Check for updates

J. Appl. Probab. **62,** 1534–1548 (2025)
*doi:*10.1017/jpr.2025.21

THE REPLICATOR COALESCENT

A. E. KYPRIANOU,* *University of Warwick*
L. PEÑALOZA,** *Universidad del Mar*
T. ROGERS,*** *University of Bath*

Abstract

We consider a stochastic model, called the *replicator coalescent*, describing a system of blocks of k different types that undergo pairwise mergers at rates depending on the block types: with rate $C_{ij} \geq 0$ blocks of type i and j merge, resulting in a single block of type i. The replicator coalescent can be seen as a generalisation of Kingman's coalescent death chain in a multi-type setting, although without an underpinning exchangeable partition structure. The name is derived from a remarkable connection between the instantaneous dynamics of this multi-type coalescent when issued from an arbitrarily large number of blocks, and the so-called *replicator equations* from evolutionary game theory. By dilating time arbitrarily close to zero, we see that initially, on coming down from infinity, the replicator coalescent behaves like the solution to a certain replicator equation. Thereafter, stochastic effects are felt and the process evolves more in the spirit of a multi-type death chain.

Keywords: Markov chain; coalescent; coming down from infinity

2010 Mathematics Subject Classification: Primary 60J80
Secondary 60E10

1. Introduction

In this article, we are interested in developing a multi-type analogue of Kingman's coalescent as a death chain, called a *replicator coalescent*, with the following interpretation. Blocks take one of k different types. Mergers within blocks may take place, as well as mergers of blocks with two different types. In the latter case, we will need to specify what type the two merging blocks of different type will take. To this end, let us introduce the $k \times k$ rate matrix $\boldsymbol{C} = (C_{i,j})$, the merger rate matrix, with $C_{i,j} \geq 0$ for all $i, j \in \{1, \ldots, k\}$. This matrix (which is not an intensity matrix) encodes the evolution of a continuous-time Markov chain, say $(\boldsymbol{n}(t), t \geq 0)$, on the state space $\mathbb{N}_*^k = \left\{\boldsymbol{\eta} \in \mathbb{N}_0^k : \ \sum_{i=1}^k \eta_i \geq 1\right\}$, where $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$, which is defined in the following way. Given that $\boldsymbol{n}(t) = (n_1, \ldots, n_k) \in \mathbb{N}_*^k$ such that $\sum_{i=1}^k n_i > 1$:

Received 27 November 2023; accepted 31 January 2025.
* Postal address: Department of Statistics, University of Warwick, Coventry, CV4 7AL, UK. Email: andreas.kyprianou@warwick.ac.uk
** Postal address: Instituto de Investigación de Matemáticas y Actuaría, Universidad del Mar, campus Huatulco. Mexico. Email: lizbethpv@aulavirtual.umar.mx
*** Postal address: Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, UK. Email: ma3tcr@bath.ac.uk

© The Author(s), 2025. Published by Cambridge University Press on behalf of Applied Probability Trust. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

https://doi.org/10.1017/jpr.2025.21 Published online by Cambridge University Press

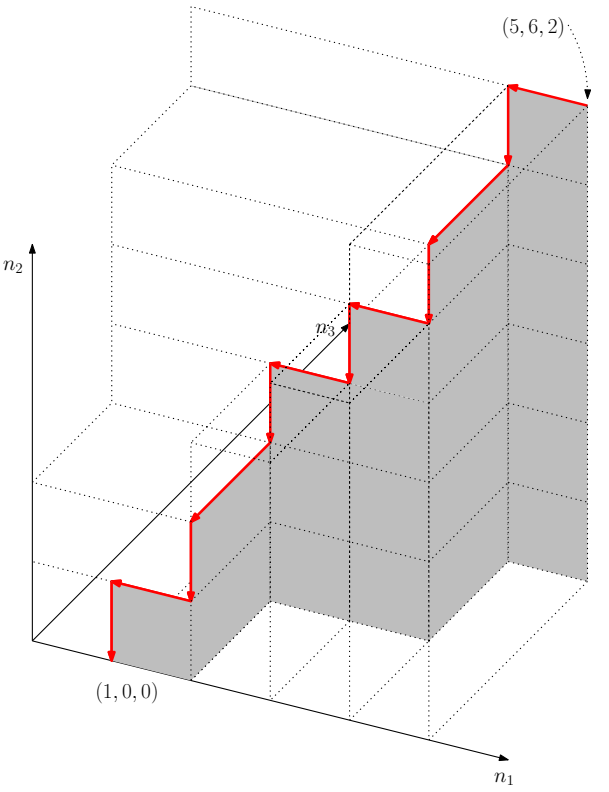


FIGURE 1. A path of replicator coalescent block numbers with $k = 3$, initiated from $\boldsymbol{n}(0) = (5, 6, 2)$ and reducing to a population of one with $\boldsymbol{n}(\gamma_1) = (1, 0, 0)$. The diagram represents the range of the process and there is no time axis.

- For $i \in \{1, \ldots, k\}$, any two specific blocks of type i will merge at rate $C_{i,i}$, and hence the total merger rate of type i blocks is equal to $C_{i,i} \binom{n_i}{2}$.
- For $i \neq j$, both selected from $\{1, \ldots, k\}$, any block of type i will merge with any block of type j, producing a single block of type i, at rate $C_{i,j}$. The total rate of events of this type is thus $C_{i,j} n_i n_j$.

We can interpret $(\boldsymbol{n}(t), t \geq 0)$ as the evolution of the population of k types that exhibit both inter-type and intra-type competition. The rate $C_{i,i}$ is the rate at which two individuals of type i compete for resource, resulting in one of them not surviving. Moreover, at rate $C_{i,j}$ individuals of types i and j encounter one another in a competition for resource, resulting in j not surviving. In this respect, our replicator coalescent echoes features of the so-called 'O.K. Corral' model describing a famous nineteenth-century Arizona shoot-out between lawmen and outlaws in [5, 6, 7], as well as the (birth–)death process in [1]. An example of the sample path of the process $\boldsymbol{n}$ is given in Figure 1 in the setting $k = 3$.

The reader will note that the rate at which inter-type mergers occur is that of Kingman's coalescent. When there is only one type, and hence only inter-type coalescence occurs, the replicator coalescent is therefore nothing more than the death chain of a Kingman coalescent.

https://doi.org/10.1017/jpr.2025.21 Published online by Cambridge University Press

In this sense, the process $(\boldsymbol{n}, \mathbb{P})$ may be thought of as a multi-type variant of the Kingman death chain.

We also note that the requirement $C_{i,j} > 0$ for all i, j is a sufficient condition to ensure that populations of different types are able to interact with one another, which, in turn, will allow for the population to collapse to a single surviving individual. If, for example, we were to take $C_{i,j} = 0$ for all $i \neq j$, then we have k independent Kingman coalescence processes that do not interact, in which case the absorbing state of $(\boldsymbol{n}(t), t \geq 0)$ is the vector $(1, \ldots, 1)$. That said, even sticking to the requirement that the model ensures that the population reduces to a single surviving individual, the condition $C_{i,j} > 0$ for all i, j may well be replaced by a weaker 'irreducibility'-type condition. Nonetheless, we refrain from exploring this further at this point as, later on in this article, we will require more conditions on the matrix $C_{i,j}$, for different reasons, that will supersede the current discussion.

The specific structure of the replicator coalescent does not permit an interpretation in terms of exchangeable partition structures, as is the case for Kingman's coalescent, neither when considering the total population nor when considered as a vector-valued process. In the former case, this is obviously because blocks are subject to different rates according to their type and therefore cannot be exchangeably labelled. In the latter case, a notion of multi-type exchangeability is possible and was discussed in the context of coalescence in [4]. Unfortunately, the way in which mergers occur across different types of blocks is just outside the definition given in [4], which insists on a random selection of multiple mergers that cannot be arranged to be a single merger via parameter choices.

Although the replicator coalescent lives in the space $\mathbb{N}_*^k$, we prefer to describe it via a so-called L^1-polar *decomposition* in the spirit of, e.g., [1]. To this end, define $\sigma(t) = \|\boldsymbol{n}(t)\|_1 = n_1(t) + \cdots + n_k(t) \in \mathbb{N}$ and let $\boldsymbol{r}(t) = \arg(\boldsymbol{n}(t)) := \sigma(t)^{-1}\boldsymbol{n}(t) \in \mathcal{S}^k$, where

$$\mathcal{S}^k := \left\{ (x_1, x_2, \ldots, x_k) \in \mathbb{R}^k : \sum_{i=1}^{k} x_i = 1,\ x_i \geq 0 \text{ for all } i \right\}$$

is the $(k-1)$-dimensional simplex, with vector entries $\boldsymbol{r}_i(t) = \sigma(t)^{-1} n_i(t)$, $i = 1, \ldots, k$. We will additionally writem and occasionally use, $\mathcal{S}_+^k$ to have the same definition as $\mathcal{S}^k$, albeit with each of the $x_i > 0$.

We often refer to the process $\boldsymbol{n}$ as $(\boldsymbol{r}, \sigma^{-1})$. In particular, if $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_k) \in \mathbb{N}_*^k$, we will use $\mathbb{P}_{\boldsymbol{\eta}}$ for the law of the replicator coalescent issued from state $\boldsymbol{n}(0) = \boldsymbol{\eta}$. The usual convention would be to think of the family of probabilities $\mathbb{P} = (\mathbb{P}_{\boldsymbol{\eta}}, \boldsymbol{\eta} \in \mathbb{N}_*^k)$; however, we interchangeably also think of $\mathbb{P} = (\mathbb{P}_{\boldsymbol{\eta}}, \boldsymbol{\eta} \in \mathcal{S}^k \times \mathbb{N}^{-1})$, where $\mathcal{S}^k \times \mathbb{N}^{-1} := \{(\boldsymbol{x}, 1/n) : \boldsymbol{x} \in \mathcal{S}^k \text{ and } n \in \mathbb{N}\}$.

In the setting of the block-counting process for Kingman's coalescent, there are three fundamental facts that are now taken for granted in the mainstream literature. First, Kingman's coalescent block-counting process comes down from infinity almost surely. Second, it comes down from infinity in such a way that the number of blocks divided by $1/t$ converges to a constant as $t \to 0$. Third, and somewhat trivially, the block-counting process is a death chain with an absorbing state that is a single block. This inspires us to address the following questions for our replicator coalescent:

(i) Does it 'come down from infinity' in an appropriately prescribed sense?

(ii) What is the distribution on $\{1, \ldots, k\}$ of the type of the terminal block?

We are interested in characterising the behaviour of the replicator coalescent as we start it from an initial population that 'tends to infinity' in a prescribed way, and as such we will

https://doi.org/10.1017/jpr.2025.21 Published online by Cambridge University Press

give a response to (i). In doing so, we will unravel a remarkable connection with the theory of evolutionary dynamical systems described by so-called *replicator equations*, hence our choice of the name *replicator coalescent*. Our response to (ii) is by no means a complete story. For example, we don't show the existence of entrance laws on Skorokhod space, but rather we focus on the behaviour of the process as we limit its initial state to a boundary state 'at infinity', which means an initial condition for $(\boldsymbol{r}, \sigma^{-1})$ in $\mathcal{S}_+^k \times \{0\}$.

Definition 1. Henceforth we will say that $(\boldsymbol{\eta}^N, N \geq 1)$ tends to $(\boldsymbol{r}_0, 0) \in \mathcal{S}_+^k \times \{0\}$ if $\boldsymbol{\eta}^N \in \mathbb{N}_*^k$ such that $\|\boldsymbol{\eta}^N\| = N$ and $\arg(\boldsymbol{\eta}^N) \to \boldsymbol{r}_0$ as $N \to \infty$.

We are unable to provide any results for (ii) and believe this to be an extremely difficult problem; even in light of related results, e.g. on the aforementioned O.K. Corral model in [5, 6, 7]. This short article is but an introduction to replicator coalescence, offering the opportunity for further analysis to take place. Indeed, in future work we aim to give a more precise statement on the convergence on the Skorokhod space of the process to a unique entrance law that exhibits continuity at time zero. We comment further on this in the final section of this paper.

2. Main results

For our first result, we show that the replicator coalescent comes down from infinity in a relatively specific sense. We study the time $\gamma_m = \inf\{t > 0 \colon \sigma(t) \leq m\}$, $m \in \mathbb{N}$, that the coalescent first reaches a state with m blocks in total, which can be bounded in probability for large N according to the following result.

Lemma 1. *Without any further requirement on $C_{i,j}$, there exists a Kingman coalescent death chain ν^- such that $\sigma \geq \nu^-$ under $\mathbb{P}_{\boldsymbol{\eta}}$ for each $\boldsymbol{\eta} = (\arg(\boldsymbol{n}(0)), \sigma(0)^{-1}) \in \mathcal{S}^k \times \mathbb{N}^{-1}$. Conversely, if $C_{i,j} > 0$ for all i,j, there exists a Kingman coalescent death chain ν^+ such that $\sigma \leq \nu^+$ under $\mathbb{P}_{\boldsymbol{\eta}}$ for each $\boldsymbol{\eta} = (\arg(\boldsymbol{n}(0)), \sigma(0)^{-1}) \in \mathcal{S}^k \times \mathbb{N}^{-1}$.*

In particular, under the assumption that $C_{i,j} > 0$ for all i, j, the replicator coalescent comes down from infinity in the sense that, for all $\varepsilon > 0$, $\lim_{m\to\infty} \lim_{N\to\infty} \mathbb{P}_{\boldsymbol{\eta}^N}(\gamma_m < \varepsilon) = 1$.

As it comes down from infinity, the standard Kingman coalescent with merger rate $c > 0$ has block count $(\nu(t), t \geq 0)$, which is approximately described by the ordinary differential equation (ODE)

$$\dot{\nu}(t) = -c\nu(t)^2/2. \tag{1}$$

It turns out that the corresponding ODE for our coalescent is already known in the evolutionary game theory literature as the *replicator equation*.

Replicator equations are used to describe a population of k types, with the proportion of the total population of type $i \in \{1, \ldots, k\}$ at time $t \geq 0$ denoted $x_i(t) \in [0, 1]$. These values sum to one, so $\boldsymbol{x}(t) := (x_1(t), \ldots, x_k(t))$ lives in $\mathcal{S}^k$. The replicator equations are then written as $\dot{x}_i(t) = x_i(t)(f_i(\boldsymbol{x}(t)) - \bar{f}(\boldsymbol{x}(t)))$, $i = 1, \ldots, k$, $t \geq 0$, where $f_i \colon \mathcal{S}^k \mapsto \mathbb{R}$ describes the 'fitness' of type i as a function of the current population density, and $\bar{f}(\boldsymbol{x}(t)) = \sum_{i=1}^n x_i f_i(\boldsymbol{x}(t))$ is the average population fitness.

Fitness is often assumed to depend linearly upon the population distribution, with coefficients organised in the 'payoff matrix' A. Specifically, let $A_{i,j}$ denote the payoff for a player of type i facing an opponent of type j. Then $f_i(\boldsymbol{x}) = \sum_{j=1}^n A_{i,j}x_j$. This replicator equation, henceforth referred to as the $\boldsymbol{A}$*-replicator equation*, may be written as

$$\dot{x}_i(t) = x_i(t)\big([\boldsymbol{A}\boldsymbol{x}(t)]_i - \boldsymbol{x}(t)^\top \boldsymbol{A}\boldsymbol{x}(t)\big), \quad i = 1, \ldots, k, \ t \geq 0. \tag{2}$$

https://doi.org/10.1017/jpr.2025.21 Published online by Cambridge University Press

If the system in (2) admits a fixed point in the simplex, i.e. $x_i(t) = x_i^*$, $i = 1, \ldots, k$, for some vector $\boldsymbol{x}^* = (x_1^*, \ldots, x_k^*) \in \mathcal{S}^k$, so that $\mathrm{d}\big(\sum_{i=1}^k x_i(t)\big)/\mathrm{d}t = 0$, then we see that, necessarily, $[\boldsymbol{A}\boldsymbol{x}^*]_i = (\boldsymbol{x}^*)^\top \boldsymbol{A}\boldsymbol{x}^*$, $i = 1, \ldots, k$. In turn, this implies that there is a constant $c > 0$ such that $\boldsymbol{A}\boldsymbol{x}^* = c\mathbf{1}$, so $\boldsymbol{x}^* = c\boldsymbol{A}^{-1}\mathbf{1}$, where $\mathbf{1}$ is the vector in $\mathbb{R}^k$ with unit entries. Since $\boldsymbol{x}^* \in \mathcal{S}^k$, it follows that $\mathbf{1}^\top \boldsymbol{x}^* = 1$, and hence $c = (\mathbf{1}^\top A^{-1}\mathbf{1})^{-1}$, thus (2) is solved by $\boldsymbol{x}^* = A^{-1}\mathbf{1}/\mathbf{1}^\top A^{-1}\mathbf{1}$.

If $\boldsymbol{x}^*$ satisfies the relation $(\boldsymbol{x}^*)^\top \boldsymbol{A}\boldsymbol{x} > \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$ for all $\boldsymbol{x} \neq \boldsymbol{x}^*$ in a neighbourhood of $\boldsymbol{x}^*$, then it is called an *evolutionary stable state* (ESS). Theorem 7.2.4 of [3] states that if $\boldsymbol{x}^*$ is an ESS, then

$$\lim_{t\to\infty} \boldsymbol{x}(t) = \boldsymbol{x}^*. \tag{3}$$

The following results give us a remarkable connection between the theory of replicator equations and our coalescent model. To this end we define our A matrix by

$$A_{i,j} = -\left(C_{j,i}1_{j\neq i} + \frac{1}{2}C_{i,i}1_{i=j}\right). \tag{4}$$

For the remainder of the paper we will assume that the rates C are such that (3) holds.

Let us now state the connection between the notion of coming down from infinity for the replicator coalescent and the corresponding replicator equations.

Theorem 1. *Suppose that $\boldsymbol{A}$ is such that (3) holds, and that $(\boldsymbol{\eta}^N, N \geq 1)$ tends to $(\boldsymbol{r}_0, 0)$. Then, for all $T > 0$, $\lim_{N\to\infty} \mathbb{E}_{\boldsymbol{\eta}^N}\big[\sup_{t\leq T} \|\boldsymbol{R}(t) - \boldsymbol{x}(t)\|_1\big] = 0$, $i = 1, \ldots, k$, where $\boldsymbol{R}(t) = \boldsymbol{r}(\tau(t))$, $t \geq 0$, $\boldsymbol{x}(t) = (x_1(t), \ldots, x_k(t))$ solves the $\boldsymbol{A}$-replicator equation with initial condition $\boldsymbol{x}(0) = \boldsymbol{r}_0$, and $\tau(t) = \inf\big\{s > 0 : \int_0^s \sigma(u)\,\mathrm{d}u > t\big\}$, $t \geq 0$. In particular,*

$$\lim_{t\uparrow\infty} \lim_{N\to\infty} \mathbb{E}_{\boldsymbol{\eta}^N}[\|\boldsymbol{R}(t) - \boldsymbol{x}^*\|_1] = 0, \quad i = 1, \ldots, k.$$

In words, Theorem 1 says that by dilating time arbitrarily close to zero, its process $\boldsymbol{r}$ in the simplex behaves deterministically like the solution to an $\boldsymbol{A}$-replicator equation. For a special choice of the matrix $C_{i,j}$, Figure 2 shows simulations of the process $(\boldsymbol{R}(t), t \geq 0)$ that resonate with the statement of Theorem 1. As part of the proof of Lemma 1, we will see that under $\mathbb{P}_{\boldsymbol{\eta}^N}$, the process σ is comparable to a Kingman coalescent with some collision rate, say $c > 0$. Noting that $\int_0^{\tau(t)} \sigma(u)\,\mathrm{d}u = t$, which implies $\dot{\tau}(t) = 1/\sigma(\tau(t))$, if, in heuristic terms, we treat σ as a solution to (1) with $\sigma(0) = N$, so that $\sigma(t)^{-1} = N^{-1} + ct/2$, then $\tau(t) \approx 2(\exp\{ct/2\} - 1)/cN \approx t/N$.

For a Kingman coalescent, say $(\nu(t), t \geq 0)$ with merger rate c, classical reasoning tells us that $(\nu(t/N)/N, t \geq 0)$ converges in an appropriate sense to the solution to the ODE (1) as $\nu(0) = N \to \infty$. In the spirit of these arguments, we would therefore expect that we can similarly control $(\sigma(t/N)/N, t \geq 0)$ as well as $(\boldsymbol{n}(t/N)/N, t \geq 0)$ as $N \to \infty$. In turn, since we can write

$$\boldsymbol{R}(t) \approx \boldsymbol{r}(t/N) = \frac{\boldsymbol{n}(t/N)/N}{\sigma(t/N)/N}, \quad t \geq 0,$$

we can therefore think of Theorem 1 as a version of the aforementioned functional scaling result for Kingman's coalescent.

The remainder of this paper is structured as follows. In the next section we discuss how we can compare the process $(\sigma(t), t \geq 0)$ with Kingman's coalescent on the same probability space when it is issued from a finite number of blocks. This comparison is used frequently in several of our proofs. In Section 4 we treat the Markov process $\boldsymbol{n}$ as a semimartingale and study its

https://doi.org/10.1017/jpr.2025.21 Published online by Cambridge University Press

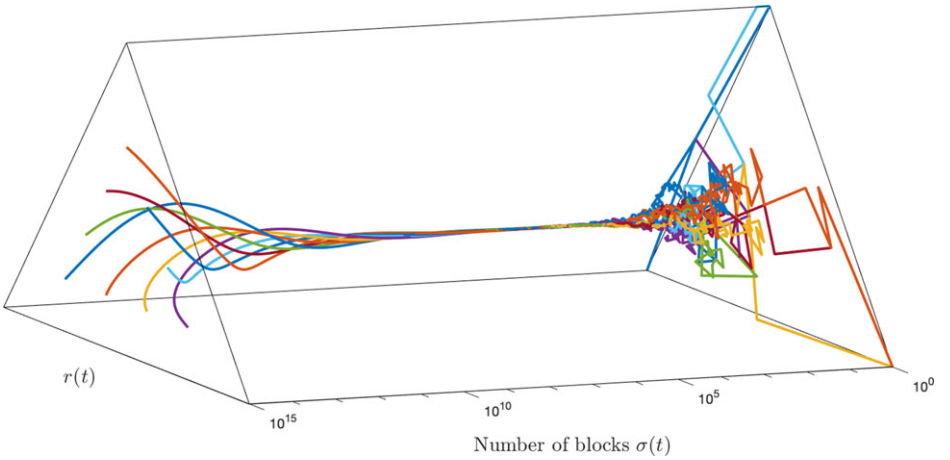


FIGURE 2. Simulations of a replicator coalescent with $k = 3$ initiated from a variety of initial states with an initial number of blocks $\sigma(0) = 10^{15}$. Each path represents a simulation from a different initial state, presented in barycentric coordinates in the 3-simplex and a logarithmic axis for the total number of blocks. The matrix C has entries $C_{i,i} = C_{i,i+1} = 1$ and other entries zero. The reader will note that this case in particular demonstrates that we clearly do not need to enforce $C_{i,j} > 0$ for all i, j.

decomposition as the sum of a martingale and a bounded variation compensator. This provides the basis for the proof of Theorem 1, given in Section 5. Finally, in Section 6 we conclude with some technical remarks and two conjectures concerning further behaviour of the entrance law.

3. Stochastic comparison with Kingman's coalescent

As previously alluded to, there are various points in our reasoning where we will compare the number of blocks in a replicator coalescent with the number of blocks in an appropriately formulated Kingman coalescent on the same probability space. The first such result gives us the proof that the replicator coalescent comes down from infinity.

Proof of Lemma 1. The process σ decreases by one with each block merger, analogously to the block-counting process of a standard Kingman coalescent. It is therefore sufficient to prove that σ decreases at least as slowly, or at least as fast, as a Kingman block-counting process with comparable rates.

Formally, from finite starting states we want to construct such a Kingman coalescent ν^+ on the same space as σ with $\sigma \leq \nu^+$ by considering the minimal rate of σ.

The total rate of mergers in state $\boldsymbol{n}$ is given by

$$\rho(\boldsymbol{n}) = \sum_{i=1}^{k} \left(\sum_{j \neq i} C_{i,j} n_j n_i + C_{i,i} \binom{n_i}{2} \right),$$

which depends not just on the total number of blocks, but also on the distribution of block types. However, since $C_{i,j} \geq 0$ for all i, j, we can choose $C > \max_{i,j} C_{i,j}$ such that

$$\rho(\boldsymbol{n}) < \sum_{i=1}^{k} \frac{1}{2} C n_i \left(\sum_{j \neq i} n_j + n_i - 1 \right) = \sum_{i=1}^{k} \frac{1}{2} C n_i (\|\boldsymbol{n}\|_1 - 1) = C \binom{\|\boldsymbol{n}\|_1}{2}. \tag{5}$$

https://doi.org/10.1017/jpr.2025.21 Published online by Cambridge University Press

Appealing to the skip-free property, it follows that we can stochastically couple a Kingman coalescent death chain $(\nu^-(t), t \geq 0)$ with collision rate C and the process $(\boldsymbol{n}(t), t \geq 0)$ on the same space such that, with $\boldsymbol{\eta} = (\arg(\boldsymbol{n}(0)), \sigma(0)^{-1}) \in \mathcal{S}^k \times \mathbb{N}^{-1}$ and $\nu^-(0) = \sigma(0)$, $\sigma(t) \geq \nu^-(t)$, $t \geq 0$.

Conversely, when $C_{i,j} > 0$ for all i, j, writing $\underline{C} = \min_{i,j} C_{i,j}$, we get

$$\rho(\boldsymbol{n}) \geq \sum_{i=1}^{k} \frac{1}{2} \underline{C} n_i \left(\sum_{j \neq i} n_j + n_i - 1 \right) = \underline{C} \binom{\|\boldsymbol{n}\|_1}{2}.$$

In the same spirit, it follows that we can stochastically embed another Kingman death chain $(\nu^+(t), t \geq 0)$ and the process $(\boldsymbol{n}(t), t \geq 0)$ on the same space such that, with $\boldsymbol{\eta} = (\arg(\boldsymbol{n}(0)), \sigma(0)^{-1}) \in \mathcal{S}^k \times \mathbb{N}^{-1}$ and $\nu^-(0) = \sigma(0)$, $\sigma(t) \leq \nu^+(t)$, $t \geq 0$.

In particular, as ν^+ comes down from infinity, since $\gamma_m \leq \beta_m^+ := \inf\{t > 0 \colon \nu^+(t) = m\}$, it follows that, for each $\varepsilon > 0$,

$$\lim_{m \to \infty} \lim_{N \to \infty} \mathbb{P}_{\boldsymbol{\eta}^N}[\gamma_m < \varepsilon] \geq \lim_{m \to \infty} \lim_{N \to \infty} \mathbb{P}[\beta_m^+ < \varepsilon \mid \nu^+(0) = N] = 1,$$

thus completing the proof. □

4. Semimartingale representation

We would like to treat the replicator coalescent $(\boldsymbol{n}(t), t \geq 0)$ as a semimartingale. It turns out to be more convenient to consider instead the vectorial process

$$\boldsymbol{y}(t) = \begin{bmatrix} \boldsymbol{r}(t \wedge \gamma_1) \\ 1/\sigma(t \wedge \gamma_1) \end{bmatrix}, \quad t \geq 0,$$

where $\gamma_1 = \inf\{t > 0 \colon \sigma(t) = 1\}$. Naturally, by expressing the evolution of $(\boldsymbol{y}(t), t \geq 0)$ as that of a semimartingale, our interest is predominantly in the process $(\boldsymbol{r}(t), t \geq 0)$ in order to make a link with the replicator equations in (2).

Lemma 2. *For each $\boldsymbol{\eta} \in \mathbb{N}_*^k$, the process $\boldsymbol{y}$ under $\mathbb{P}_{\boldsymbol{\eta}}$ has a semimartingale decomposition $\boldsymbol{y}(t) = \boldsymbol{y}(0) + \boldsymbol{m}(t) + \boldsymbol{\alpha}(t)$, $t \geq 0$, where $(\boldsymbol{m}(t), t \geq 0)$ is a martingale taking the form $\boldsymbol{m}(t) = \sum_{s \leq t \wedge \gamma_1} \Delta \boldsymbol{y}(s) - \boldsymbol{\alpha}(t)$, $t \geq 0$, such that $\Delta \boldsymbol{y}(t) = \boldsymbol{y}(t) - \boldsymbol{y}(t-)$ and $(\boldsymbol{\alpha}(t), t \geq 0)$ is a compensator taking the form*

$$\boldsymbol{\alpha}(t) = \int_0^{t \wedge \gamma_1} \frac{\sigma(s)}{\sigma(s) - 1} \sum_{i=1}^{k} \begin{bmatrix} \sigma(s)(\boldsymbol{r}(s) - \boldsymbol{e}_i) \\ 1 \end{bmatrix} r_i(s) [\sigma(s)^{-1} \mathrm{diag}(\boldsymbol{A})\mathbf{1} - \boldsymbol{A}\boldsymbol{r}(s)]_i \, \mathrm{d}s.$$

Proof. A standard computation using the compensation formula tells us that $\boldsymbol{m}$ is a martingale provided that $\sum_{s \leq t} \|\Delta \boldsymbol{m}(s)\|_1 = \sum_{s \leq t} \|\Delta \boldsymbol{y}(s)\|_1$ has finite expectation for each $t \geq 0$, which is equivalent to the existence of the compensator $\boldsymbol{\alpha}(t)$. The latter is given by the rates that define the replicator coalescent. More precisely, recalling (4),

https://doi.org/10.1017/jpr.2025.21 Published online by Cambridge University Press

$$\begin{aligned}\boldsymbol{\alpha}(t) &= \int_0^{t\wedge\gamma_1} \sum_{i=1}^{k} \begin{bmatrix} \dfrac{\boldsymbol{n}(s)-\boldsymbol{e_i}}{\sigma(s)-1} - \dfrac{\boldsymbol{n}(s)}{\sigma(s)} \\ \dfrac{1}{\sigma(s)-1} - \dfrac{1}{\sigma(s)} \end{bmatrix} \left[\sum_{j=1,\, j\neq i}^{k} n_j(s)n_i(s)C_{ji} + \frac{1}{2}n_i(s)(n_i(s)-1)C_{i,i} \right] \mathrm{d}s \\ &= \int_0^{t\wedge\gamma_1} \sum_{i=1}^{k} \begin{bmatrix} \dfrac{\boldsymbol{n}(s)-\boldsymbol{e_i}\sigma(s)}{(\sigma(s)-1)\sigma(s)} \\ \dfrac{1}{(\sigma(s)-1)\sigma(s)} \end{bmatrix} \left[n_i(s)A_{i,i} - \sum_{j=1}^{k} n_j(s)n_i(s)A_{i,j} \right] \mathrm{d}s \\ &= \int_0^{t\wedge\gamma_1} \frac{\sigma(s)}{\sigma(s)-1} \sum_{i=1}^{k} \begin{bmatrix} \sigma(s)(\boldsymbol{r}(s)-\boldsymbol{e_i}) \\ 1 \end{bmatrix} r_i(s)[\sigma(s)^{-1}\mathrm{diag}(\boldsymbol{A})\mathbf{1} - \boldsymbol{A}\boldsymbol{r}(s)]_i \, \mathrm{d}s,\end{aligned}$$

as required. Note that if we identify the compensator via the density $(\boldsymbol{\lambda}(t),\, t \geq 0)$, where

$$\boldsymbol{\alpha}(t) =: \int_0^{t\wedge\gamma_1} \boldsymbol{\lambda}(s)\, \mathrm{d}s \tag{6}$$

then, in the above representation of $\boldsymbol{\alpha}$, the largest term is of order $\sigma(s)$, from which, because the process $\boldsymbol{n}$ is non-increasing, we can easily conclude that, for all $\boldsymbol{\eta} \in \mathbb{N}_*^k$ and any time $t \geq 0$,

$$\mathbb{E}_{\boldsymbol{\eta}^N}\left[\sum_{s\leq t\wedge\gamma_1} \|\Delta \boldsymbol{y}(s)\|_1 \right] \leq \mathbb{E}_{\boldsymbol{\eta}}\left[\int_0^{t\wedge\gamma_1} \|\boldsymbol{\lambda}(s)\|_1 \, \mathrm{d}s \right] \leq C\|\boldsymbol{\eta}\|_1 \mathbb{E}_{\boldsymbol{\eta}^N}[t\wedge\gamma_1] \leq C\|\boldsymbol{\eta}\|_2 t \tag{7}$$

for an unimportant constant $C > 0$. This ensures that $\boldsymbol{m}$ is a martingale and that $\boldsymbol{\alpha}$ is well defined. □

For the proof of Theorem 1, we are interested in the behaviour of the process $\boldsymbol{r}$ under $\mathbb{P}_{\boldsymbol{\eta}}$ for any $\boldsymbol{\eta}$ such that $\|\boldsymbol{\eta}\|_1 \to \infty$ and $\arg(\boldsymbol{\eta}) \to \boldsymbol{r}$ for some $\boldsymbol{r} \in \mathcal{S}_+^k$. Heuristically speaking, the term $\sigma(s)(\boldsymbol{r}(s) - \boldsymbol{e_i})$ in the expression for $\boldsymbol{\alpha}$ suggests that $\alpha(t)$ explodes as $t \to 0$. The undesirable factor $\sigma(s)$ can be removed, however, by an appropriate time change; in doing so, we begin to see where the relationship with the replicator equations emerges.

Lemma 3. *Suppose we define the sequence of stopping times* $(\tau(t),\, t \geq 0)$*, which are defined by the right inverse,* $\tau(t) = \inf\{u > 0 \colon \int_0^u \sigma(s)\, \mathrm{d}s > t\}$*,* $t \geq 0$*. Then* $\boldsymbol{y}^\tau := \boldsymbol{y} \circ \tau$ *has the semi-martingale decomposition* $\boldsymbol{y}^\tau = \boldsymbol{m}^\tau + \boldsymbol{\alpha}^\tau$*, where* $\boldsymbol{m}^\tau := \boldsymbol{m} \circ \tau$ *is a martingale and, for* $t \geq 0$*,*

$$\begin{aligned}\boldsymbol{\alpha}^\tau(t) = \int_0^{t\wedge\tau^{-1}(\gamma_1)} & \frac{\sigma(\tau(s))}{\sigma(\tau(s))-1} \\ & \times \sum_{i=1}^{k} \begin{bmatrix} \boldsymbol{r}(\tau(s)) - \boldsymbol{e_i} \\ \dfrac{1}{\sigma(\tau(s))} \end{bmatrix} r_i(\tau(s))[\sigma(\tau(s))^{-1}\mathrm{diag}(\boldsymbol{A})\mathbf{1} - \boldsymbol{A}\boldsymbol{r}(\tau(s))]_i \, \mathrm{d}s.\end{aligned} \tag{8}$$

Proof. We use basic Stieltjes calculus to tell us that $\mathrm{d}\boldsymbol{\alpha}^\tau(t) = \mathrm{d}\boldsymbol{\alpha}(s)|_{s=\tau(t)}\, \mathrm{d}\tau(t)$. Moreover,

$$\int_0^{\tau(t)} \sigma(s)\, \mathrm{d}s = t, \quad \text{and hence } \sigma(\tau(t))\, \mathrm{d}\tau(t) = \mathrm{d}t. \tag{9}$$

Combining these observations with the conclusion of Lemma 2, the result follows. In particular, from (6),

$$\int_0^{\tau(t)\wedge\gamma_1} \boldsymbol{\lambda}(s)\, \mathrm{d}s = \int_0^{t\wedge\tau^{-1}(\gamma_1)} \frac{\boldsymbol{\lambda}(\tau(u))}{\sigma(\tau(u))}\, \mathrm{d}u. \tag{10}$$

https://doi.org/10.1017/jpr.2025.21 Published online by Cambridge University Press

Technically, we need to verify that $\boldsymbol{m}^\tau$ is a martingale rather than a local martingale; however, a computation similar to (7) taking advantage of (8) is easily executed, affirming the required martingale status. □

Next we look at how to control the second moment of the martingale $\boldsymbol{m}^\tau$ in the semimartingale decomposition of $\boldsymbol{y}^\tau$.

Lemma 4. *For each $\boldsymbol{\eta} \in \mathbb{N}_*^k$, the martingale $\boldsymbol{m}^\tau$ under $\mathbb{P}_{\boldsymbol{\eta}}$ satisfies*

$$\mathbb{E}_{\boldsymbol{\eta}^N}\big[\|\boldsymbol{m}^\tau(t)\|_2^2\big] \le C\mathbb{E}_{\boldsymbol{\eta}}\bigg[\int_0^{\tau(t)\wedge\gamma_1} \frac{\sigma(s)^2}{(\sigma(s)-1)^2}\,\mathrm{d}s\bigg], \quad t \ge 0.$$

Proof. Steiltjes calculus, or equivalently general semi-martingale calculus (see, for example, [8, Theorem II.33]), tells us that, since $\boldsymbol{m}$ has bounded variation,

$$\begin{aligned}\|\boldsymbol{m}(t)\|_2^2 &= 2\int_0^{t\wedge\gamma_1} \boldsymbol{m}(s-)\cdot \mathrm{d}\boldsymbol{m}(s) + \sum_{0<s\le t\wedge\gamma_1} \big\{\Delta\|\boldsymbol{m}(s)\|_2^2 - 2\boldsymbol{m}(s-)\cdot\Delta\boldsymbol{m}(s)\big\}\\ &= 2\int_0^{t\wedge\gamma_1} \boldsymbol{m}(s-)\cdot \mathrm{d}\boldsymbol{m}(s) + \sum_{0<s\le t\wedge\gamma_1} (\Delta\boldsymbol{m}(s))^2.\end{aligned}$$

As all vector entries are bounded, it is easy to show that $\int_0^{t\wedge\gamma_1} \boldsymbol{m}(s-)\cdot \mathrm{d}\boldsymbol{m}(s)$, $t \ge 0$, is a martingale.

Next, we identify the adapted increasing bounded variation process, say $\boldsymbol{\beta}$, that is the compensator of $\sum_{0<s\le t\wedge\gamma_1}(\Delta\boldsymbol{m}(s))^2$, $t \ge 0$, so that $\|\boldsymbol{m}(t)\|_2^2 - \boldsymbol{\beta}(t)$, $t \ge 0$, is a martingale with mean 0. To this end, note that $\Delta\boldsymbol{m}(t) = \Delta\boldsymbol{y}(t)$. Hence, we have, on the event that t is a time at which the number of blocks of type i decreases, that $(\Delta\boldsymbol{m}(t))^2$ is given by

$$\chi_i(t) := \begin{bmatrix} \dfrac{\boldsymbol{n}(t)-\sigma(t)\boldsymbol{e_i}}{(\sigma(t)-1)\sigma(t)} \\ \dfrac{1}{(\sigma(t)-1)\sigma(t)} \end{bmatrix} \cdot \begin{bmatrix} \dfrac{\boldsymbol{n}(t)-\sigma(t)\boldsymbol{e_i}}{(\sigma(t)-1)\sigma(t)} \\ \dfrac{1}{(\sigma(t)-1)\sigma(t)} \end{bmatrix} = \frac{\boldsymbol{n}(t)\cdot\boldsymbol{n}(t) - 2\sigma(t)n_i(t) + \sigma(t)^2 + 1}{\sigma(t)^2(\sigma(t)-1)^2}.$$

It is now straightforward to see that there exists a $C > 0$ such that

$$\begin{aligned}\boldsymbol{\beta}(t) &= \int_0^{t\wedge\gamma_1} \sum_{i=1}^k \chi_i(s)\bigg[\sum_{j=1,\,j\ne i}^k n_j(s)n_i(s)C_{ji} + \frac{1}{2}n_i(s)(n_i(s)-1)C_{i,i}\bigg]\mathrm{d}s\\ &\le C\int_0^{t\wedge\gamma_1} \frac{\sigma(s)^2}{(\sigma(s)-1)^2}\,\mathrm{d}s,\end{aligned}$$

where we have used that $\sigma(t)^2 = (n_1(t)+\cdots+n_k(t))^2 \ge \boldsymbol{n}(t)\cdot\boldsymbol{n}(t)$. Replacing t by $\tau(t)$ and taking expectation gives the desired inequality. □

As the reader may now expect, our ultimate objective is to show that for any sequence of starting initial configurations $\boldsymbol{\eta}^N \in \mathbb{N}_*^k$ such that $\|\boldsymbol{\eta}^N\|_1 \to \infty$ as $N \to \infty$ and $\arg(\boldsymbol{\eta}^N) \to \boldsymbol{r} \in \mathcal{S}_+^k$, the martingale component $\boldsymbol{m}^\tau$ disappears. This tells us that the behaviour of $(\boldsymbol{r}(t), t \ge 0)$ behaves increasingly like the compensator term, which is a further key to controlling its behaviour. To this end, we conclude this section with two more results that provide us with the desired control of the aforesaid martingale component.

Lemma 5. *Fix $t > 0$ and suppose that $(\boldsymbol{\eta}^N, N \ge 1)$ tends to $(\boldsymbol{r}_0, 0)$. Then $\tau^{-1}(t) = \int_0^t \sigma(s)\,\mathrm{d}s \to \infty$, $\tau(t) \to 0$, and $|\tau(t)\wedge\gamma_1 - \tau(t)| \to 0$ weakly as $N \to \infty$.*

https://doi.org/10.1017/jpr.2025.21 Published online by Cambridge University Press

Proof. Recall from the proof of Lemma 1, and specifically (5), that there is a death chain $(\nu(t), t \geq 0)$ representing the number of blocks in a Kingman coalescent with merger rate C such that, for any $\boldsymbol{\eta} \in \mathbb{N}_*^k$, on the same probability space, we can stochastically bound $\sigma(t) \geq \nu(t)$, $t \geq 0$.

We now note that for any large $M > 0$ there exists a constant $C > 0$ (not necessarily the same as before) such that, for any m sufficiently large,

$$\begin{aligned}\lim_{N\to\infty} \mathbb{P}_{\boldsymbol{\eta}^N}\left(\int_0^{\gamma_m} \sigma(s)\,\mathrm{d}s > M\right) &\geq \lim_{N\to\infty} \mathbb{P}_{\boldsymbol{\eta}^N}\left(\int_0^{\beta_m} \nu(s)\,\mathrm{d}s > M\right)\\ &= \Pr\left(\sum_{n=m+1}^{\infty} \frac{n}{C\binom{n}{2}}\mathbf{e}_1^{(n)} > M\right)\\ &= \Pr\left(\sum_{n=m+1}^{\infty} \frac{1}{n-1}\mathbf{e}_1^{(n)} > \frac{CM}{2}\right), \end{aligned} \tag{11}$$

where $\beta_m = \inf\{t > 0\colon \nu(t) = m\}$ and $(\mathbf{e}_1^n, n \geq 1)$ is a sequence of independent and identically distributed unit-mean exponentially distributed random variables. If we write $(N(t), t \geq 0)$ for a Poisson process with unit rate, then almost surely we have

$$\sum_{n=m+1}^{\infty} \frac{1}{n-1}\mathbf{e}_1^{(n)} = \int_0^{\infty} \frac{1}{N(s)+m}\,\mathrm{d}s = \int_0^{\infty} \frac{s}{N(s)+m}\,\frac{\mathrm{d}s}{s} = \infty,$$

where the final equality follows by the strong law of large numbers for Poisson processes. As such, the right-hand side of (11) is equal to 1.

Since M and m can be arbitrarily large, this shows the first claim as soon as we note that $\tau^{-1}(t) = \int_0^t \sigma(u)\,\mathrm{d}u$, which is an easy consequence of (9). On the other hand, note that since $\int_0^{\tau(t)} \sigma(s)\,\mathrm{d}s = t$ when $\sigma(0) = \boldsymbol{\eta}^N \to \infty$, the above comparison with Kingman's coalescent shows almost surely that, since $\int_0^u \sigma(s)\,\mathrm{d}s$ converges weakly to infinity for all $u > 0$, then $\tau(t)$ converges weakly to 0. Indeed, if with positive probability $\tau(t) > \varepsilon$ in the limit as $N \to \infty$, then, on that event, $\int_0^{\tau(t)} \sigma(s)\,\mathrm{d}s \geq \int_0^{\varepsilon} \sigma(s)\,\mathrm{d}s$, which explodes in distribution. This in turn contradicts the definition of $\tau(t)$. This proves the second and third statements of the lemma. □

Since the second moment of the martingale $\boldsymbol{m}^\tau$ can be controlled by its associated time change, we also get a helpful L^2 corollary from Lemma 5.

Corollary 1. *From Lemmas 4 and 5, we deduce that, if* $(\boldsymbol{\eta}^N, N \geq 1)$ *tends to* $(\mathbf{r}_0, 0)$*, then, for each* $t > 0$*,* $\lim_{N\to\infty} \mathbb{E}_{\boldsymbol{\eta}^N}\left[\sup_{s\leq t} \|\boldsymbol{m}^\tau(s)\|_2^2\right] = 0$.

Proof. From Lemma 4 and a change of variable similar to (10),

$$\begin{aligned}\mathbb{E}_{\boldsymbol{\eta}^N}\left[\|\boldsymbol{m}^\tau(t)\|_2^2\right] &\leq C\mathbb{E}_{\boldsymbol{\eta}^N}\left[\int_0^{\tau(t)\wedge\gamma_1} \frac{\sigma(s)^2}{(\sigma(s)-1)^2}\,\mathrm{d}s\right]\\ &= C\mathbb{E}_{\boldsymbol{\eta}^N}\left[\int_0^{t\wedge\tau^{-1}(\gamma_1)} \frac{\sigma(\tau(u))^2}{(\sigma(\tau(u))-1)^2}\frac{1}{\sigma(\tau(u))}\,\mathrm{d}u\right] \leq Ct\mathbb{E}_{\boldsymbol{\eta}^N}\left[\frac{1}{\sigma(\tau(t))}\right].\end{aligned}$$

We can choose N sufficiently large that $\tau(t)$ is less than δ with probability at least $1-\varepsilon$, so

$$\mathbb{E}_{\boldsymbol{\eta}^N}\left[\frac{1}{\sigma(\tau(t))}\right] \leq \mathbb{E}_{\boldsymbol{\eta}^N}\left[\frac{1}{\sigma(\delta)}; \tau(t) < \delta\right] + \mathbb{P}_{\boldsymbol{\eta}^N}(\tau(t) \geq \delta) \leq \delta\mathbb{E}_{\boldsymbol{\eta}^N}\left[\frac{1}{\delta\nu(\delta)}\right] + \varepsilon,$$

https://doi.org/10.1017/jpr.2025.21 Published online by Cambridge University Press

where we have again compared with a lower bounding Kingman coalescent $(\nu(t), t \geq 0)$ on the same space, as in Lemma 5. Recall the classical result for Kingman's coalescent coming down from infinity that, when the collision rate is $C > 0$, $\delta\nu(\delta) \to 2/C$ almost surely as $\delta \to 0$ [2]. We can now easily conclude with the help of dominated convergence that $\lim_{N\to\infty} \mathbb{E}_{\boldsymbol{\eta}^N}\big[\|\boldsymbol{m}^\tau(t)\|_2^2\big] = 0$, and this concludes the proof once we invoke Doob's submartingale inequality. □

5. Proof of Theorem 1

Recall that we have required from the $\boldsymbol{A}$-replicator equation that $\boldsymbol{x}(t) \to \boldsymbol{x}^*$ holds. Reinterpreting (2) in its integral form, this tells us that

$$x_i(t) = x_i(0) + \int_0^t x_i(s)([\boldsymbol{A}\boldsymbol{x}(s)]_i - \boldsymbol{x}(s)^\top \boldsymbol{A}\boldsymbol{x}(s))\,\mathrm{d}s, \quad t \geq 0. \tag{12}$$

This representation makes it easier to give the heuristic basis of the proof of Theorem 1.

Following our earlier heuristic reasoning, we can now see that, under $\mathbb{P}_{\boldsymbol{\eta}}$ as $\|\boldsymbol{\eta}\|_1 \to \infty$ the integrand in the expression for $\boldsymbol{\alpha}^\tau$ appears to have a similar structure to the replicator equations (2). That is, under $\mathbb{P}_{\boldsymbol{\eta}}$ as $\|\boldsymbol{\eta}\|_1 \to \infty$ and as $t \to 0$,

$$\frac{\mathrm{d}\boldsymbol{\alpha}^\tau(t)}{\mathrm{d}t} \approx \begin{bmatrix} \boldsymbol{\theta}(t) \\ 0 \end{bmatrix},$$

where $\boldsymbol{\theta}_i(t) = r_i(\tau(t))\big([\boldsymbol{A}\boldsymbol{r}(\tau(t))]_i - \boldsymbol{r}(\tau(t))^\top \boldsymbol{A}\boldsymbol{r}(\tau(t))\big)$. On the other hand, if we can show that $\boldsymbol{m}^\tau \to 0$ as $\|\boldsymbol{\eta}\|_1 \to \infty$, then, since $\boldsymbol{y}^\tau = \boldsymbol{m}^\tau + \boldsymbol{\alpha}^\tau$, reading off the first component of $\boldsymbol{y}^\tau$, i.e. $\boldsymbol{R}(t) := \boldsymbol{r}(\tau(t))$, roughly speaking we see that

$$R_i(t) \approx r_i(0) + \int_0^t R_i(s)\big([\boldsymbol{A}\boldsymbol{R}(s)]_i - \boldsymbol{R}(s)^\top \boldsymbol{A}\boldsymbol{R}(s)\big)\,\mathrm{d}s$$

as $\|\boldsymbol{\eta}\|_1 \to \infty$, given that Corollary 1 shows the martingale component is negligible. In other words, the process $(\boldsymbol{R}(t), t \geq 0)$, begins to resemble the replicator equation in its integral form (12). It now looks like a reasonable conjecture that $\boldsymbol{R}(t) \to \boldsymbol{x}^*$, just as the solution to the replicator equation does.

Let us thus move to the proof of Theorem 1, which, as alluded to earlier, boils down to the control we have on the martingale $\boldsymbol{m}^\tau$ under $\mathbb{P}_{\boldsymbol{\eta}}$ as $\|\boldsymbol{\eta}\|_1 \to \infty$, thanks to Corollary 1.

Proof of Theorem 1. Write $\boldsymbol{R}_i(t) = \boldsymbol{r}(\tau(t))$ on the event $A_t := \{t < \tau^{-1}(\gamma_1)\}$, $t \geq 0$, and note from Lemma 5 and the proof of Lemma 1 that $\lim_{N\to\infty} \mathbb{P}_{\boldsymbol{\eta}^N}(A_t) = 1$. We have, for each $T > 0$,

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\eta}^N}\left[\sup_{t\leq T}\left\|\boldsymbol{R}(t) - \boldsymbol{r}(0) - \int_0^t \left(\sum_{i=1}^k \boldsymbol{e}_i R_i(s)[\boldsymbol{A}\boldsymbol{R}(s)]_i - (\boldsymbol{R}(s)^\top \boldsymbol{A}\boldsymbol{R}(s))\boldsymbol{R}(s)\right)\mathrm{d}s\right\|_1 1_{A_t}\right] \\
&\leq \mathbb{E}_{\boldsymbol{\eta}^N}\left[\sup_{t\leq T}\left\|\boldsymbol{R}(t) - \boldsymbol{r}(0)\right.\right. \\
&\qquad \left.\left. - \int_0^t \frac{\sigma(\tau(s))}{\sigma(\tau(s)) - 1}\sum_{i=1}^k (\boldsymbol{R}(s) - \boldsymbol{e}_i)R_i(s)[\sigma(\tau(s))^{-1}\mathrm{diag}(\boldsymbol{A})\mathbf{1} - \boldsymbol{A}\boldsymbol{R}(s)]_i\,\mathrm{d}s\right\|_1 1_{A_t}\right]
\end{aligned}$$

https://doi.org/10.1017/jpr.2025.21 Published online by Cambridge University Press

$$+\mathbb{E}_{\boldsymbol{\eta}^N}\left[\sup_{t\le T}\left\|\int_0^t \frac{1}{\sigma(\tau(s))-1}\left(\sum_{i=1}^k \boldsymbol{e}_i R_i(s)[\boldsymbol{A}\boldsymbol{R}(s)]_i-(\boldsymbol{R}(s)^\top \boldsymbol{A}\boldsymbol{R}(s))\boldsymbol{R}(s)\right)\mathrm{d}s\right\|_1 1_{A_t}\right]$$
$$+\mathbb{E}_{\boldsymbol{\eta}^N}\left[\sup_{t\le T}\left\|\int_0^t \frac{1}{\sigma(\tau(s))-1}\left((\boldsymbol{R}(s)^\top \mathrm{diag}(\boldsymbol{A})\mathbf{1})\boldsymbol{R}(s)-\sum_{i=1}^k \boldsymbol{e}_i R_i(s)[\mathrm{diag}(\boldsymbol{A})\mathbf{1}]_i\right)\mathrm{d}s\right\|_1 1_{A_t}\right].$$

From Corollary 1, and the fact that $\|\boldsymbol{y}\|_1 \le \sqrt{k}\|\boldsymbol{y}\|_2$ for $\boldsymbol{y}\in\mathbb{R}^k$, the first term after the inequality tends to zero as $N\to\infty$. Up to a multiplicative constant, the second and third terms after the inequality can be bounded by

$$T\mathbb{E}_{\boldsymbol{\eta}^N}\left[\frac{1}{(\sigma(\tau(T))-1)}\wedge 1\right],$$

where we have used the monotonicity of $\tau(\cdot)$ and $\sigma(\cdot)$. As noted in the proof of Corollary 1, the latter tends to zero as $N\to\infty$.

It now follows that

$$\lim_{N\to\infty}\mathbb{E}_{\boldsymbol{\eta}^N}\left[\sup_{t\le T}\left\|\boldsymbol{R}(t)-\boldsymbol{r}(0)-\int_0^t\sum_{i=1}^k \boldsymbol{e}_i R_i(s)[\boldsymbol{A}\boldsymbol{R}(s)]_i-(\boldsymbol{R}(s)^\top \boldsymbol{A}\boldsymbol{R}(s))\boldsymbol{R}(s)\,\mathrm{d}s\right\|_1 1_{A_t}\right]=0.$$

As all of the vectorial and matrix terms in this equation are bounded, it is also easy to see, with the help of Lemma 5, that

$$\lim_{N\to\infty}\mathbb{E}_{\boldsymbol{\eta}^N}\left[\sup_{t\le T}\left\|\boldsymbol{R}(t)-\boldsymbol{r}(0)-\int_0^t\sum_{i=1}^k \boldsymbol{e}_i R_i(s)[\boldsymbol{A}\boldsymbol{R}(s)]_i-(\boldsymbol{R}(s)^\top \boldsymbol{A}\boldsymbol{R}(s))\boldsymbol{R}(s)\,\mathrm{d}s\right\|_1 1_{A_t^c}\right]$$
$$\le \lim_{N\to\infty} CT\mathbb{P}_{\boldsymbol{\eta}^N}(A_T^c)=0$$

for some constant $C>0$, which gives us

$$\lim_{N\to\infty}\mathbb{E}_{\boldsymbol{\eta}^N}\left[\sup_{t\le T}\left\|\boldsymbol{R}(t)-\boldsymbol{r}(0)-\int_0^t\sum_{i=1}^k \boldsymbol{e}_i R_i(s)[\boldsymbol{A}\boldsymbol{R}(s)]_i-(\boldsymbol{R}(s)^\top \boldsymbol{A}\boldsymbol{R}(s))\boldsymbol{R}(s)\,\mathrm{d}s\right\|_1\right]=0. \quad (13)$$

Consider the replicator equation initiated from any $\boldsymbol{r}(0)\in\mathcal{S}_+^k$. Similarly to (12), albeit in vectorial form, we can write the solution to (2) when issued from $\boldsymbol{r}(0)$ as

$$\boldsymbol{x}(t)-\boldsymbol{r}(0)-\int_0^t\sum_{i=1}^k \boldsymbol{e}_i x_i(s)[\boldsymbol{A}\boldsymbol{x}(s)]_i-(\boldsymbol{x}(s)^\top \boldsymbol{A}\boldsymbol{x}(s))\boldsymbol{x}(s)\,\mathrm{d}s=0.$$

Subtracting this from (13), we see that, for each $T>0$,

https://doi.org/10.1017/jpr.2025.21 Published online by Cambridge University Press

$$
\begin{aligned}
&\mathbb{E}_{\boldsymbol{\eta}^N}\Big[\sup_{t\le T}\|\boldsymbol{R}(t)-\boldsymbol{x}(t)\|_1\Big]\\
&\quad\le \mathbb{E}_{\boldsymbol{\eta}^N}\left[\sup_{t\le T}\left\|\boldsymbol{R}(t)-\boldsymbol{r}(0)-\int_0^t\sum_{i=1}^k \boldsymbol{e}_i R_i(s)([\boldsymbol{A}\boldsymbol{R}(s)]_i-\boldsymbol{R}(s)^\top\boldsymbol{A}\boldsymbol{R}(s))\,\mathrm{d}s\right\|_1\right]\\
&\qquad+\int_0^T\sum_{i=1}^k \mathbb{E}_{\boldsymbol{\eta}^N}\Big[|R_i(s)-x_i(s)|\,|[\boldsymbol{A}\boldsymbol{R}(s)]_i-\boldsymbol{R}(s)^\top\boldsymbol{A}\boldsymbol{R}(s)|\Big]\,\mathrm{d}s\\
&\qquad+\int_0^T\sum_{i=1}^k x_i(s)\mathbb{E}_{\boldsymbol{\eta}^N}\Big[|[\boldsymbol{A}(\boldsymbol{R}(s)-\boldsymbol{x}(s))]_i-\boldsymbol{R}(s)^\top\boldsymbol{A}(\boldsymbol{R}(s)-\boldsymbol{x}(s))|\Big]\,\mathrm{d}s\\
&\qquad+\int_0^T\sum_{i=1}^k x_i(s)\mathbb{E}_{\boldsymbol{\eta}^N}[|(\boldsymbol{R}(s)-\boldsymbol{x}(s))^\top\boldsymbol{A}\boldsymbol{x}(s)|]\,\mathrm{d}s.
\end{aligned}
$$

Noting that $0\le R_i(s), x_i(s)\le 1$ for all $i=1,\dots,k$, $s\ge 0$, we have

$$
\begin{aligned}
&\mathbb{E}_{\boldsymbol{\eta}^N}[\sup_{t\le T}\|\boldsymbol{R}(t)-\boldsymbol{x}(t)\|_1]\\
&\quad\le \mathbb{E}_{\boldsymbol{\eta}^N}\left[\sup_{t\le T}\left\|\boldsymbol{R}(t)-\boldsymbol{r}(0)-\int_0^t\sum_{i=1}^k \boldsymbol{e}_i R_i(s)\big([\boldsymbol{A}\boldsymbol{R}(s)]_i-\boldsymbol{R}(s)^\top\boldsymbol{A}\boldsymbol{R}(s)\big)\,\mathrm{d}s\right\|_1\right]\\
&\qquad+C\int_0^T\mathbb{E}_{\boldsymbol{\eta}^N}[\sup_{u\le s}\|\boldsymbol{R}(u)-\boldsymbol{x}(u)\|_1]\,\mathrm{d}s,
\end{aligned}
$$

where $C>0$ is an unimportant constant. Hence, using (13), the monotonicity of norms, and dominated convergence,

$$
\begin{aligned}
u(T)&:=\lim_{N\to\infty}\sup_{N'\ge N}\mathbb{E}_{\boldsymbol{\eta}^N}\Big[\sup_{t\le T}\|\boldsymbol{R}(t)-\boldsymbol{x}(t)\|_1\Big]\\
&\le C\int_0^T\lim_{N\to\infty}\sup_{N'\ge N}\mathbb{E}_{\boldsymbol{\eta}^N}\Big[\sup_{u\le s}\|\boldsymbol{R}(u)-\boldsymbol{x}(u)\|_1\Big]\,\mathrm{d}s=C\int_0^T u(s)\,\mathrm{d}s,
\end{aligned}
$$

where $C>0$ is an unimportant constant. Grönwall's lemma now tells us that

$$
\lim_{N\to\infty}\mathbb{E}_{\boldsymbol{\eta}^N}\Big[\sup_{t\le T}\|\boldsymbol{R}(t)-\boldsymbol{x}(t)\|_1\Big]=0.
$$

By taking $t\to\infty$, we easily deduce that

$$
\lim_{t\to\infty}\lim_{N\to\infty}\mathbb{E}_{\boldsymbol{\eta}^N}[\|\boldsymbol{R}(t)-\boldsymbol{x}^*\|_1]\le\lim_{t\to\infty}\lim_{N\to\infty}\mathbb{E}_{\boldsymbol{\eta}^N}[\|\boldsymbol{R}(t)-\boldsymbol{x}(t)\|_1]+\lim_{t\to\infty}\|\boldsymbol{x}(t)-\boldsymbol{x}^*\|_1=0.
$$

This completes the proof of the theorem. □

6. Concluding remarks

For the purpose of the following discussion, we can assume that $C_{i,j}>0$ for all $i,j\in\{1,\dots,k\}$. In further work, it is possible pursue the issue of Skorokhod continuity with respect to any 'entrance laws at infinity' that the process can come down from.

Suppose $\mathbb{D}$ is the space of càdlàg paths from $[0,\infty)$ to $\mathbb{N}_*^k$, with $\mathcal{D}$ as the Borel sigma algebra on $\mathbb{D}$ generated from the usual Skorokhod metric. A Markovian definition of coming

https://doi.org/10.1017/jpr.2025.21 Published online by Cambridge University Press

down from infinity would require the existence of a law $\mathbb{P}^\infty$ on $(\mathbb{D}, \mathcal{D})$ that is consistent with $\mathbb{P}$ in the sense that

$$\mathbb{P}^\infty(\boldsymbol{n}(t+s)=\boldsymbol{n}) = \sum_{\substack{\boldsymbol{n}'\in\mathbb{N}_*^k \\ \|\boldsymbol{n}'\|_1\geq\|\boldsymbol{n}\|_1}} \mathbb{P}^\infty(\boldsymbol{n}(t)=\boldsymbol{n}')\mathbb{P}_{\boldsymbol{n}'}(\boldsymbol{n}(s)=\boldsymbol{n}), \quad s,t>0,\ \boldsymbol{n}\in\mathbb{N}_*^k,$$

with $\mathbb{P}^\infty(\sigma(t)<\infty)=1$ for all $t>0$, and $\mathbb{P}^\infty(\lim_{t\downarrow 0}\sigma(t)=\infty)=1$.

As there is no single point on the boundary of $\mathbb{N}_*^k$ that represents an appropriate 'infinity' to come down from, one of the associated issues is whether a unique entrance law exists or whether, e.g., there is an entrance law of $(\boldsymbol{r}, \sigma^{-1})$ for each 'infinity' of the form $(\boldsymbol{r}_0, 0)$, where $\boldsymbol{r}_0 \in \mathcal{S}_+^k$. We believe the latter to hold.

Conjecture 1. *An entrance law, say* $\mathbb{P}^{(\boldsymbol{r}_0,0)}$*, exists for each* $\boldsymbol{r}_0 \in \mathcal{S}_+^k$.

There is also the question of how we can see these different entrance laws in terms of the behaviour of the process at arbitrarily small times. The following conjecture suggests that looking backwards in time, it will be difficult to differentiate between the different entrance laws proposed in Conjecture 1.

Conjecture 2. *Suppose* $(\boldsymbol{\eta}^N, N\geq 1)$ *tends to* $(\boldsymbol{r}_0, 0)$ *for some* $\boldsymbol{r}_0 \in \mathcal{S}_+^k$*. Then*

$$\lim_{m\to\infty}\lim_{N\to\infty}\mathbb{E}_{\boldsymbol{\eta}^N}[\|\boldsymbol{r}(\gamma_m)-\boldsymbol{x}^*\|_1]=0,$$

where we recall that $\gamma_m = \inf\{t>0\colon \sigma(t)\leq m\}$ *for* $m\geq 1$.

In contrast to Theorem 1, Conjecture 2 claims that, moving backwards through time towards the instantaneous event at which the replicator coalescent comes down from infinity at the origin of time, the process $\boldsymbol{r}$ necessarily approaches $\boldsymbol{x}^*$. As such, working backwards in time, the replicator coalescent never gets to see the 'initial state' $(\boldsymbol{r}_0, 0)$ from which its entrance law is constructed.

Put together, Theorem 1 and Conjecture 2 are really claiming that $\boldsymbol{x}^*$ is a 'bottleneck' for the replicator coalescent. Figure 1 simulates an example where $k=3$, in which the bottleneck phenomenon can clearly be seen.

Theorem 1 also shows that under $\mathbb{P}_{\boldsymbol{\eta}^N}$ as $\boldsymbol{\eta}^N \to (\boldsymbol{r}_0, 0)$, in an arbitrarily small amount of time (on the natural time scale of the original Markov process), the process $\boldsymbol{r}$ will effectively jump from $\boldsymbol{r}_0$ to $\boldsymbol{x}^*$. Taking Conjecture 2 into account, it is for this reason we pose our final conjecture.

Conjecture 3. *Suppose* $(\boldsymbol{\eta}^N, N\geq 1)$ *tends to* $(\boldsymbol{r}_0, 0)$ *for some* $\boldsymbol{r}_0 \in \mathcal{S}_+^k$*. Then,* $\lim_{N\to\infty}\mathbb{P}_{\boldsymbol{\eta}^N} \to \mathbb{P}^{(\boldsymbol{r}_0,0)}$ *continuously on* $(\mathbb{D}, \mathcal{D})$ *if and only if* $\boldsymbol{r}_0 = \boldsymbol{x}^*$.

Acknowledgements

TR would like to thank Chris Guiver for useful discussions. Referees provided some extremely helpful feedback leading to the improvement of an earlier version of this document. AEK would also like to thank Jon Warren and Jason Schweinsberg for insightful discussion.

Funding information

LP was visiting TR and AEK as part of a doctoral visiting programme sponsored by the Internationalisation Research Office at the University of Bath. She would like to thank

https://doi.org/10.1017/jpr.2025.21 Published online by Cambridge University Press

the university for their support. AEK and TR acknowledge EPSRC grant support from EP/S036202/1.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] BARBOUR, A. D., ETHIER, S. N. AND GRIFFITHS, R. C. (2000). A transition function expansion for a diffusion model with selection. *Ann. Appl. Prob.* **10**, 123–162.
[2] BERESTYCKI, N. (2009). *Recent Progress in Coalescent Theory* (Ensaios Matemáticos [Mathematical Surveys] 16). Sociedade Brasileira de Matemática, Rio de Janeiro.
[3] HOFBAUER, J. AND SIGMUND, K. (1998). *Evolutionary Games and Population Dynamics*. Cambridge University Press.
[4] JOHNSTON, S., KYPRIANOU, A. E. AND ROGERS, T. (2023). Multi-type Λ-coalescents. *Ann. Appl. Prob.* **33**, 4210–4237.
[5] KINGMAN, J. F. C. (1999). Martingales in the OK Corral. *Bull. London Math. Soc.* **31**, 601–606.
[6] KINGMAN, J. F. C. AND VOLKOV, S. E. (2003). Solution to the OK Corral model via decoupling of Friedman's urn. *J. Theoret. Prob.* **16**, 267–276.
[7] KINGMAN, J. F. C. AND VOLKOV, S. E. (2019). Correction to: Solution to the OK Corral model via decoupling of Friedman's Urn. *J. Theoret. Prob.* **32**, 1614–1615.
[8] PROTTER, P. E. (2004). *Stochastic Integration and Differential Equations* (Appl. Math. (New York) 21), 2nd edn. Springer, Berlin.

https://doi.org/10.1017/jpr.2025.21 Published online by Cambridge University Press