definition 3. Then, there is an $\eta > 0$ such that $\sum_{k=2}^{n}(\gamma/\delta)^k |\tilde{\mathcal{R}}_1(\boldsymbol{a})_k - \tilde{\mathcal{R}}_1(\boldsymbol{b})_k| < \varepsilon/2$ for $\boldsymbol{a}, \boldsymbol{b} \in X_{\alpha,\delta}$ with $d(\boldsymbol{a}, \boldsymbol{b}) < \eta$. Thus, for such $\boldsymbol{a}$ and $\boldsymbol{b}$,

$$
d(\tilde{\mathcal{R}}_1(\boldsymbol{a}), \tilde{\mathcal{R}}_1(\boldsymbol{b})) \leqslant \sum_{k=0}^{n}\left(\frac{\gamma}{\delta}\right)^k |\tilde{\mathcal{R}}_1(\boldsymbol{a})_k - \tilde{\mathcal{R}}_1(\boldsymbol{b})_k| + \sum_{k>n}(2\gamma)^k < \varepsilon,
$$

which proves the claim. □

Note that the fixed point equation on the level of the coefficients $\boldsymbol{a}$ is always satisfied for a_0 and a_1. If $k > 1$, one obtains the recursion

$$
a_k = \frac{1}{k-1}\sum_{n=1}^{k-1} a_n a_{k-n},
$$

which shows that at most one fixed point with given mean can exist.

Let us now consider the case of discrete time first. Analogously to (5), define $\boldsymbol{a}(t) = \boldsymbol{a}(\boldsymbol{p}(t))$ as the coefficients belonging to $\boldsymbol{p}(t)$, which are assumed to exist. It is clear from (16), (26) and (27) that $\boldsymbol{a}(t+1) = \tilde{\mathcal{R}}_1(\boldsymbol{a}(t))$. We then have the following two propositions.

Proposition 11. *Assume $\boldsymbol{a}(0)$ to exist. Then, in discrete time,*

$$\lim_{t\to\infty} a_k(t) = \alpha^k \qquad \textit{for all } k \geqslant 0.$$

This result indicates that a weaker condition than the one of Theorem 3 may be sufficient for convergence of $\boldsymbol{p}(t)$.

Proof. Clearly, $a_0(t) \equiv 1$, $a_1(t) \equiv \alpha$. Furthermore, by the assumption and (26), the coefficients $a_k(t)$ exist for all $k, t \in \mathbb{N}_0$. Now, assume that the claim holds for all $k \leqslant n$ with some n and let $k = n+1$. According to the general properties of $\limsup$ and $\liminf$, we then have

$$\begin{aligned}\limsup_{t\to\infty}\, a_k(t+1) &\leqslant \frac{1}{k+1}\sum_{\ell=0}^{k} \limsup_{t\to\infty}\big(a_\ell(t)a_{k-\ell}(t)\big)\\ &= \frac{k-1}{k+1}\alpha^k + \frac{2}{k+1}\limsup_{t\to\infty}\, a_k(t)\end{aligned}$$

and analogously with $\geqslant$ for $\liminf$. This leads to

$$\frac{k-1}{k+1}\limsup_{t\to\infty}\, a_k(t) \leqslant \frac{k-1}{k+1}\alpha^k \leqslant \frac{k-1}{k+1}\liminf_{t\to\infty} a_k(t),$$

from which the claim follows for all $k \leqslant n+1$ and, by induction over n, for all $k \geqslant 0$. □

Proposition 12. *The recombinator $\tilde{\mathcal{R}}_1$, acting on $X_{\alpha,\delta}$, is a strict contraction with respect to the metric d from* (19)*, i.e., there is a $C < 1$ such that, for all elements $\boldsymbol{a}, \boldsymbol{b} \in X_{\alpha,\delta}$,*

$$d(\tilde{\mathcal{R}}_1(\boldsymbol{a}), \tilde{\mathcal{R}}_1(\boldsymbol{b})) \leqslant C\, d(\boldsymbol{a}, \boldsymbol{b}).$$

Proof. First consider, for $k \geqslant 2$, without using the special choice of the d_k,

$$\begin{aligned}d(\tilde{\mathcal{R}}_1(\boldsymbol{a}), \tilde{\mathcal{R}}_1(\boldsymbol{b})) &= \sum_{k\geqslant 2} d_k \frac{1}{k+1}\Big|\sum_{\ell=0}^{k}(a_\ell\, a_{k-\ell} - b_\ell\, b_{k-\ell})\Big|\\ &= \sum_{k\geqslant 2} d_k \frac{1}{k+1}\Big|\sum_{\ell=0}^{k}(a_\ell - b_\ell)(a_{k-\ell} + b_{k-\ell})\Big|\\ &\leqslant \sum_{k\geqslant 2} d_k \frac{2}{k+1}\sum_{\ell=2}^{k}\delta^{k-\ell}|a_\ell - b_\ell|\\ &= \sum_{\ell\geqslant 2} d_\ell |a_\ell - b_\ell| \sum_{k\geqslant \ell}\frac{2}{k+1}\delta^{k-\ell}\frac{d_k}{d_\ell}.\end{aligned} \tag{28}$$

With the choice $d_k = (\gamma/\delta)^k$, where we had $\gamma < \frac{1}{3}$, we can find, for $\ell \geqslant 2$, an upper bound for the inner sum,

$$\sum_{k\geqslant\ell} \frac{2}{k+1}\delta^{k-\ell}\frac{d_k}{d_\ell} \leqslant \frac{2}{3}\sum_{k\geqslant\ell}\gamma^{k-\ell} = \frac{2}{3-3\gamma} =: C < 1,$$

which, together with (28), proves the claim. $\square$

Together with Banach's fixed point theorem (compare [18, Thm. V.18]), the two propositions imply that $\boldsymbol{a}(t)$ converges to $(1, \alpha, \alpha^2, \ldots)$ with respect to the metric d, and that convergence is exponentially fast.

In continuous time, we consider the time derivative of $\boldsymbol{a}(t) := \boldsymbol{a}(\boldsymbol{p}(t))$, which is, by (18),

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{a}(t) = \tfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{a}\big(\boldsymbol{p}(t)\big) = \boldsymbol{a}\big(\mathcal{R}_1(\boldsymbol{p}(t)) - \boldsymbol{p}(t)\big) = \tilde{\mathcal{R}}_1(\boldsymbol{a}(t)) - \boldsymbol{a}(t). \tag{29}$$

The following lemma ensures, together with [1, Thm. 7.6 and Rem. 7.10(b)], that this initial value problem has a unique solution for all $\boldsymbol{a}(0) = \boldsymbol{a}_0 \in X_{\alpha,\delta}$.

Lemma 9. *Consider the Banach space $H_{\gamma/\delta}$ from* (20)*, with some $0 < \gamma < \frac{1}{3}$, and its open subset $Y = \{\boldsymbol{x} \in H_{\gamma/\delta} : |x_k| < (2\delta)^k\}$. Then, the recombinator $\tilde{\mathcal{R}}_1$ from* (27) *maps Y into itself, satisfies a global Lipschitz condition, and is bounded on Y. Furthermore, it is infinitely differentiable, $\tilde{\mathcal{R}}_1 \in \mathrm{C}^\infty(Y, Y)$.*

Proof. For $\boldsymbol{x} \in Y$, one has $|x_k| < (2\delta)^k$, hence $|\tilde{\mathcal{R}}_1(\boldsymbol{x})_k| < (2\delta)^k$, with a similar argument as in the proof of Lemma 8. Consequently, $\tilde{\mathcal{R}}_1(Y) \subset Y$. So, let $\boldsymbol{x}, \boldsymbol{y} \in Y$. Then, similarly to the proof of Proposition 12, one shows the Lipschitz condition

$$\|\tilde{\mathcal{R}}_1(\boldsymbol{x}) - \tilde{\mathcal{R}}_1(\boldsymbol{y})\| \leqslant \sum_{\ell\geqslant 0}\Big(\frac{\gamma}{\delta}\Big)^{\ell}|x_\ell - y_\ell|\sum_{k\geqslant\ell}\frac{2}{k+1}(2\gamma)^{k-\ell} \leqslant \frac{2}{1-2\gamma}\|\boldsymbol{x} - \boldsymbol{y}\|$$

and, since $\|\boldsymbol{x}\| < 1/(1-2\gamma)$ in Y, the boundedness,

$$\|\tilde{\mathcal{R}}_1(\boldsymbol{x})\| \leqslant \frac{1}{1-2\gamma}\|\boldsymbol{x}\| < \frac{1}{(1-2\gamma)^2}.$$

With respect to differentiability, consider, for sufficiently small $\boldsymbol{h} \in Y$,

$$\tilde{\mathcal{R}}_1(\boldsymbol{x}+\boldsymbol{h})_k = \tilde{\mathcal{R}}_1(\boldsymbol{x})_k + \frac{2}{k+1}\sum_{\ell=0}^{k} x_{k-\ell}\, h_\ell + \tilde{\mathcal{R}}_1(\boldsymbol{h})_k.$$

Since

$$\begin{aligned}\|\tilde{\mathcal{R}}_1(\boldsymbol{h})\| &\leqslant \sum_{k\geqslant 0}\Big(\frac{\gamma}{\delta}\Big)^{k}\frac{1}{k+1}\sum_{\ell=0}^{k}|h_{k-\ell}|\,|h_\ell| \\ &= \sum_{\ell\geqslant 0}\Big(\frac{\gamma}{\delta}\Big)^{\ell}|h_\ell|\sum_{k\geqslant\ell}\Big(\frac{\gamma}{\delta}\Big)^{k-\ell}\frac{|h_{k-\ell}|}{k+1} \leqslant \|\boldsymbol{h}\|^2,\end{aligned}$$

it is clear that $\tilde{\mathcal{R}}_1$ is differentiable with linear (and thus continuous) derivative, whose Jacobi matrix is explicitly $\tilde{\mathcal{R}}_1'(\boldsymbol{x})_{k\ell} = \frac{\partial}{\partial x_\ell}\tilde{\mathcal{R}}_1(\boldsymbol{x})_k = \frac{2}{k+1}x_{k-\ell}$ if $k \geqslant \ell$ and zero otherwise, hence one has $\tilde{\mathcal{R}}_1 \in \mathrm{C}^1(Y, Y)$. It is now trivial to show that $\tilde{\mathcal{R}}_1 \in \mathrm{C}^2(Y, Y)$ with constant second derivative and thus $\tilde{\mathcal{R}}_1 \in \mathrm{C}^\infty(Y, Y)$. $\square$

Proposition 13. *If* $\boldsymbol{a}_0 \in X_{\alpha,\delta}$ *for some* α, δ, *then* $\boldsymbol{a}(t) \in X_{\alpha,\delta}$ *for all* $t \geqslant 0$ *and* $\lim_{t\to\infty} d(\boldsymbol{a}(t), \boldsymbol{\alpha}) = 0$ *with* $\boldsymbol{\alpha} = (1, \alpha, \alpha^2, \alpha^3, \ldots)$.

Proof. The first statement follows from [14, Thm. VI.2.1] (see also [1, Thm. 16.5]) since, due to the convexity of $X_{\alpha,\delta}$, we have $\boldsymbol{a} + t(\tilde{\mathcal{R}}(\boldsymbol{a}) - \boldsymbol{a}) \in X_{\alpha,\delta}$ for every $\boldsymbol{a} \in X_{\alpha,\delta}$ and $t \in [0,1]$, hence a subtangent condition is satisfied. For the second, observe that $\tilde{\mathcal{R}}_1(\boldsymbol{\alpha}) = \boldsymbol{\alpha}$. We now show that

$$L(\boldsymbol{a}_0) = d(\boldsymbol{a}_0, \boldsymbol{\alpha}) \tag{30}$$

is a Lyapunov function, cf. Definition 1. With the notation of Lemma 9, note that the compact metric space $X_{\alpha,\delta}$ is contained in the open subset Y of the Banach space $H_{\gamma/\delta}$. The continuity of L is obvious. Now, let $\boldsymbol{a}_0 \in X_{\alpha,\delta}$ be given. By Lemma 9 and [1, Thm. 9.5 and Rem. 9.6(b)], the solution $\boldsymbol{a}(t)$ of (29) is infinitely differentiable. Thus, for $t \in [0,1]$,

$$\begin{aligned} L(\boldsymbol{a}(t)) - L(\boldsymbol{a}_0) &= \|\boldsymbol{a}_0 + t(\tilde{\mathcal{R}}_1(\boldsymbol{a}_0) - \boldsymbol{a}_0) + \mathcal{O}(t) - \boldsymbol{\alpha}\| - \|\boldsymbol{a}_0 - \boldsymbol{\alpha}\| \\ &\leqslant t\big(\|\tilde{\mathcal{R}}_1(\boldsymbol{a}_0) - \tilde{\mathcal{R}}_1(\boldsymbol{\alpha})\| - \|\boldsymbol{a}_0 - \boldsymbol{\alpha}\|\big) + \mathcal{O}(t), \end{aligned} \tag{31}$$

where $\mathcal{O}(t)$ is the usual Landau symbol and represents some function that vanishes faster than t as $t \to 0$. From this, by the strict contraction property of $\tilde{\mathcal{R}}_1$ (Proposition 12), the Lyapunov property (8) follows, with equality if and only if $\boldsymbol{a}_0 = \boldsymbol{\alpha}$. Since $X_{\alpha,\delta}$ is compact, Theorem 1 implies the claim. □

We are now able to give the previously postponed

Proof of Theorem 3. By Proposition 6, we have $\boldsymbol{a}(0) = \boldsymbol{a}(\boldsymbol{p}(0)) \in X_{\alpha,\delta}$ with $\alpha = \frac{1}{2}m$ and some δ. In discrete time, according to Propositions 11 and 12 and Banach's fixed point theorem (compare [18, Thm. V.18]), it then follows that $\boldsymbol{a}(t) \to \boldsymbol{\alpha} = (1, \alpha, \alpha^2, \ldots)$ with respect to the metric d. Inserting (23) into (18) and letting $x = m/(m+2)$ yields

$$\begin{aligned} a_k &= \sum_{\ell \geqslant k} \frac{\ell!}{(\ell-k)!(k+1)!}(1-x)^2(\ell+1)x^\ell \\ &= (1-x)^2 \sum_{\ell \geqslant k} \binom{\ell+1}{k+1} x^\ell = \left(\frac{x}{1-x}\right)^k = \alpha^k. \end{aligned}$$

The claim now follows from Lemma 5. Similarly, in continuous time, the claim follows from Proposition 13. □

Let us finally note

Proposition 14. *For the dynamics described by* (29)*, the fixed point* $\boldsymbol{\alpha}$ *from Proposition* 13 *is exponentially stable.*

Proof. Let $\boldsymbol{a}_0 \in X_{\alpha,\delta}$ be arbitrary. The Lyapunov function from the proof of Proposition 13 satisfies, as a consequence of (31) and Proposition 12,

$$\dot{L}(\boldsymbol{a}_0) \leqslant d(\tilde{\mathcal{R}}_1(\boldsymbol{a}_0), \tilde{\mathcal{R}}_1(\boldsymbol{\alpha})) - d(\boldsymbol{a}_0, \boldsymbol{\alpha}) \leqslant -(1-C)\, d(\boldsymbol{a}_0, \boldsymbol{\alpha}),$$

with $0 < C < 1$. From this, together with (30) and [1, Thm. 18.7], the claim follows. □

Remark. In a related UC model introduced by Takahata [24], for which

$$T_{ij,k\ell} = \delta_{i+j,k+\ell} \frac{1}{k+\ell+1},$$

the recombinator $\tilde{\mathcal{R}}_1$ appears for the coefficients $\boldsymbol{b}(\boldsymbol{p})_k = (k+1)\,\boldsymbol{a}(\boldsymbol{p})_k$, where $\boldsymbol{b}(\boldsymbol{p})_1$ is the mean copy number m. The above results then imply, under the appropriate condition on $\boldsymbol{p}(0)$, that $\boldsymbol{b}(t) \to (1, m, m^2, \ldots)$ as $t \to \infty$ both in discrete and in continuous time. This corresponds to convergence of $\boldsymbol{p}(t)$ to the fixed point $\boldsymbol{p}$ with $p_k = \frac{1}{m+1}(\frac{m}{m+1})^k$.

6. The intermediate parameter regime

In this section, q may take any value in $[0,1]$. With respect to reversibility of fixed points, one finds

Proposition 15. *For parameter values $q \in \,]0,1[$, any fixed point $\boldsymbol{p} \in \mathcal{M}_1^+$ of the recombinator $\mathcal{R}_q$, given by* (2) *and* (7)*, satisfies $p_k > 0$ for all $k \geqslant 0$* (*unless it is the trivial fixed point $\boldsymbol{p} = (1,0,0,\ldots)$ we excluded*)*. None of these extra fixed points is reversible.*

Proof. Let a non-trivial fixed point $\boldsymbol{p}$ be given and choose any $n > 0$ with $p_n > 0$. Observe that $T^{(q)}_{n+1\,n-1,nn} > 0$ for $0 < q < 1$ and hence

$$p_{n\pm1} = \mathcal{R}_q(\boldsymbol{p})_{n\pm1} = \sum_{j,k,\ell \geqslant 0} T^{(q)}_{n\pm1\,j,k\ell}\, p_k\, p_\ell \geqslant T^{(q)}_{n+1\,n-1,nn}\, p_n\, p_n > 0.$$

The first statement follows now by induction.

For the second statement, evaluate the reversibility condition (10) for all combinations of i, j, k, ℓ with $i+j = k+\ell \leqslant 4$. This leads to four independent equations. Three of them can be transformed to the recursion

$$p_k = \frac{(k+1)q}{2(k-1)+2q} \frac{p_1}{p_0}\, p_{k-1}, \qquad k \in \{2,3,4\},$$

from which one derives explicit equations for all p_k with $k \in \{2,3,4\}$ in terms of p_0 and p_1. Inserting the one for p_2 into the remaining equation yields another equation for p_4 in terms of p_0 and p_1, which contradicts the first equation for all $q \in \,]0,1[$, as is easily verified. □

So, non-trivial fixed points for $0 < q < 1$ are not reversible, and thus much more difficult to determine. Our most general result so far is

Theorem 4. *If $\boldsymbol{p}(0) \in P_{\alpha,\delta}$ for some α, δ, then $\boldsymbol{p}(t) \in P_{\alpha,\delta}$ for all times $t \in \mathbb{N}_0$, respectively $t \in \mathbb{R}_{\geqslant 0}$, and $\mathcal{R}_q$ has a fixed point in $P_{\alpha,\delta}$.*

The proof is based on the fact that $\mathcal{R}_q$ is, in a certain sense, monotonic in the parameter q. This is stated in

Proposition 16. *Assume $\boldsymbol{a}(\boldsymbol{p}) \in X_{\alpha,\delta}$ for some α, δ. Then, with respect to the partial order introduced before Proposition 7, $\boldsymbol{a}(\mathcal{R}_q(\boldsymbol{p})) \leqslant \boldsymbol{a}(\mathcal{R}_{q'}(\boldsymbol{p}))$ for all $0 \leqslant q \leqslant q' \leqslant 1$. In particular, $\boldsymbol{a}(\mathcal{R}_q(\boldsymbol{p})) \in X_{\alpha,\delta}$ for all $0 \leqslant q \leqslant 1$.*

To show this, we need three rather technical lemmas. The first one collects formal conditions on the difference of two discrete probability distributions $T^{(q)}_{ij,k\ell}$ with different parameter values (but $j = k + \ell - i$ and the same fixed k, ℓ). These are then verified in our case.

Lemma 10. *Let the numbers $x_i \in \mathbb{R}$ ($0 \leqslant i \leqslant r$ with some $r \in \mathbb{N}_0$) satisfy the following three conditions:*

$$\sum_{i=0}^{r} x_i = 0. \tag{32}$$

$$x_{r-i} = x_i \quad \textit{for all } 0 \leqslant i \leqslant r. \tag{33}$$

$$\textit{There is an integer } n \textit{ such that } \begin{cases} x_i \geqslant 0 \ : \ 0 \leqslant i \leqslant n \\ x_i < 0 \ : \ n < i \leqslant \lfloor \frac{r}{2} \rfloor \end{cases} . \tag{34}$$

Further, let $f_i \in \mathbb{R}$ ($0 \leqslant i \leqslant r$) be given with

$$0 \leqslant f_1 - f_0 \leqslant f_2 - f_1 \leqslant \ldots \leqslant f_r - f_{r-1}. \tag{35}$$

Then, we have

$$\sum_{i=0}^{r} f_i x_i \geqslant 0.$$

Proof. Let us first consider the trivial cases. If $x_i \equiv 0$, everything is clear, so let $x_i \not\equiv 0$. If $r \leqslant 1$ then $x_i \equiv 0$, so let $r \geqslant 2$, and thus $n \leqslant \frac{r}{2} - 1$. Define $x_{\frac{r}{2}} = f_{\frac{r}{2}} = 0$ for odd r. Then, we can write

$$\sum_{i=0}^{r} f_i x_i = \sum_{i=0}^{n} (f_i + f_{r-i})\, x_i + \sum_{i=n+1}^{\lceil \frac{r}{2} \rceil - 1} (f_i + f_{r-i})\, x_i + f_{\frac{r}{2}} x_{\frac{r}{2}}. \tag{36}$$

Furthermore, for $r - i \geqslant i$, due to (35),

$$f_i + f_{r-i} = f_{i-1} + f_{r-i+1} + (f_i - f_{i-1}) - (f_{r-i+1} - f_{r-i}) \leqslant f_{i-1} + f_{r-i+1}.$$

Now, define $C := \sum_{i=0}^{n} x_i = -\sum_{i=n+1}^{\lceil \frac{r}{2} \rceil - 1} x_i - \frac{1}{2} x_{\frac{r}{2}} > 0$, and the claim follows with (36), since $r - n \geqslant n + 1$ by assumption:

$$\begin{aligned} \sum_{i=0}^{r} f_i x_i &\geqslant C\,[f_n + f_{r-n} - f_{n+1} - f_{r-n-1}] \\ &= C\,[(f_{r-n} - f_{r-n-1}) - (f_{n+1} - f_n)] \geqslant 0. \end{aligned}$$

□

Lemma 11. *Let $j \in \mathbb{N}_0$ be fixed and $f_i = (i)_j$, $i \in \mathbb{N}_0$, where $(i)_j$ is the falling factorial, which equals 1 for $j = 0$ and $i(i-1)\cdots(i-j+1)$ for $j > 0$, hence $\frac{i!}{(i-j)!}$ for $i \geqslant j$. Then condition (35) is satisfied.*

Proof. For $j = 0$, condition (35) is trivially true. Otherwise, each f_i is a polynomial of degree j in i with zeros $\{0, 1, \ldots, j-1\}$, hence we have the equality $0 = f_1 - f_0 = \ldots = f_{j-1} - f_{j-2}$. Then, for $i \geqslant j-1$, the polynomial and all its derivatives are increasing functions since $\lim_{i\to\infty} f_i = \infty$. Therefore, for $i \geqslant j-1$, we have $0 \leqslant f_{i+1} - f_i \leqslant f_{i+2} - f_{i+1}$. Hence (35) holds. □

Lemma 12. *For $0 \leqslant q \leqslant q' \leqslant 1$ and all k, ℓ, Equations* (32)–(34) *are true for $r = k + \ell$ and $x_i = T^{(q')}_{ik\ell} - T^{(q)}_{ik\ell}$, where $T^{(q)}_{ik\ell} = T^{(q)}_{ij,k\ell}$ with $j = k + \ell - i$.*

Proof. The validity of (32) and (33) is clear from the normalization (3) and the symmetry of the $T^{(q)}_{ik\ell}$. For (34), let $k \leqslant \ell$ without loss of generality. In the trivial cases $q = q'$ or $k = 0$, choose $n = \lfloor \frac{r}{2} \rfloor$. Otherwise, $x_i = T^{(q')}_{ik\ell} - T^{(q)}_{ik\ell} < 0$ for $k \leqslant i \leqslant \lfloor \frac{r}{2} \rfloor$, since $C^{(q')}_{k\ell} < C^{(q)}_{k\ell}$, and $x_0 > 0$. For $0 \leqslant i \leqslant k$, consider

$$y_i = \frac{x_i}{T^{(q)}_{ik\ell}} + 1 = \frac{C^{(q')}_{k\ell}}{C^{(q)}_{k\ell}} \left(\frac{q'}{q}\right)^{k-i}.$$

Here, the first factor is less than 1, the second is equal to 1 for $k = i$, greater than 1 for $0 \leqslant k < i$, and strictly decreasing with i. Since $x_i \geqslant 0$ if and only if $y_i \geqslant 1$, there is an index n with the properties needed. □

Proof of Proposition 16. We assume $0 \leqslant q \leqslant q' \leqslant 1$. Lemmas 10–12 imply, for all $k, \ell, j \in \mathbb{N}_0$ with $k + \ell \geqslant j$,

$$\sum_{i=j}^{k+\ell} \tfrac{i!}{(i-j)!} T^{(q)}_{ik\ell} \leqslant \sum_{i=j}^{k+\ell} \tfrac{i!}{(i-j)!} T^{(q')}_{ik\ell}.$$

Then, since $T^{(q)}_{ik\ell} = 0$ for $i > k + \ell$,

$$\begin{aligned}
\boldsymbol{a}(\mathcal{R}_q(\boldsymbol{p}))_j &= \tfrac{1}{(j+1)!} \sum_{i\geqslant j} \tfrac{i!}{(i-j)!} \mathcal{R}_q(\boldsymbol{p})_i = \tfrac{1}{(j+1)!} \sum_{i\geqslant j} \tfrac{i!}{(i-j)!} \sum_{k,\ell\geqslant 0} T^{(q)}_{ik\ell} p_k p_\ell \\
&= \tfrac{1}{(j+1)!} \sum_{k,\ell\geqslant 0} p_k p_\ell \sum_{i\geqslant j} \tfrac{i!}{(i-j)!} T^{(q)}_{ik\ell} \\
&\leqslant \tfrac{1}{(j+1)!} \sum_{k,\ell\geqslant 0} p_k p_\ell \sum_{i\geqslant j} \tfrac{i!}{(i-j)!} T^{(q')}_{ik\ell} = \boldsymbol{a}(\mathcal{R}_{q'}(\boldsymbol{p}))_j .
\end{aligned}$$

From this, together with Lemma 8, the claim follows. □

Proof of Theorem 4. According to Proposition 16, $\mathcal{R}_q$ maps $P_{\alpha,\delta}$ into itself, and thus, in discrete time, $\boldsymbol{p}(t) \in P_{\alpha,\delta}$ for every $t \in \mathbb{N}_0$. The analogous statement is true for continuous time $t \in \mathbb{R}_{\geqslant 0}$. To see this, consider $P_{\alpha,\delta}$ as a closed subset of ℓ^1. Recall that $\mathcal{R}_q - \mathbb{1}$ is globally Lipschitz on ℓ^1 by Proposition 1. Moreover, for any $\boldsymbol{p} \in P_{\alpha,\delta}$ and $t \in [0, 1]$, Proposition 7 tells us that

$$\boldsymbol{p} + t(\mathcal{R}_q(\boldsymbol{p}) - \boldsymbol{p}) = (1-t)\boldsymbol{p} + t\mathcal{R}_q(\boldsymbol{p}) \in P_{\alpha,\delta}.$$

This implies the positive invariance of $P_{\alpha,\delta}$ by [14, Thm. VI.2.1] (see also [1, Thm. 16.5]). The existence of a fixed point once again follows from the Leray–Schauder–Tychonov theorem [18, Thm. V.19]. □

On the basis of the above analysis, and further numerical work done to investigate the fixed point properties [17,22], it is plausible that, given the mean copy number m, never more than one fixed point for $\mathcal{R}_q$ exists. Due to the global convergence results at $q = 0$ and $q = 1$, any non-uniqueness in the vicinity of these parameter values could only come from a bifurcation, not from an independent source. Numerical investigations indicate that no bifurcation is present, but this needs to be analyzed further.

Furthermore, the Lipschitz constant for $\tilde{\mathcal{R}}_q$ can be expected to be continuous in the parameter q, hence to remain strictly less than 1 on the sets $X_{\alpha,\delta}$ in a neighborhood of $q = 1$. So, at least locally, the contraction property should be preserved. Nevertheless, we do not expand on this here since it seems possible to use a rather different approach [7], which has been used for similar problems in game theory, to establish a slightly weaker type of convergence result for all $0 < q < 1$, and probably even on the larger compact set $\mathcal{M}^+_{1,m,C}$ of Lemma 2.

7. Concluding remarks

In this article, we have shown that, for the extreme parameter values $q = 0$ (internal UC) and $q = 1$ (random UC), any initial configuration satisfying a specific condition converges to one of the known fixed points, both in discrete and continuous time. The condition to be met is, for $q = 0$, the existence of the r-th moment ($r > 1$, see Theorem 2), respectively, for $q = 1$, that the corresponding generating function has a radius of convergence $\rho > 1$ (Theorem 3). Convergence takes place in the total variation norm in all cases. As argued in the previous section, similar results can be expected for the intermediate parameter values as well.

These results are valid for deterministic dynamics and thus correspond to the case of infinite populations. With respect to biological relevance, however, we add some arguments that it is reasonable to expect this to be a good description for large but finite populations as well, i.e., for the underlying (multitype) branching process. For finite state spaces, such as in the mutation–selection models discussed in [6], the results by Ethier and Kurtz [4, Thm. 11.2.1] and the generalization [2, Thm. V.7.2] of the Kesten–Stigum theorem [10,11] guarantee that in the infinite population limit the relative genotype frequencies of the branching process converge almost surely to the deterministic solution (if the population does not go to extinction). Since for the UC models considered here the equilibrium distributions are exponentially small for large copy numbers (owing to Theorem 4 also for $q \in \,]0, 1[$), one can expect these systems to behave very much like ones with finitely many genotypes. This is also supported by several simulations. Nevertheless, this questions deserves further attention.

Acknowledgements. We thank Ellen Baake, Joachim Hermisson, and Josef Hofbauer for their cooperation, and Odo Diekmann and Anton Wakolbinger for clarifying discussions. O.R. gratefully acknowledges support by a PhD scholarship of the Studienstiftung des deutschen Volkes. The authors express their gratitude to the Erwin Schrödinger International Institute for Mathematical Physics in Vienna for support during a stay in December 2002, where the manuscript for this article was completed.

References

1. Amann, H.: Ordinary Differential Equations. De Gruyter, Berlin, 1990
2. Athreya, K.B., Ney, P.E.: Branching Processes. Springer, New York, 1972
3. Billingsley, P.: Convergence of Probability Measures. 2nd ed., Wiley, New York, 1999
4. Ethier, S.N., Kurtz, T.G.: Markov Processes – Characterization and Convergence. Wiley, New York, 1986
5. Graur, D., Li, W.-H.: Fundamentals of Molecular Evolution. 2nd ed., Sinauer, Sunderland, 2000
6. Hermisson, J., Redner, O., Wagner, H., Baake, E.: Mutation–selection balance: Ancestry, load, and maximum principle. Theor. Pop. Biol. **62**, 9–46 (2002); `cond-mat/0202432`
7. Hofbauer, J.: In preparation
8. Hofbauer, J.: Private communication (2002)
9. Kallenberg, O.: Foundations of Modern Probability. Springer, New York, 1997
10. Kesten, H., Stigum, B.P.: A limit theorem for multidimensional Galton–Watson processes. Ann. Math. Statist. **37**, 1211–1233 (1966)
11. Kurtz, T., Lyons, R., Pemantle, R., Peres, Y.: A conceptual proof of the Kesten–Stigum theorem for multi-type branching processes. In: Athreya, K.B., Jagers, P. (eds.), Classical and Modern Branching Processes, pp. 181–185, Springer, New York, 1997
12. Lang, S.: Real and Functional Analysis. 3rd ed., Springer, New York, 1993
13. Lang, S.: Complex Analysis. 4th ed., Springer, New York, 1999
14. Martin, R.H.: Nonlinear Operators and Differential Equations in Banach Spaces. Wiley, New York, 1976
15. Ohta, T.: On the evolution of multigene families. Theor. Pop. Biol. **23**, 216–240 (1983)
16. Pedersen, G.K.: Analysis Now. Revised ed., Springer, New York, 1995
17. Redner, O.: Models for Mutation, Selection, and Recombination in Infinite Populations. PhD Thesis, Univ. Greifswald (2003). Shaker, Aachen, 2003.
18. Reed, M., Simon, B.: Methods of Modern Mathematical Physics I: Functional Analysis. Academic Press, San Diego, 1980
19. Remmert, R.: Classical Topics in Complex Function Theory. Springer, New York, 1998
20. Rudin, W.: Real and Complex Analysis. 3rd ed., McGraw-Hill, New York, 1986
21. Shiryaev, A.N.: Probability. 2nd ed., Springer, New York, 1996
22. Shpak, M., Atteson, K.: A survey of unequal crossover systems and their mathematical properties. Bull. Math. Biol. **64**, 703–746 (2002)
23. Shpak, M., Wagner, G.P.: Asymmetry of configuration space induced by unequal crossover: Implications for a mathematical theory of evolutionary innovation. Artificial Life **6**, 25–43 (2000)
24. Takahata, N.: A mathematical study on the distribution of the number of repeated genes per chromosome. Genet. Res. **38**, 97–102 (1981)
25. Walsh, J.B.: Persistence of tandem arrays: Implications for satellite and simple-sequence DNA's. Genetics **115**, 553–567 (1987)
26. Walter, W.: Ordinary Differential Equations. Springer, New York, 1998
27. Yosida, K.: Functional Analysis. 6th ed., Springer, Berlin, 1980