

Supplementary Material for:

Fixation properties of rock-paper-scissors games in fluctuating populations

Robert West and Mauro Mobilia *

In this Supplementary Material (SM), we provide additional information about the relationships between various rock-paper-scissors models (Section 1), and further technical details concerning the stages 1 and 2 dynamics (Section 2 and 3). We also analyze the population composition at the inception of Stage 2 (Section 4), as well as the mean extinction, absorption and fixation times (Section 5) and discuss the average number of switches occurring in Stages 1 and 2. The notation in this SM is the same as in the main text; all equations not given in this SM refer to those of the main text.

1 Various cyclic Lotka-Volterra models (zero-sum rock-paper-scissors games): general properties, similarities and differences

In the literature, there are various formulations of the zero-sum rock-paper-scissors game, here generically referred to as “cyclic Lotka-Volterra” models. Here, we consider the birth-death cyclic Lotka Volterra model (BDCLV), defined in the main text by (2)-(7), the cyclic Lotka Volterra model formulated in terms of a Moran process (MCLV), and finally the so-called chemical cyclic Lotka volterra model (cCLV). These models are characterized by many similar features, but also some important differences. Below, we outline some of the main properties of these models and discuss their similarities and differences.

1.1 The birth-death cyclic Lotka-Volterra model (BDCLV): Mean-field equations and piecewise deterministic Markov process

The BDCLV is here defined in terms of the six reactions

$$N_i \xrightarrow{T_i^+} N_i + 1 \quad \text{and} \quad N_i \xrightarrow{T_i^-} N_i - 1, \quad \text{with} \quad i \in \{1, 2, 3\}, \quad (\text{S1})$$

the first set of reactions corresponds to the birth of an individual of species i and the other reaction is associated with the death of an i -individual. These reactions occur with transition rates

$$T_i^+ = f_i N_i = (1 + s\Pi_i) N_i = (1 + \{\alpha_i x_{i+1} - \alpha_{i-1} x_{i-1}\}) N_i \quad \text{and} \quad T_i^- = \frac{N}{K(t)} N_i, \quad \text{where} \quad N = \sum_{i=1}^3 N_i \quad (\text{S2})$$

is the total population size and $K(t)$ is the carrying capacity. In this work, we consider the case of a constant and randomly switching carrying capacity, namely

$$K(t) = \begin{cases} K & \text{constant,} & \text{see Sec. 3 in main text} \\ \frac{1}{2} [(K_+ + K_-) + \xi(t) (K_+ - K_-)], & \text{with dichotomous noise } \xi \in \{-1, +1\}, & \text{see Sec. 4 in main text} \end{cases}$$

The formulation of the cyclic competition in terms of the BDCLV allows us to conveniently introduce the carrying capacity through the death rate T_i^- and, the population size not being conserved, also enables us to aptly model the cyclic dynamics when the population size fluctuates and possibly varies greatly in time.

The BDCLV dynamics is fully described by the underpinning master equation (7) from which the equation of motion of the average number of individual of species i in the environmental state ξ can be derived as usual[†] [1, 2]

$$\frac{d}{dt} \langle N_i \rangle = \frac{d}{dt} \sum_{\vec{N}} N_i P(\vec{N}, \xi, t) = \langle T_i^+ \rangle - \langle T_i^- \rangle,$$

*Department of Applied Mathematics, School of Mathematics, University of Leeds, Leeds LS2 9JT, United Kingdom.
Email: mmrw@leeds.ac.uk, m.mobilia@leeds.ac.uk

[†]In this section, for notational convenience $\langle X(\vec{N}) \rangle = \sum_{\vec{N}} X(\vec{N}) P(\vec{N}, \xi, t)$ denotes the average of the observable $X(\vec{N})$ when the environment remains in the state ξ . This should not be confused with the notation used in the main text where the angular bracket refers to the average over the environmental noise ξ .

where

$$\langle T_i^- \rangle \equiv \begin{cases} \left\langle \frac{N}{K} N_i \right\rangle & , \text{ when } K \text{ is constant} \\ \left\langle T_i^- | + \right\rangle = \left\langle \frac{N}{K^+} N_i \right\rangle & , \text{ in state } \xi = +1 \text{ when } K \text{ is switching} \\ \left\langle T_i^- | - \right\rangle = \left\langle \frac{N}{K^-} N_i \right\rangle & , \text{ in state } \xi = -1 \text{ when } K \text{ is switching} \end{cases}$$

This readily leads to the following equations for the average population size $\langle N \rangle$ in a static environment (constant K):

$$\frac{d}{dt} \langle N \rangle = \sum_{i=1}^3 (\langle T_i^+ \rangle - \langle T_i^- \rangle), \quad \text{and in a varying environment with a randomly switching } K: \quad (\text{S3})$$

$$\frac{d}{dt} \langle N \rangle = \begin{cases} \sum_{i=1}^3 (\langle T_i^+ \rangle - \langle T_i^- | + \rangle) & \text{if } \xi = +1 \\ \sum_{i=1}^3 (\langle T_i^+ \rangle - \langle T_i^- | - \rangle) & \text{if } \xi = -1 \end{cases}. \quad (\text{S4})$$

For the population composition, we can proceed similarly to derive the equation motion for $\langle x_i \rangle \equiv \langle N_i/N \rangle = \sum_{\vec{N}} (N_i/N) P(\vec{N}, \xi, t)$, paying due attention to the fact that now both N_i and N vary in time:

$$\begin{aligned} \frac{d}{dt} \langle x_i \rangle &= \sum_{\vec{N}} \frac{N_i}{N} \frac{d}{dt} P(\vec{N}, \xi, t) \\ &= \sum_{\vec{N}} \frac{N_i}{N} \left\{ T_i^+(\vec{N} - \vec{e}_i) P(\vec{N} - \vec{e}_i, \xi, t) + T_i^-(\vec{N} + \vec{e}_i, K) P(\vec{N} + \vec{e}_i, \xi, t) - \left(T_i^+(\vec{N}) + T_i^-(\vec{N}, K) \right) P(\vec{N}, \xi, t) \right\} \\ &+ \sum_{j \in \{1,2,3\} \neq i, \vec{N}} \frac{N_i}{N} \left\{ T_j^+(\vec{N} - \vec{e}_j) P(\vec{N} - \vec{e}_j, \xi, t) + T_j^-(\vec{N} + \vec{e}_j, K) P(\vec{N} + \vec{e}_j, \xi, t) \right\} \\ &- \sum_{j \in \{1,2,3\} \neq i, \vec{N}} \frac{N_i}{N} \left(T_j^+(\vec{N}) + T_j^-(\vec{N}, K) \right) P(\vec{N}, \xi, t) \quad (\text{S5}) \\ &= \sum_{\vec{N}} \left\{ \frac{N_i + 1}{N + 1} T_i^+(\vec{N}) P(\vec{N}, \xi, t) + \frac{N_i - 1}{N - 1} T_i^-(\vec{N}, K) P(\vec{N}, \xi, t) - \frac{N_i}{N} \left(T_i^+(\vec{N}) + T_i^-(\vec{N}, K) \right) P(\vec{N}, \xi, t) \right\} \\ &+ \sum_{j \in \{1,2,3\} \neq i, \vec{N}} \left\{ \frac{N_i}{N + 1} T_j^+(\vec{N}) P(\vec{N}, \xi, t) + \frac{N_i}{N - 1} T_j^-(\vec{N}, K) P(\vec{N}, \xi, t) - \frac{N_i}{N + 1} \left(T_j^+(\vec{N}) + T_j^-(\vec{N}, K) \right) P(\vec{N}, \xi, t) \right\}, \end{aligned}$$

where \vec{e}_i is the unit vector such that $\vec{e}_1 = (1, 0, 0)$, etc. By rearranging the right-hand-side of (S5) and, for notational convenience, by writing $T^+ \equiv T^+(\vec{N})$ and $T^- \equiv T^-(\vec{N}, K)$, we obtain

$$\begin{aligned} \frac{d}{dt} \langle x_i \rangle &= \left\langle \left(\frac{N_i + 1}{N + 1} - \frac{N_i}{N} \right) T_i^+ \right\rangle + \left\langle \left(\frac{N_i - 1}{N - 1} - \frac{N_i}{N} \right) T_i^- \right\rangle \\ &+ \sum_{j \in \{1,2,3\} \neq i} \left\{ \left\langle \left(\frac{N_i}{N + 1} - \frac{N_i}{N} \right) T_j^+ \right\rangle + \left\langle \left(\frac{N_i}{N - 1} - \frac{N_i}{N} \right) T_j^- \right\rangle \right\} \\ &= \left\langle \frac{T_i^+ - T_i^-}{N} \left(1 + \mathcal{O} \left(\frac{1}{N} \right) \right) \right\rangle - \left\langle \frac{x_i}{N} \left(1 + \mathcal{O} \left(\frac{1}{N} \right) \right) \sum_{j=1}^3 (T_j^+ - T_j^-) \right\rangle. \quad (\text{S6}) \end{aligned}$$

We can now derive the mean-field equations (constant K) and the stochastic differential equation defining the piecewise-deterministic Markov process for the evolution of the population size. For this, as usual, we ignore all demographic fluctuations and factorize all terms appearing on the right-hand-side of (S3) and (S6) in terms of $\langle x_i \rangle$ and $\langle N \rangle$, respectively denoted by x_i and N , e.g. $\langle x_i x_j \rangle \rightarrow x_i x_j$, $\langle f_i x_j N \rangle \rightarrow f_i x_j N$ and $\langle N_i N \rangle \rightarrow N_i N$. In the case of a constant

carrying capacity, making the natural mean-field assumption that N is always sufficiently large for contributions of order $\mathcal{O}(x_i/N)$ to be negligible, using (S3), we obtain:

$$\begin{aligned}\frac{d}{dt}N &= \sum_{i=1}^3 (T_i^+ - T_i^-) = N \left(1 - \frac{N}{K}\right), \\ \frac{d}{dt}x_i &= \frac{T_i^+ - T_i^-}{N} - x_i \left(\frac{dN/dt}{N}\right) = s\Pi_i x_i = x_i (\alpha_i x_{i+1} - \alpha_{i-1} x_{i-1}),\end{aligned}\quad (\text{S7})$$

where we have used $\bar{f} = 1$ and $\alpha_i \equiv sr_i$. These mean-field equations coincide with the decoupled rate equations (8) and (9) discussed in the main text. In the case of a randomly switching carrying capacity, the x_i 's still obey (S7) while the population size evolves according to

$$\frac{d}{dt}N = \begin{cases} N \left(1 - \frac{N}{K_+}\right) & \text{if } \xi = +1 \\ N \left(1 - \frac{N}{K_-}\right) & \text{if } \xi = -1, \end{cases}\quad (\text{S8})$$

which can be rewritten as the stochastic differential equation (20) defining the piecewise deterministic Markov process governing the evolution of the population size $N(t)$ when demographic is ignored and whose stationary marginal probability density is given (21).

Similar derivations also hold in the general (non-zero-sum) rock-paper-scissors game, whose birth-death formulation is given by the rates T_i^\pm of Eq. (34) and leads to the mean-field equations of Sec. 5 of the main text.

A comment on our choice of the transition rates and of the model formulation is here in order: With (5) and (6) we have arguably chosen the simplest formulation of the RPS dynamics subject to a carrying capacity. It is however worth noting that other choices are of course also possible. Another natural possibility would be to use the transition rates $T_i^+ = f_i N_i / \bar{f}$ and $T_i^- = (N/K)N_i$ [3]. Clearly, for the BDCLV these transition rates coincide with (6) since $\bar{f} = 1$ when $\epsilon = 0$. A difference however arises when $\epsilon \neq 0$ and $\bar{f} = 1 - \epsilon \sum_{i=1}^3 \alpha_i x_i x_{i+1}$. In fact, proceeding as above and using the rates $T_i^+ = f_i N_i / \bar{f}$ and $T_i^- = (N/K)N_i$ in the master equation (7), we obtain the following mean-field rate equations: $\dot{N} = N(1 - N/K)$ and $\dot{x}_i = x_i(f_i - \bar{f})/\bar{f} = x_i[\alpha_i x_{i+1} - (1 + \epsilon)\alpha_{i-1} x_{i-1} + 1 - \bar{f}]/\bar{f}$. While these equations are decoupled, the rate equations for the x_i 's do *not* coincide with the celebrated replicator equations (34) of the general RPS game [4]: The x_i 's mean-field rate equations obtained with the above alternative transition rates [5], differ from (34) due to the nonlinear \bar{f} term appearing in the denominator on their right-hand-side. The mean-field rate equations $\dot{x}_i = x_i(f_i - \bar{f})/\bar{f}$ and Eqs. (34) however coincide to leading order in $s\epsilon$.

1.2 The Moran cyclic Lotka-Volterra model (MCLV)

We now outline the main features of the Moran cyclic Lotka Volterra model (MCLV) in a static environment (no environmental noise). The MCLV is defined by six pairwise reactions and is characterized by the *conservation* of the population size N [3, 6, 7, 8, 9]. Each of the six reactions corresponds to the *simultaneous* death of an individual of species i and the birth of an individual of species $j \neq i \in \{1, 2, 3\}$ [3]. This occurs with a rate $T_{i \rightarrow j}$. If the state of the system consisting of N_1 individuals of type 1, N_2 of species 2, and $N_3 = N - N_1 - N_2$ of the third type is denoted by $[N_1, N_2]$, the six reactions of the MCLV are [10, 11, 12]

$$\begin{aligned}[N_1, N_2] &\xrightarrow{T_{1 \rightarrow 2}} [N_1 - 1, N_2 + 1]; [N_1, N_2] \xrightarrow{T_{2 \rightarrow 1}} [N_1 + 1, N_2 - 1] \\ [N_1, N_2] &\xrightarrow{T_{1 \rightarrow 3}} [N_1 - 1, N_2]; [N_1, N_2] \xrightarrow{T_{3 \rightarrow 1}} [N_1 + 1, N_2] \\ [N_1, N_2] &\xrightarrow{T_{2 \rightarrow 3}} [N_1, N_2 - 1]; [N_1, N_2] \xrightarrow{T_{3 \rightarrow 2}} [N_1, N_2 + 1],\end{aligned}$$

with the transition rates [10, 11]

$$T_{j \rightarrow i} = f_i x_i x_j N = (1 + s\Pi_i) x_i x_j N = (1 + \{\alpha_i x_{i+1} - \alpha_{i-1} x_{i-1}\}) x_i x_j N, \quad (\text{S9})$$

where f_i and Π_i are given by (2) and (1). Interestingly, the transition rates of the MCLV can be expressed in terms of those of the BDCLV for a population of constant size $N = K$. In fact, using (S2) and $N = K$, we have $T_{j \rightarrow i} =$

$T_i^+ T_j^- / K$. This means that the BDCLV coincides with the MCLV in a population of constant size $N = K$, see below. Proceeding as above, we can readily find the mean-field rate equations for the MCLV:

$$\frac{d}{dt} x_i = \frac{1}{N} \sum_{j=1; j \neq i}^3 (T_{j \rightarrow i} - T_{i \rightarrow j}) = s \Pi_i x_i = x_i (\alpha_i x_{i+1} - \alpha_{i-1} x_{i-1}),$$

which coincide with the mean-field (replicator) equations for the population composition in the BDCLV, see (S7) and (9). Clearly therefore, in the constant- K BDCLV the dynamics of the population composition coincides with that of the MCLV in the mean-field limit $K \rightarrow \infty$: both are characterized by the same neutrally stable fixed point $\vec{x}^* = (r_2, r_3, r_1) = (r_2, 1 - r_1 - r_2, r_1)$ and constant of motion $\mathcal{R} = x_1^{r_2} x_2^{r_3} x_3^{r_1}$.

Since in the constant- K BDCLV dynamics the population size obeys a logistic equation, after a short transient $N(t) \approx K$, see Eq. (8) and Fig. 1. This establishes a useful relationship between the BDCLV and MCLV: Except for a short transient (on a timescale $t \sim \mathcal{O}(1)$), corresponding to the so-called exponential phase of the logistic equation, the evolution of the constant- K BDCLV is similar to the dynamics of the MCLV in a population of constant size $N = K$. The BDCLV and MCLV relation is particularly useful to determine the absorption/fixation properties of the former in terms of the well-studied fixation properties of latter, see Secs. 3.1.2 and 3 of this SM. In Fig. S1, we show that the survival and absorption probabilities $\phi_{i,j}$ and ϕ_i in the constant- K BDCLV are almost indistinguishable from those obtained in the MCLV (with $N = K$). Since the overall fixation probabilities $\tilde{\phi}_i = \phi_{i,i+1} \phi_i + \phi_{i-1,i} (1 - \phi_i)$, see Eq. (16), we can consider that the absorption and total fixation probabilities in the constant- K BDCLV and those of the MCLV with $N = K \gg 1$ coincide. Similarly, the mean extinction and absorption times T_1 and T_2 in the BDCLV with constant- K and MCLV with $N = K \gg 1$ are indistinguishable, see the insets of Fig. S1 and below.

To study the absorption/fixation properties of the BDCLV and MCLV, it is useful to write down the two-dimensional forward Fokker-Planck equation (FPE) obeyed by the probability density $P_{\text{MCLV}} \equiv P_{\text{MCLV}}(\vec{x}, t)$ of the latter. Using standard methods, see, e.g. Refs. [1, 2, 13, 10, 11] we have the forward FPE

$$[\partial_t - \mathcal{G}_{\text{fMCLV}}(\vec{x})] P_{\text{MCLV}}(\vec{x}, t) = 0, \quad \text{where} \quad \mathcal{G}_{\text{fMCLV}}(\vec{x}) \equiv - \sum_{i=1}^2 \partial_i A_i^{\text{MCLV}}(\vec{x}) + \frac{1}{2} \sum_{i,j=1}^2 \partial_i \partial_j B_{ij}^{\text{MCLV}}(\vec{x}), \quad (\text{S10})$$

is the forward FPE generator, with $\partial_i \equiv \partial / \partial x_i$ [‡], defined by

$$A_i^{\text{MCLV}}(\vec{x}) \equiv \sum_{j=1, j \neq i}^3 (T_{j \rightarrow i} - T_{i \rightarrow j});$$

$$B_{ii}^{\text{MCLV}}(\vec{x}) \equiv \sum_{j=1, j \neq i}^3 \left(\frac{T_{j \rightarrow i} + T_{i \rightarrow j}}{N} \right) \quad \text{and} \quad B_{12}^{\text{MCLV}}(\vec{x}) = B_{21}^{\text{MCLV}}(\vec{x}) \equiv - \left(\frac{T_{1 \rightarrow 2} + T_{2 \rightarrow 1}}{N} \right). \quad (\text{S11})$$

Within the linear noise approximation [1, 2], upon linearising A_i^{MCLV} about the coexistence fixed point \vec{x}^* and by evaluating $B_{ij}^{\text{MCLV}}(\vec{x})$ at \vec{x}^* , in the variables $\vec{y} = \mathbf{S}\vec{x} = \frac{\sqrt{3}}{2} \begin{pmatrix} (r_1+r_2)\omega_0^{\text{MCLV}} & \omega_0^{\text{MCLV}} \\ r_1 r_2 & r_1 \\ 0 & 1 \end{pmatrix} \vec{x}$, the forward FPE reads [13, 11]

$$\partial_t P_{\text{MCLV}}(\vec{y}, t) = -\omega_0^{\text{MCLV}} [y_1 \partial_{y_1} - y_2 \partial_{y_2}] P_{\text{MCLV}}(\vec{y}, t) + D^{\text{MCLV}} [\partial_{y_1}^2 + \partial_{y_2}^2] P_{\text{MCLV}}(\vec{y}, t), \quad (\text{S12})$$

where $\omega_0^{\text{MCLV}} = s \sqrt{r_1 r_2 (1 - r_1 - r_2)}$ and $D^{\text{MCLV}} = 3[r_1 + r_2 - 4r_1 r_2 - (r_1 - r_2)^2] / (4N)$. To study the fixation properties of the MCLV, the FPEs (S11) and (S12) have to be supplemented with absorbing boundaries at the corners of S_3 [13, 14, 15].

1.3 The chemical cyclic Lotka-Volterra model (cCLV)

The chemical cyclic Lotka-Volterra model (cCLV) is defined by three pairwise (“bimolecular”) reactions involving the simultaneous death and birth of individuals of different species, therefore conserving the total population size N . Hence, in the cCLV, in contrast to the BDCLV and MCLV, species i is the predator of species $i + 1$ and the

[‡]In Eq. (S10), the indices $i, j \in \{1, 2\}$ since $x_3 = 1 - x_1 - x_2$ and, as usual in the diffusion theory, we have rescaled the time $t \rightarrow t/N$.

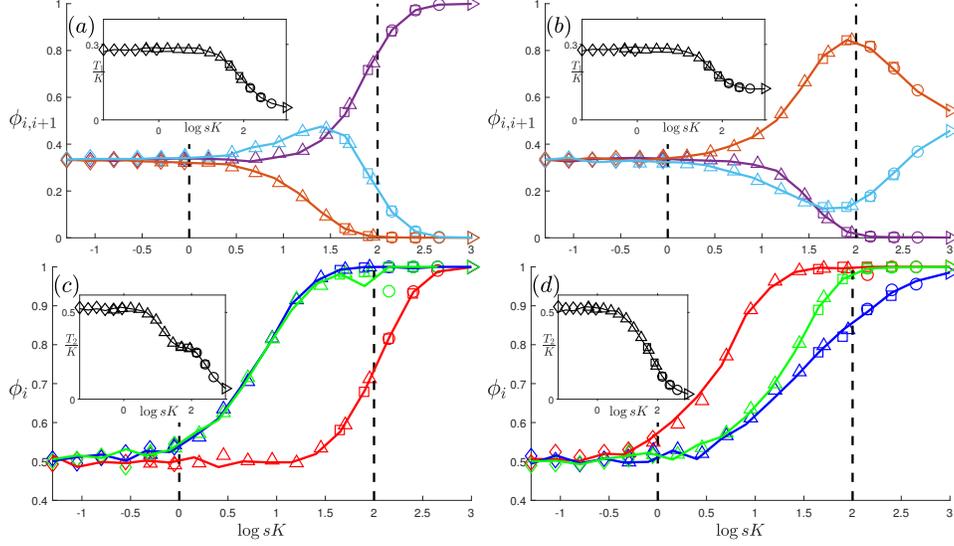
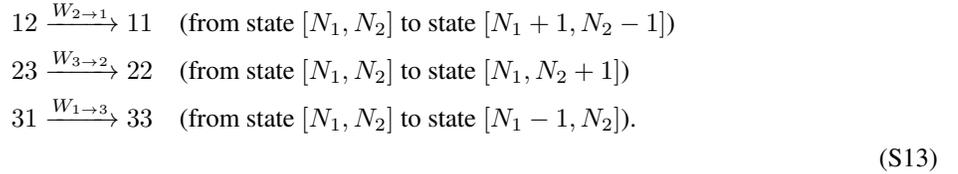


Figure S1: Comparison of the fixation properties vs. sK in the BDCLV (solid lines) with constant carrying capacity K and in the MCLV (symbols) with a constant population size $N = K \in \{1000 (\triangleright), 450 (\circ), 250 (\diamond), 90 (\square), 50 (\triangle)\}$, with $\bar{r} = \bar{r}^{(1)}$ in (a,c) and $\bar{r} = \bar{r}^{(2)}$ in (b,d) and different values of selection intensity: $s \in \{10^{-j/4}, j \in J_K^{\text{MCLV}}\}$ with $J_{1000}^{\text{MCLV}} = \{0\}$, $J_{450}^{\text{MCLV}} = \{0, \dots, 3\}$, $J_{250}^{\text{MCLV}} = \{0, \dots, 4\}$, $J_{90}^{\text{MCLV}} = \{0, \dots, 10\}$, $J_{50}^{\text{MCLV}} = \{7, \dots, 12\}$ for the MCLV and $s \in \{10^{-j/4}, j \in J_K^{\text{BDCLV}}\}$ with $J_{1000}^{\text{BDCLV}} = \{1\}$, $J_{450}^{\text{BDCLV}} = \{0, \dots, 12\}$, $J_{90}^{\text{BDCLV}} = \{10, 11, 12\}$, $J_{50}^{\text{BDCLV}} = \{12\}$ for the BDCLV. (a,b) Stage 1 survival probabilities $\phi_{1,2}$ (purple), $\phi_{2,3}$ (light blue) and $\phi_{3,1}$ (orange) vs. sK : BCLV results (lines) match perfectly with those obtained for the MCLV (symbols). Insets: Rescaled mean extinction times T_1/K vs. sK for the BDCLV (solid lines) and MCLV (symbols) virtually coincide, see text. (c,d) Stage 2 conditional fixation probabilities ϕ_1 (red), ϕ_2 (blue) and ϕ_3 (green) vs. sK : BCLV results (lines) agree perfectly with those obtained for the MCLV (symbols). Insets: Rescaled mean absorption times T_2/K vs. sK for the BDCLV (solid lines) and MCLV (symbols) almost coincide, see text. In all panels: $\bar{x}_0 = \bar{x}_c$, $\epsilon = 0$; regimes (i)-(iii), from left to right, are indicatively separated by dashed lines. Simulation results for the fixation probabilities of in the constant- K BDCLV and MCLV with $N = K$ are almost indistinguishable, see text.

prey of species $i - 1$: an i -individual kills and replaces an $(i + 1)$ -individual with one of its offspring, while it is killed and replaced by individual of type $i - 1$ according to the following “bimolecular chemical reactions”, with $N_3 = N - N_1 - N_2$:



These reactions occur with the transition rates [13, 14, 15]

$$W_{i+1 \rightarrow i} = k_i \frac{N_i N_{i+1}}{N} = k_i x_i x_{i+1} N, \quad \text{where } k_i \geq 0. \tag{S14}$$

Clearly, the reactions (S13) and transition rates (S14) differ from those of the BDCLV and MCLV. Yet, as discussed below many of the features of the BDCLV, MCLV and cCLV are similar. The cCLV mean-field equations for the x_i 's are given by

$$\frac{d}{dt} x_i = \frac{W_{i+1 \rightarrow i} - W_{i \rightarrow i-1}}{N} = x_i (k_i x_{i+1} - k_{i-1} x_{i-1}). \tag{S15}$$

We notice that upon rescaling the time as $t \rightarrow st/(k_1 + k_2 + k_3)$, the reaction rates become $k_i \rightarrow k_i/(k_1 + k_2 + k_3) = r_i$ and Eq. (S15) is identical to Eq. (S7). Hence, upon time rescaling, the MCLV and cCLV are identical at mean-field

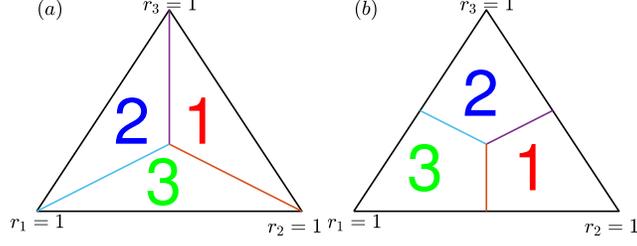


Figure S2: Law of the weakest (a) and law of stay out (b) in the simplex S_3 spanned by r_i , divided into three regions where the most likely species to survive is labelled. On the lines separating these regions, both adjacent species are equally likely to survive. (a) Law of the weakest (LOW): In the cCLV, the most likely species to survive in a large population is that with the lowest r_i . The LOW becomes asymptotically a zero-one law and also applies to the constant- K BDCLV and MCLV when $N = K$ and $sK \gg 1$ (regime (iii)), see text. (b) Law of stay out (LOSO) when all species initially coexist with the same density: In the cCLV, no species is guaranteed to survive in small populations, see text. The LOSO also applies to the constant- K BDCLV and MCLV when $N = K$ and $sK = \mathcal{O}(10)$ (regime (ii)), see text.

level and their dynamics coincide with the REs (9) of the BDCLV. Moreover, Eqs. (S7) and (S10) admit the same marginally stable coexistence fixed point $\vec{x}^* = (k_2, k_3, k_1)/(k_1 + k_2 + k_3) = (r_2, r_3, r_1)$ and the same constant of motion $\mathcal{R} = \left(x_1^{k_2} x_2^{k_3} x_3^{k_1}\right)^{1/(k_1+k_2+k_3)}$. The mean-field dynamics of the x_i 's is therefore identical for the BDCLV, MCLV and cCLV.

It is useful to proceed as above and consider the two-dimensional forward Fokker-Planck equation (FPE) obeyed by the cCLV probability density $P_{\text{cCLV}} \equiv P_{\text{cCLV}}(\vec{x}, t)$ (with $t \rightarrow t/N$):

$$[\partial_t - \mathcal{G}_{\text{cCLV}}(\vec{x})] P_{\text{cCLV}}(\vec{x}, t) = 0, \quad \text{where} \quad \mathcal{G}_{\text{cCLV}}(\vec{x}) \equiv - \sum_{i=1}^2 \partial_i A_i^{\text{cCLV}}(\vec{x}) + \frac{1}{2} \sum_{i,j=1}^2 \partial_i \partial_j B_{ij}^{\text{cCLV}}(\vec{x}), \quad (\text{S16})$$

with $A_i^{\text{cCLV}}(\vec{x}) \equiv W_{i+1 \rightarrow i} - W_{i \rightarrow i-1}$, $B_{ii}^{\text{cCLV}}(\vec{x}) \equiv (W_{i+1 \rightarrow i} + W_{i \rightarrow i-1})/N$ where $i \in \{1, 2\}$, and $B_{12}^{\text{cCLV}}(\vec{x}) \equiv B_{21}^{\text{cCLV}}(\vec{x}) \equiv -(W_{1 \rightarrow 2} + W_{2 \rightarrow 1})/N$. It is worth noting that the drift terms of the cCLV and MCLV are simply related by $A_i^{\text{cCLV}} = sA_i^{\text{MCLV}}/(k_1 + k_2 + k_3)$. In the case of symmetric rates, $k_1 = k_2 = k_3 = 1$, within the linear noise approximation, this forward FPE in the variables $\vec{y} = S\vec{x}$ reads:

$$\partial_t P_{\text{cCLV}}(\vec{y}, t) = -\omega_0^{\text{cCLV}} [y_1 \partial_{y_1} - y_2 \partial_{y_2}] P_{\text{cCLV}}(\vec{y}, t) + D^{\text{cCLV}} [\partial_{y_1}^2 + \partial_{y_2}^2] P_{\text{cCLV}}(\vec{y}, t), \quad (\text{S17})$$

where $\omega_0^{\text{cCLV}} = 1/\sqrt{3}$ and $D^{\text{cCLV}} = 1/(12N)$ [13]. This FPE is similar to Eq. (S11). The comparison with the MCLV with equal rates $r_i = 1/3$ is particularly illuminating: $\omega_0^{\text{MCLV}} = s\omega_0^{\text{cCLV}}/3$ and $D^{\text{MCLV}} = 2D^{\text{cCLV}}$. Hence, upon a suitable rescaling of the timescale, the MCLV and cCLV deterministic drift and diffusive terms (about \vec{x}^*) can be mapped onto each other.

1.3.1 Fixation probabilities in the cCLV: The law of the weakest and the law of stay out

Due to the predator-prey interactions underpinning the cCLV, its fixation properties of the cCLV are entirely set by the stage 1 of its dynamics: the probability $\phi_{i,i+1}^{\text{cCLV}}$ that species i and $i+1$ survive the stage 1 coincides with the fixation probability ϕ_i^{cCLV} of species i : $\phi_{i,i+1}^{\text{cCLV}} = \phi_i^{\text{cCLV}}$. The survival/fixation probability ϕ_i^{cCLV} of the cCLV can be explained by two simple laws called the *law of the weakest* (LOW) and the *law of stay out* (LOSO) [14, 17, 16, 15], see Fig.S2. The former applies to populations of large size and the latter to small populations. The LOW says that in a sufficiently large population (when $N \gtrsim 10^2$) evolving according to the cCLV, the most likely species to survive is the one with the lowest rate k_i [14] ($i \in \{1, 2, 3\}$), see Fig.S2 (a):

$$\phi_i^{\text{cCLV}} > \phi_{i+1}^{\text{cCLV}}, \phi_{i-1}^{\text{cCLV}} \text{ if } k_i < k_{i\pm 1}, \quad \text{and} \quad \phi_i^{\text{cCLV}} \approx \phi_{i+1}^{\text{cCLV}} > \phi_{i-1}^{\text{cCLV}} \text{ if } k_i = k_{i+1} < k_{i-1}, \text{ with } i \in \{1, 2, 3\}. (\text{S18})$$

The LOW becomes asymptotically a zero-one law (when $N \gtrsim 10^3$):

$$\phi_i^{\text{cCLV}} \rightarrow 1, \phi_{i\pm 1} \rightarrow 0 \text{ if } k_i < k_{i\pm 1}, \quad \text{while} \quad \phi_i^{\text{cCLV}} = \phi_{i+1}^{\text{cCLV}} \rightarrow 1/2 \text{ and } \phi_{i-1}^{\text{cCLV}} \rightarrow 0 \text{ if } k_i = k_{i+1} < k_{i-1}. (\text{S19})$$

The LOW is independent of the initial condition and results from the fact that in large populations, due to the effect of weak demographic noise the cCLV trajectories perform random walks between the deterministic orbits until they reach the so-called “outermost orbit”. This is obtained from the constant of motion \mathcal{R} as the deterministic orbit that lying at a distance $1/N$ from the closest edge of S_3 [14, 15]. In the cCLV, the extinction of a first species occurs when a chance fluctuation pushes a trajectory along the edge of S_3 from where the absorbing state corresponding to the fixation of the “weakest species” (with lowest k_i) and death of its “prey” is attained exponentially quickly.

The LOSO is a non-zero-one law prescribing which species is most likely to survive in small populations ($3 \leq N \lesssim 50$). The LOSO results from the interplay between the deterministic drift and demographic fluctuations and its prescriptions depend on the initial condition. In the cCLV, when initially all species have the same density, i.e. $\vec{x}_0 = \vec{x}_c$, the LOSO says that the most likely species to survive is/are the one(s) predated on the species with the highest k_i 's, see Fig.S2 (b) [14, 15][§]:

$$\phi_{i-1}^{\text{cCLV}} > \phi_i^{\text{cCLV}}, \phi_{i+1}^{\text{cCLV}} \quad \text{if} \quad k_i > k_{i+1}, k_{i-1}, \quad \text{and} \quad \phi_i^{\text{cCLV}} \approx \phi_{i+1}^{\text{cCLV}} > \phi_{i-1}^{\text{cCLV}} \quad \text{if} \quad k_{i+1} = k_{i-1} > k_i. \quad (\text{S20})$$

The LOSO can be understood by estimating the initial drift at \vec{x}_0 with the Jacobian \mathbf{J}^* of (S15) evaluated at \vec{x}^* . When, as here, $\vec{x}_0 \neq \vec{x}^*$, the rate of the bias from \vec{x}_0 towards a corner i of S_3 is $(\mathbf{J}^* \vec{x}_0)_i = k_i x_{i+1}(0) - k_{i-1} x_{i-1}(0)$. Hence, $k_i^- \equiv k_i x_{i+1}(0) - k_{i-1} x_{i-1}(0)$, gives the initial *deterministic rate* in the direction i . The most likely species to die out first is therefore the one with the smallest k_i^- (edge $(i-1, i+1)$ as the most likely to be hit first). With this reasoning, and $k_i^- = (k_i - k_{i-1})/3$ when $\vec{x}_0 = \vec{x}_c$, we find that the species that is the least likely to survive/fixate in the cCLV satisfies (S20) when all species initially coexist with the same density $x_i(0) = 1/3$.

1.3.2 Mean extinction and fixation times in the cCLV

The cCLV dynamics is also characterized by two stages: in Stage 1, the three species coexist until an edge of S_3 is hit and one of the species dies out (see Sec. 3.1.1) after a mean extinction time $T_1^{\text{cCLV}} = \mathcal{O}(N)$, see Sec. E.1.1. While the stage 1 dynamics of the cCLV, MCLV and constant- K BDCLV are similar (when $s = \mathcal{O}(1)$), a major difference arises in Stage 2, when two species, say i and its weak opponent $i+1$, compete along the edge $(i, i+1)$ of S_3 . According to the cCLV, the interaction between species i (predator) and $i+1$ (prey) is of predator-prey type, and the outcome of Stage 2 is certain: Contrarily to the MCLV and BDCLV, species i always wins against $i+1$ exponentially quickly in time. The overall cCLV mean fixation time $T_F^{\text{cCLV}} = T_1^{\text{cCLV}} + T_2^{\text{cCLV}}$ therefore coincides with T_1^{cCLV} to leading order, yielding $T_F^{\text{cCLV}} \simeq T_1^{\text{cCLV}} = \mathcal{O}(N)$, when $N \gg 1$ [13, 15].

It has been shown that the mean extinction/fixation time T_1^{cCLV} can be obtained from the linear approximation about \vec{x}^* [13] (see also [11, 18]). For this, it is useful to consider the FPE (S17) in polar coordinates, via $y_1 = r \cos \theta$ and $y_2 = r \sin \theta$. Since there is no angular dependence when $\vec{x}_0 = \vec{x}^*$, one has $P_{\text{cCLV}}(\vec{y}, t) \rightarrow P_{\text{cCLV}}(r, t)$ with

$$\partial_t P_{\text{cCLV}}(r, t) = D^{\text{cCLV}} [r^{-1} \partial_r + \partial_r^2] P_{\text{cCLV}}(r, \theta, t), \quad (\text{S21})$$

which is the two-dimensional diffusion equation in polar coordinates with only radial dependence and diffusion constant $D^{\text{cCLV}} = 1/(12N)$. By supplementing this FPE with an absorbing boundary at ∂S_3 , approximated as a circle of radius R in order to exploit the symmetry about \vec{x}^* , the mean extinction time was found to scale with N :

$$T_1^{\text{cCLV}} \simeq 3R^2 N, \quad \text{where} \quad R = \frac{1}{2\sqrt{3}} \left(1 + \frac{1}{\sqrt{3}} \right). \quad (\text{S22})$$

Hence, in the cCLV with equal rates ($k_i = 1$), the mean fixation and extinction time when the dynamics starts at \vec{x}^* is $T_F^{\text{cCLV}} \simeq T_1^{\text{cCLV}} \simeq 0.62N$. Qualitatively, the same conclusion $T_F^{\text{cCLV}} \simeq T_1^{\text{cCLV}} = \mathcal{O}(N)$ also holds when the rates k_i 's are unequal [13].

2 Stage 1 dynamics in the constant- K BDCLV and MCLV: similarities with the cCLV

We have seen that the cCLV survival/fixation probabilities are set in Stage 1 by the outermost orbit and follow the LOW in large populations. The MCLV and cCLV obey the same mean-field equations (up to time rescaling), with the same constant of motion \mathcal{R} and fixed points, see Eqs. (S15) and (S10), and as they admit the same outermost orbits. Furthermore, when the same timescale, the diffusion constant in the MCLV is $1/(Ns)$ and $1/N$ in the cCLV.

[§]In the cCLV, when the population size is $N = 3$, we have $\phi_i^{\text{cCLV}} = x_{i+1}^*$ [14].

the survival probabilities $\phi_{i,i+1}^{\text{MCLV}}$ of a population evolving with the MCLV are therefore expected to correspond to those of the cCLV in a population of effective size $\mathcal{O}(Ns)$, with rates related according to $r_i = k_i/(k_1 + k_2 + k_3)$. We have also seen that in the constant- K BDCLV the population size rapidly fluctuates about K , i.e. $N(t) \simeq K$, see Eq. (9) and Fig. 1, and its survival probabilities are the same as in the MCLV with $N = K \gg 1$, see Fig. S1. The survival probabilities $\phi_{i,i+1}$ in the constant- K BDCLV are therefore the same as those, $\phi_{i,i+1}^{\text{cCLV}}|_{Ks}$, in the cCLV with a population of size $\mathcal{O}(Ks)$: $\phi_{i,i+1} \approx \phi_{i,i+1}^{\text{MCLV}}|_K \approx \phi_{i,i+1}^{\text{cCLV}}|_{Ks} = \phi_i^{\text{cCLV}}|_{Ks}$. We therefore expect that the survival probabilities of the constant- K BDCLV obey the LOW when $Ks \gtrsim 100$, whereas they obey the LOSO when $Ks = \mathcal{O}(10)$, see Fig. S2. This is confirmed by the results discussed in Sec. 3.1.1, see Fig. 3 (a,b). We have also seen that the mean extinction time in the cCLV scales with N to leading order and can be obtained within a linear noise approximation about \bar{x}^* . We can proceed similarly with the MCLV, and since the linear noise approximation about \bar{x}^* of the cCLV and MCLV is similar, see Eqs. (S17) and (S11), we can obtain the mean extinction time T_1^{MCLV} by solving the radial diffusion equation $\partial_t P_{\text{MCLV}}(r, t) = D^{\text{MCLV}} [r^{-1} \partial_r + \partial_r^2] P_{\text{MCLV}}(r, \theta, t)$, with absorbing boundary on ∂S_3 and $D^{\text{MCLV}} = 2D^{\text{cCLV}}$. This yields $T_1^{\text{MCLV}} \simeq \frac{3}{2} R^2 N \approx 0.3N$ when $r_i = r = 1/3$ (symmetric rates). A similar relation, with a different expression of R , holds when the rates r_i are asymmetric. Since $N(t) \simeq K$ in the constant- K BDCLV (after a time $t = \mathcal{O}(1)$), we readily obtain its mean extinction time: $T_1 \simeq \frac{3}{2} R^2 K \approx 0.3K$ to leading order in $K \gg 1$, when $r_i = 1/3$. The insets of Fig. S1 confirm that T_1 in constant- K BDCLV is almost indistinguishable from T_1^{MCLV} obtained in the MCLV with $N = K \gg 1$. This result also holds when the dynamics towards extinction is driven by diffusion (weak demographic noise). This is certainly the case when $\bar{x}_0 = \bar{x}^*$ and also when $\bar{x}_0 \neq \bar{x}^*$ and $s \ll 1$. In fact, under weak selection, the deterministic drift arising when $\bar{x}_0 \neq \bar{x}^*$ is weak and extinction is driven by weak demographic fluctuations when $s \ll 1$. We therefore find $T_1 \simeq \frac{3}{2} R^2 N \approx 0.3N$ when $r_i = r = 1/3$ and when $s \ll 1$ and $sK = \mathcal{O}(1)$, as reported in Fig. S4(a).

3 Stage 2 dynamics in a population with constant carrying capacity

In stark contrast to the cCLV, the outcome of Stage 2 in the MCLV/BDCLV is not certain. This is because the interactions in the MCLV/BDCLV are *not* of predator-prey type: In Stage 2, the dynamics boils down to the competition between species i and its “weak opponent”, species $i + 1$, that the latter has a non-zero chance to win it.

To study this two-species competition, we focus on the stage 2 dynamics along the edge $(i, i + 1)$. Since species $i - 1$ has died out at the end of Stage 1, we have $x_i + x_{i+1} = 1$ and $x_{i-1} = 0$, and the constant- K BDCLV transition rates in Stage 2 are $T_j^+ = (1 + s\Pi_j) N x_j$ and $T_j^- = N^2 x_j / K$, with $j \in \{i, i + 1\}$, see (S2). Similarly, the transition rates of the MCLV along the edge $(i, i + 1)$ for a population of size $N = K$ are obtained from (S9) with $x_{i+1} = 1 - x_i$ and $x_{i-1} = 0$:

$$T_{i+1 \rightarrow i} = \frac{T_i^+ T_{i+1}^-}{K} = K x_i (1 - x_i) (1 + \alpha_i (1 - x_i)) \quad \text{and} \quad T_{i \rightarrow i+1} = \frac{T_i^- T_{i+1}^+}{K} = K x_i (1 - x_i) (1 - \alpha_i x_i). \quad (\text{S23})$$

It is clear from these transition rates that, $x_i = 0, 1$ are the possible outcome of the stage 2 dynamics and correspond to either the absorption of species i with probability $\phi_i|_K$ ($x_i = 1, x_{i+1} = 0$), or the absorption of $i + 1$ ($x_i = 0, x_{i+1} = 1$) with probability $1 - \phi_i|_K$.

Clearly, (S23) define a one-dimensional Moran process whose fixation properties can be computed exactly [6, 8]. For our purposes, the diffusion theory allows us to obtain a concise and reliable characterization of $\phi_i|_K$. In fact, the backward version of the FPE generator (S10) for the MCLV (with $N = K$) along the edge $(i, i + 1)$ is [1, 2]

$$\mathcal{G}|_K(x_i) \equiv \frac{x_i(1-x_i)}{K} \left[K \alpha_i \frac{\partial}{\partial x_i} + \left\{ 1 + \frac{\alpha_i}{2} (1 - 2x_i) \right\} \frac{\partial^2}{\partial x_i^2} \right].$$

When Stage 2 starts with a fraction \hat{x}_i of individuals of species i , the fixation probability $\phi_i|_K$ of the underpinning MCLV is obtained in the realm of the diffusion theory by solving $\mathcal{G}|_K(\hat{x}_i) \phi_i|_K(\hat{x}_i) = 0$ with $\phi_i|_K(0) = 0$ and $\phi_i|_K(1) = 1$. This yields

$$\phi_i|_K(\hat{x}_i) = \frac{(2 + \alpha_i)^{K+1} - \{2 + \alpha_i(1 - 2\hat{x}_i)\}^{K+1}}{(2 + \alpha_i)^{K+1} - (2 - \alpha_i)^{K+1}}.$$

When $s \ll 1$, i.e. $\alpha_i \ll 1$, the backward FPE generator takes the classical form [19, 6, 7]

$$\mathcal{G}|_K(x_i) = \frac{x_i(1-x_i)}{K} \left[K \alpha_i \frac{\partial}{\partial x_i} + \frac{\partial^2}{\partial x_i^2} \right], \quad \text{yielding the familiar expression} \quad \phi_i|_K(\hat{x}_i) = \frac{1 - e^{-\alpha_i K \hat{x}_i}}{1 - e^{-\alpha_i K}}. \quad (\text{S24})$$

In the realm of the diffusion theory, the MCLV mean absorption time T_2^{MCLV} (from the inception of Stage 2) with an initial fraction \hat{x}_i of i individuals, $T_2^{\text{MCLV}}(\hat{x}_i)$, is obtained from the FPE generator $\mathcal{G}_{(i,i+1)}|_K$ by solving $\mathcal{G}|_K(\hat{x}_i)T_2^{\text{MCLV}}(\hat{x}_i) = -1$ with boundary conditions $T_2^{\text{MCLV}}(\hat{x}_i = 0) = T_2^{\text{MCLV}}(\hat{x}_i = 1) = 0$ [1, 2, 6, 19]. The insets of Fig. S1 confirm that the mean absorption time $T_2^{(i,i+1)}$ along the edge $(i, i+1)$ in the constant- K BDCLV virtually coincide with T_2^{MCLV} when $N = K \gg 1$: $T_2^{(i,i+1)} \simeq T_2^{\text{MCLV}}$. The FPE for T_2^{MCLV} can be solved by standard methods and generally yields a cumbersome expression. In the limit of weak selection, $s \ll 1$, we can use the simpler form (S24) and find that $T_2^{(i,i+1)} \simeq T_2^{\text{MCLV}} \sim (\log K)/s$ when $s \ll 1$ and $sK \gg 1$. In this case, $T_2^{(i,i+1)}$ scales as $1/s$ to leading order, with a subleading dependence on the population size via the prefactor $\log K$. When $s \ll 1$ and $sK \lesssim 1$, $T_2^{(i,i+1)} = \mathcal{O}(K)$ at quasi-neutrality, while $T_2^{(i,i+1)} \sim \log K$ when $sK \gg 1$ (strong selection), as shown in Fig. S4 (b).

A similar analysis can be carried out when $\epsilon > 0$, see Section 5. In this case, the non-zero-sum birth-death dynamics defined by (6) with constant carrying capacity $K \gg 1$ is similar to the dynamics of Moran model defined by the transition rates

$$\begin{aligned} T_{i+1 \rightarrow i} &= \frac{T_i^+ T_{i+1}^-}{K} = K x_i (1 - x_i) (1 + \alpha_i (1 - x_i)) \quad \text{and} \\ T_{i \rightarrow i+1} &= \frac{T_i^- T_{i+1}^+}{K} = K x_i (1 - x_i) (1 - \alpha_i (1 + \epsilon) x_i). \end{aligned} \quad (\text{S25})$$

in a population of constant size $N = K$. In this case, the stage 2 dynamics along the edge $(i, i+1)$ is characterized by the backward FPE generator $\mathcal{G}^\epsilon|_K(x_i) \equiv \frac{x_i(1-x_i)}{K} [K\alpha_i(1+\epsilon x_i)\partial_i + \{1 + \frac{\alpha_i}{2}(1 - (2+\epsilon)x_i)\}\partial_i^2]$. The absorption probability is thus obtained by solving $\mathcal{G}^\epsilon|_K(\hat{x}_i)\phi_i|_K(\hat{x}_i) = 0$ with $\phi_i|_K(0) = 1 - \phi_i|_K(1) = 0$ yielding

$$\phi_i|_K(\hat{x}_i) = \frac{(2 + \alpha_i)^{Kh(\epsilon, \alpha_i)+1} - \{2 + \alpha_i(1 - (2 + \epsilon)\hat{x}_i)\}^{Kh(\epsilon, \alpha_i)+1}}{(2 + \alpha_i)^{Kh(\epsilon, \alpha_i)+1} - (2 - \alpha_i(1 + \epsilon))^{Kh(\epsilon, \alpha_i)+1}}, \quad \text{where} \quad h(\epsilon, \alpha_i) \equiv \frac{1 + \epsilon(1 + 1/\alpha_i)}{(1 + \epsilon/2)^2}. \quad (\text{S26})$$

When $0 < \epsilon \ll 1$, this expression simplifies in the weak selection regime ($s \ll 1$) where it takes the form

$$\phi_i|_K(\hat{x}_i) \simeq \frac{1 - e^{-K\alpha_i(1+\epsilon/2)\hat{x}_i}}{1 - e^{-K\alpha_i(1+\epsilon/2)}}. \quad (\text{S27})$$

Hence, the stage 2 fixation probability in the weak selection regime when $0 < \epsilon \ll 1$ is the same as in the constant- K BDCLV with a selection intensity is rescaled by a factor $1 + \epsilon/2$ ($s \rightarrow s(1 + \epsilon/2)$). This suggests to consider the following effective backward FPE generator when $s \ll 1$ and $\epsilon \ll 1$: $\mathcal{G}^\epsilon|_K(x_i) \equiv \frac{x_i(1-x_i)}{K} [K\alpha_i(1 + \epsilon/2)\partial_i + \partial_i^2]$. In this case, the mean absorption time is given $\mathcal{G}^\epsilon|_K(\hat{x}_i)T_2(\hat{x}_i) = -1$ with $T_2(\hat{x}_i = 0) = T_2(\hat{x}_i = 1) = 0$. Clearly, this implies that the mean absorption time is obtained from T_2 of the constant- K BDCLV with a rescaled selection intensity $s \rightarrow s(1 + \epsilon/2)$. We have checked in our simulations that this rescaling also applies to Stage 1 and therefore to the overall mean fixation time, see Sec. 5.4 of this SM and Fig. S6 (b,c).

4 Population composition at the inception of Stage 2

The stage 2 dynamics of the BDCLV and MCLV, as well as their fixation properties, depend on the population composition at the end of Stage 1 which coincides with the inception of Stage 2. In the main text, we have seen that the initial fraction \hat{x}_i of i individuals along the edge $(i, i+1)$ of S_3 is given by the probability density $P_{(i,i+1)}(\hat{x}_i)$ which can be approximated by a uniform distribution $P_{(i,i+1)}(\hat{x}_i) \approx 1$ when $sK \lesssim 10$ (constant K) and $s(K) \lesssim 10$ (switching K), yielding an average initial fraction $\mu_i = \int_0^1 \hat{x}_i P_{(i,i+1)}(\hat{x}_i) d\hat{x}_i \approx 1/2$ of i individuals along $(i, i+1)$, see Fig. S3. The same holds true also when $0 < \epsilon \ll 1$, see Sec. 5 in the main text and Sec. 3 of this SM.

This is no longer the case under strong selection, when the $P_{(i,i+1)}$'s are skewed and far from being uniform, see the lower insets of Fig. S3. When $K \gg 1$ is constant and the LOW holds, the extinction of the first species in Stage 1 occurs from the outermost orbit as in the cCLV [14, 15], see Sec. 1.3 of this SM, and μ_i can be estimated as follows: Along the outermost orbit that is closest ($x_{i-1} = 1/K$) to the edge $(i, i+1)$ in the constant- K BDCLV, from the rate equations (9) we have $x_i/x_{i+1} = r_{i+1}/r_{i-1}$ yielding $\mu_i = r_{i+1}/(r_{i+1} + r_{i-1})$. The results of Fig. S3 (a,b) for $sK \gg 1$ are in satisfying agreement with this prediction.

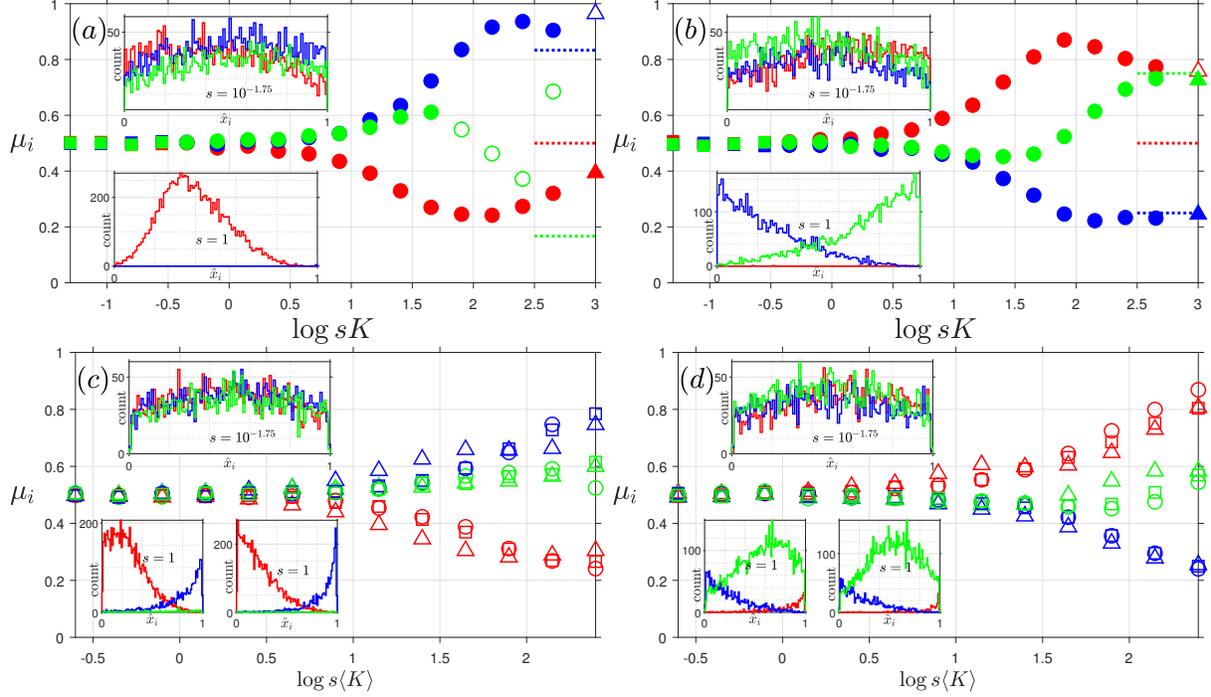


Figure S3: Population composition at the inception of Stage 2 vs. sK (a,b) and $s\langle K \rangle$ (c,d) with $\bar{r} = \bar{r}^{(1)}$ in (a,c) and $\bar{r} = \bar{r}^{(2)}$ in (b,d). In all panels: $\mu_i = \int_0^1 \hat{x}_i P_{(i,i+1)}(\hat{x}_i) d\hat{x}_i$ is the mean value of \hat{x}_i for species $i = 1$ (red), 2 (blue), 3 (green), with $\bar{x}_0 = \bar{x}_c$ and $\epsilon = 0$. (a,b) μ_i vs. sK in the constant- K BDCLV with $K = 1000$ (Δ), 450 (\circ), 50 (\square) and $s \in (10^{-3}, 1)$. (Empty symbols denote data arising from small survival probability $\phi_{i,i+1} < 0.01$ that would require additional sampling). When $sK \lesssim 10$, $\mu_i \approx 1/2$ and $P_{(i,i+1)} \approx 1$ is approximately uniform. When $sK \gg 1$, the dynamics is dominated by the LOW and $\mu_i \approx r_{i+1}/(r_{i+1} + r_{i-1})$ shown as dotted lines, see text. Upper insets: Histograms corresponding to $P_{(i,i+1)}(\hat{x}_i)$ with $s = 10^{-7/4}$ and $K = 250$, is approximately uniform, corresponding to $P_{(i,i+1)} \approx 1$, along the three edges. Lower insets: Same with $s = 1$ and $K = 1000$, showing that $P_{(i,i+1)}$ is no longer uniform when $sK \gg 1$ and how it changes with $\nu = 10$ (left) and $\nu = 0.1$ (right). (c,d) μ_i vs. $s\langle K \rangle$ in the switching- K BDCLV with $\langle K \rangle = 250$ and $\gamma = 0.8$ kept fixed and s varies with $\nu = 10$ (\square), $\nu = 1$ (\circ) and $\nu = 0.001$ (Δ). Insets: (Upper) Histograms corresponding to $P_{(i,i+1)}(\hat{x}_i)$ with $s = 10^{-7/4}$, $\nu = 0.1$ and $\langle K \rangle = 250$, $\gamma = 0.8$ for $i = 1, 2, 3$. (Lower) Same with $s = 1$, $\langle K \rangle = 250$, $\gamma = 0.8$, $\nu = 0.1$ (left) and $\nu = 10$ (right).

The results reported in Fig. S3 (c,d) show that the averages μ_i 's are closer to 1/2 in regime (ii) than in the constant- K BDCLV. This stems from the environmental variability operating to balance the effect of selection and implies that $P_{(i,i+1)} \approx 1$ is a better approximation in the regime (ii) when K is randomly switching than when it is constant. In the lower insets of Fig. S3 (c,d), we find very similar probability densities $P_{(i,i+1)}$ for very different switching rates ($\mu = 0.1$ and $\mu = 10$), showing that in the switching- K BDCLV $P_{(i,i+1)}$ varies little with ν .

5 Extinction, absorption and fixation times & number of switches

We study the overall mean fixation time T_F , which is the average time after which one species takes over the entire population, in the constant- K and switching- K BDCLV. $T_F = T_1 + T_2$ consists of the *mean extinction time* T_1 and the *mean absorption time* T_2 arising from Stages 1 and 2, respectively. We also compute the average number of switches occurring in Stages 1 and 2 of the switching- K BDCLV.

5.1 Mean extinction, absorption and fixation times in the constant- K BDCLV

We first consider the case of the constant- K BDCLV and show that the overall mean fixation time $T_F = \mathcal{O}(K)$ across all regimes (i)-(iii), see Fig. S4(a).

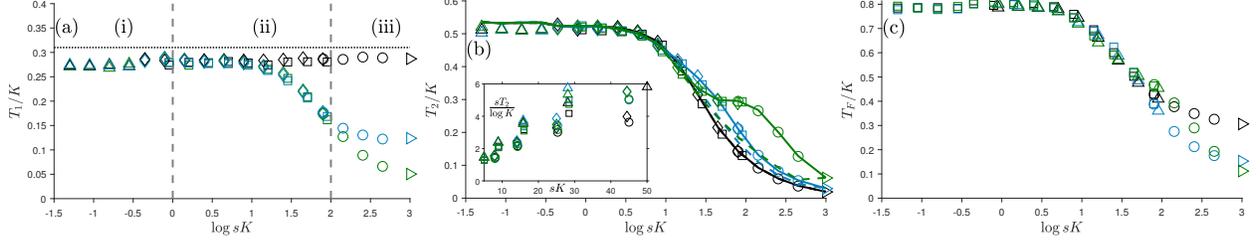


Figure S4: Mean extinction and absorption times T_1 and T_2 , and mean fixation time T_F in the constant- K BDCLV for $K \in \{1000 (\triangleright), 450 (\circ), 250 (\diamond), 90 (\square), 50 (\triangle)\}$ and the same values of s as in Figs. 3 and 4: (a) T_1/K vs. sK ; showing $T_1 = \mathcal{O}(K)$ when $K \gg 1$ and $T_1 \approx 0.31K$ (dotted line) when $r_i = r$ and under weak selection ($sK \lesssim 10$) when $\vec{x}_c \neq \vec{x}^*$ (unequal r_i 's), see text. (b) T_2/K vs. sK ; solid and dashed lines show the respective predictions of $T_2|_K = \sum_i \phi_{i,i+1} T_2^{(i,i+1)}|_K$ and (S29), see text. Inset: $sT_2/\log K = \mathcal{O}(1)$ when $s \ll 1$ and $sK \gg 1$, see text. (c) T_F/K vs. sK showing that $T_F = \mathcal{O}(K)$ across all regimes with subleading prefactors in regime (iii) shorter than in (i) and (ii). In all panels: symbols are from stochastic simulations, $\vec{x}_0 = \vec{x}_c$, $\epsilon = 0$ and $r_1 = 1/11$ (green), 1 (black), $3/5$ (blue) and $r_2 = r_3 = 1$.

5.1.1 Stage 1: Mean extinction time T_1 in the constant- K BDCLV

The mean extinction time T_1 is the average time for one of the species to go extinct at the end of Stage 1. As explained in Sec. 2 of this SM, with the results obtained for the cCLV, we find $T_1 \simeq T_1^{\text{cCLV}}/2 \approx 0.3K$ when $s \ll 1$ (regimes (i,ii)) and for arbitrary s when all $r_i = 1/3$, see Fig. S4 (a). Deviations from $T_1 \approx 0.3K$, and a weak dependence on s and on the r_i 's, are found near the boundary of regimes (ii)-(iii) and in regime (iii), where $T_1 \simeq \beta_c(s, \vec{r})K$, where β_c is a decreasing function of s when the r_i 's are unequal, see Fig. S4 (a).

5.1.2 Stage 2: Mean absorption time T_2 in the constant- K BDCLV

The stage 2 mean absorption time T_2 is given by

$$T_2 = \sum_{i=1}^3 \phi_{i,i+1} T_2^{(i,i+1)}, \quad (\text{S28})$$

where the mean absorption time along the edge $(i, i+1)$ of S_3 , denoted by $T_2^{(i,i+1)}$, is weighed by the probability $\phi_{i,i+1}$ that Stage 1 ends on that edge.

The expression of $T_2^{(i,i+1)}$ is obtained from the mean fixation time of the MCLV with $N = K$, here denoted by $T_2^{(i,i+1)}|_K$ with $T_2^{(i,i+1)} \simeq T_2^{(i,i+1)}|_K$, see Sec. 3 of this SM. For a given initial fraction \hat{x}_i of i 's at the start of Stage 2 is (\hat{x}_i) , $T_2^{(i,i+1)}(\hat{x}_i)|_K$ when $s \ll 1$ is obtained by solving $\mathcal{G}_{(i,i+1)}|_K(\hat{x}_i) T_2^{(i,i+1)}(\hat{x}_i)|_K = -1$, with $T_2^{(i,i+1)}|_K(0) = T_2^{(i,i+1)}|_K(1) = 0$, see (S24). Since the exact population composition along the edge $(i, i+1)$ at the inception of Stage 2 is given by $P_{(i,i+1)}(\hat{x}_i)$, we have:

$$T_2 \simeq \sum_{i=1}^3 \phi_{i,i+1} T_2^{(i,i+1)}|_K = \sum_{i=1}^3 \phi_{i,i+1} \int_0^1 P_{(i,i+1)}(\hat{x}_i) T_2^{(i,i+1)}(\hat{x}_i)|_K d\hat{x}_i,$$

with $T_2^{(i,i+1)}|_K \equiv \int_0^1 P_{(i,i+1)}(\hat{x}_i) T_2^{(i,i+1)}(\hat{x}_i)|_K d\hat{x}_i$. A simpler expression for T_2 is obtained when $s \ll 1$ and $sK = \mathcal{O}(10)$ upon substituting $\phi_{i,i+1} \approx 1/3$ and $P_{(i,i+1)} \approx 1$ in (S29):

$$T_2 \simeq \frac{1}{3} \sum_{i=1}^3 \int_0^1 T_2^{(i,i+1)}(\hat{x}_i)|_K d\hat{x}_i. \quad (\text{S29})$$

While the expression of $T_2^{(i,i+1)}(\hat{x}_i)$ is not particularly illuminating, its asymptotic behavior is simple and allows us to determine the behavior of T_2 : In the weak-selection regime (ii) where $s \ll 1$ and $sK = \mathcal{O}(10)$, we obtain the classical

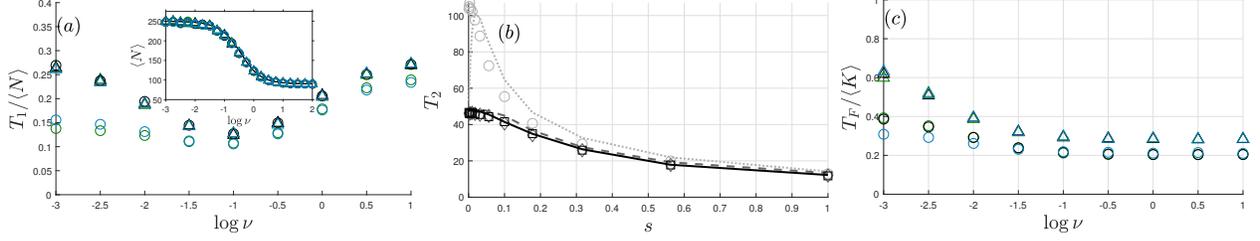


Figure S5: (a) $T_1/\langle N \rangle$ vs. ν for $r_1 = 1/11$ (green), $1/3$ (black), $3/5$ (blue) and $r_2 = r_3 = (1 - r_1)/2$, with $s = 10^{-1/2}$ (circles) and $s = 10^{-3/2}$ (triangles). In agreement with (S30), $T_1/\langle N \rangle = \beta_s = \mathcal{O}(1)$ and slowly varies with ν and s . Inset: $\langle N \rangle$ vs ν ; solid lines are from the average over the marginal probability density (21) of the process defined by (S8) and symbols are from stochastic simulations with $s = 10^{-1/2}$ (circles) and $s = 10^{-3/2}$ (triangles), showing $\langle N \rangle = \mathcal{O}(\langle K \rangle)$, see text. (b) T_2 vs. s for $\nu = 10^{-3}$ (circles, light dotted gray), 10^{-1} (diamonds, dashed gray), 10 (squares, solid black) and $\vec{r} = (1/3, 1/3, 1/3)$. Symbols are from stochastic simulations and lines are from (S31). T_2 scales as $1/s$ with subleading prefactor $\sim \log \langle K \rangle$ when $s \ll 1$ and $s \langle K \rangle = \mathcal{O}(10)$, see text. (c) Same as in (a) but for the overall mean fixation time: $T_F/\langle K \rangle$ vs. ν with $s = 10^{-1/2}$ (circles) and $s = 10^{-3/2}$ (triangles), showing $T_F = \mathcal{O}(\langle K \rangle)$ over a broad range of values ν , see text. In all panels: $\langle K \rangle = 250$, $\gamma = 0.8$ ($K_- = 50$, $K_+ = 450$) and $\vec{x}_0 = \vec{x}_c$; $\epsilon = 0$.

result $T_2^{(i,i+1)}|_K = \mathcal{O}((\log K)/s)$ according to which T_2 scales as $1/s$ with a subleading prefactor $\sim \log K$ [6, 7], which is confirmed by the results of Fig. S4 (c).

On the other hand, since the mean fixation time in the neutral Moran model scales linearly with the population size [6, 19, 7], we readily find $T_2 = \mathcal{O}(K)$ in the quasi-neutral regime (i). The mean fixation time in the Moran model with strong selection favoring species i against $i + 1$ scales logarithmically with the population size [8], from which we infer that $T_2 = \mathcal{O}(\log K)$ in regime (iii).

Putting the asymptotic behaviors of T_1 and T_2 together, we find that to leading order in $N \simeq K \gg 1$ the overall mean fixation time $T_F = T_1 + T_2 = \mathcal{O}(K)$ scales linearly with the population size across the regimes (i)-(iii), with different subleading prefactors in each regime. We also notice that in regime (iii) $T_1 \gg T_2$: The extinction of a second species (Stage 2) occurs much faster than the death of a first species in Stage 1, see Fig.1 (a). In regime (i) $T_1/T_2 = \mathcal{O}(1)$ and $T_1/T_2 = \mathcal{O}(sK/\log K)$ in regime (ii), see Fig.1 (b).

5.2 Mean extinction, absorption and fixation times in the switching- K BDCLV

We study the effect of random switching on the mean extinction and absorption times, T_1 and T_2 characterizing Stages 1 and 2, respectively. This allows us to show that the mean fixation time $T_F = T_1 + T_2 = \mathcal{O}(\langle N \rangle) = \mathcal{O}(\langle K \rangle)$ scales linearly with the average population size, and to compute the average number of switches occurring in Stages 1 and 2.

5.2.1 Stage 1: Mean extinction time in the switching- K BDCLV

Guided by the results of the constant- K BDCLV, where T_1 scales linearly with $N \approx K$ to leading order in $\langle K \rangle \gg 1$, we expect

$$T_1 = \beta_s \langle N \rangle \quad \text{with} \quad \beta_s = \beta_s(s, \vec{r}, \nu), \quad (\text{S30})$$

where $\langle N \rangle = \mathcal{O}(\langle K \rangle)$ is the long-time average population size that is in principle obtained by averaging N over the N -QSD. In the inset of Fig. S5, this quantity is accurately computed in the realm of the piecewise deterministic Markov process approximation as $\langle N \rangle = \int_{K_-}^{K_+} N p_\nu^*(N) dN$, see the inset of Fig. S5 (a), and is shown to be independent of s and a decreasing function of ν . For fast/slow switching, we have $\langle N \rangle = (1 - \gamma^2)\langle K \rangle$ when $\nu \rightarrow \infty$ and $\langle N \rangle = \langle K \rangle$ when $\nu \rightarrow 0$ [20, 21]. Comparison with simulation results of Fig. S5 confirm that $T_1/\langle N \rangle = \beta_s = \mathcal{O}(1)$ is a slowly varying function of ν and a weakly decreasing function of s . Since $\langle N \rangle = \mathcal{O}(\langle K \rangle)$ when $\gamma = \mathcal{O}(1)$, we obtain $T_1 = \mathcal{O}(\langle N \rangle) = \mathcal{O}(\langle K \rangle)$ to leading order in $\langle K \rangle$.

5.2.2 Stage 2 mean absorption time and overall mean fixation time in the switching- K BDCLV

Proceeding as in Sec. 5.1.2 of this SM, the Stage 2 mean absorption time is given by $T_2 = \sum_{i=1}^3 \phi_{i,i+1} T_2^{(i,i+1)}$. In the realm of the piecewise deterministic Markov process approximation, when $s \ll 1$ and $s\langle K \rangle \gg 1$, $T_2^{(i,i+1)}$ is obtained by averaging the constant- $\langle K \rangle$ mean absorption time $T_2^{(i,i+1)}|_{\langle K \rangle}$ along the edge $(i, i+1)$ over the probability density function (21) [20, 21]:

$$T_2^{(i,i+1)} \simeq \int_0^1 \int_{K_-}^{K_+} P_{(i,i+1)}(\hat{x}_i) T_2^{(i,i+1)}(\hat{x}_i)|_{\langle K \rangle} p_{\nu/\alpha_i}^*(N) d\hat{x}_i dN.$$

As in Sec. 4.1.2 in the main text, the switching rate is rescaled $\nu \rightarrow \nu/\alpha_i$ due to the average number $\mathcal{O}(\nu/\alpha_i)$ of switches occurring in Stage 2 along the edge $(i, i+1)$ when $s \ll 1$ and $s\langle K \rangle \gg 1$ [20, 21]. The above equation can be simplified using $\phi_{i,i+1} \approx 1/3$ and $P_{(i,i+1)}(\hat{x}_i) \approx 1$ when $s \ll 1$ and $s\langle K \rangle \lesssim 10$ (see Sec. 4 of this SM):

$$T_2 \approx \frac{1}{3} \sum_{i=1}^3 T_2^{(i,i+1)} \simeq \frac{1}{3} \sum_{i=1}^3 \int_{K_-}^{K_+} T_2^{(i,i+1)}(\hat{x}_i)|_{\langle K \rangle} p_{\nu/\alpha_i}^*(N) dN, \quad (\text{S31})$$

where $T_2^{(i,i+1)} \sim T_2^{(i,i+1)}|_{\langle K \rangle}(\hat{x}_i)$ which scales as $1/\alpha_i$ with a prefactor $\sim \log \langle K \rangle$ and a weak dependence on ν when $s \ll 1$ and $s\langle K \rangle \gg 1$ [20]. This yields $T_2^{(i,i+1)} = \mathcal{O}((\log \langle K \rangle)/s)$ in regime (ii): In agreement with the results of Fig. S5 (b), $T_2 = \mathcal{O}(1/s)$ with a subleading prefactor $\sim \log \langle K \rangle$ when $s \ll 1$ and $s\langle K \rangle \lesssim 10$. As in the constant- K BDCLV, the quasi-neutral regime (i), where $s\langle K \rangle \ll 1$, $T_2 = \mathcal{O}(\langle K \rangle)$, whereas under strong selection, $s\langle K \rangle \gg 1$, $T_2 = \mathcal{O}(\log \langle K \rangle)$, see Fig. S5 (b).

Putting together the results for T_1 and T_2 , we obtain the overall mean fixation time $T_F = T_1 + T_2 \sim \langle N \rangle$. Since $\langle N \rangle = \mathcal{O}(\langle K \rangle)$, we have $T_F = \mathcal{O}(\langle K \rangle)$ which, with subleading prefactors that vary slowly with ν and s , as illustrated by Fig. S5(c).

5.3 Average number of switches in Stages 1 and 2 of the switching- K BDCLV

Since the average duration of Stage 1 in the the switching- K BDCLV is $T_1 = \beta_s \langle N \rangle = \mathcal{O}(\langle K \rangle)$, see Eq. (S30), the average number of switches occurring prior one of the species die out scales as $\mathcal{O}(\nu \langle K \rangle)$, as shown in Fig. S6 (a), i.e. the average number of switches increases as $\nu \langle K \rangle$, with a prefactor that depends on s via β_s which is a weakly decreasing function of s (i.e. the number of switches is greater for smaller values of s). Hence, for any non-vanishingly small switching rate $\nu \gg 1/\langle K \rangle$ and $\langle K \rangle \gg 1$, a large number of switches occur during Stage 1 prior to the extinction of the first species and the DMN self averages, see Sec. 4.1.1.

In Refs. [20, 21], it has been shown that that under weak selection the population experiences, on average, $\mathcal{O}(\nu/\alpha_i)$ switches during the two-species competition characterizing the stage 2 dynamics along the edge $(i, i+1)$. This supports the rescaling $\nu \rightarrow \nu/\alpha_i$ in formula (26) which has been found to be actually valid when the selection intensity s is neither vanishingly small nor too large [21].

5.4 Mean fixation time of a close-to-zero-sum rock-paper-scissors game in fluctuating populations

The mean fixation time of the close-to-zero-sum rock-paper-scissors game ($|\epsilon| \ll 1$) under weak selection can be obtained with a similar argument used in Sec. 5 for the fixation probabilities. In fact, the mean absorption time T_2 and the mean fixation time $T_F = T_1 + T_2$ (T_1 varies little with s in regime (ii), see Fig. S4 (a)) under weak selection can be obtained from their values in the BDCLV with a rescaled selection intensity $s \rightarrow s(1 + (\epsilon/2))$, as shown in Fig. S6 (b)-(d). This is valid both for the case of a constant K , see Fig. S6 (b), and a randomly switching carrying capacity, see Fig. S6 (c,d). This confirms that the effect of $0 < \epsilon \ll 1$ on the fixation properties simply boils down to increasing the selection intensity by a factor $1 + (\epsilon/2)$ with respect to the BDCLV when sK and $s\langle K \rangle$ are in regimes (i) and (ii). When $sK \gg 1$ and $s\langle K \rangle \gg 1$ (regime (iii)), the above argument breaks down and rescaling the the selection intensity of the BDCLV's mean fixation time is no longer a good approximation: Under strong selection, the actual T_F is systematically overestimated and underestimated by the $s \rightarrow s(1 + (\epsilon/2))$ rescaling when $\epsilon > 0$ and $\epsilon < 0$. Deviations from the rescaled BDCLV results are particularly pronounced under strong selection in the case $\epsilon < 0$ and $\vec{x}_c = \vec{x}^*$ (with $r_i = 1/3$).

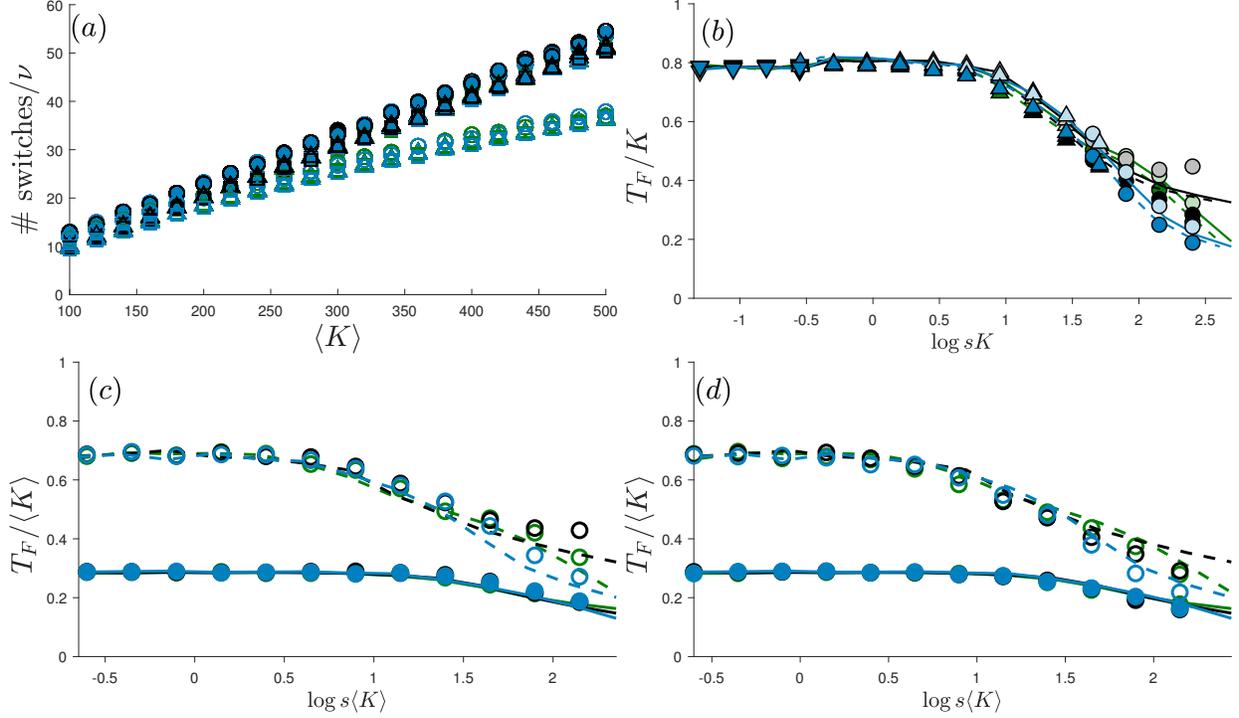


Figure S6: (a) Average number of switches in Stage 1 of the BDCLV for $\nu = 0.1$ (circles), 1 (triangles), 10 (squares). Selection intensity is $s = 10^{-3/2}$ (filled symbols) and $s = 10^{-1/2}$ (open symbols). Data for (average number of switches in Stage 1)/ ν vs $\langle K \rangle$ and different values of ν and \vec{r} essentially collapse onto a curve (almost a line). (b) Rescaled mean fixation time T_F/K vs. sK in the close-to-zero-sum game ($|\epsilon| \ll 1$) and constant K for values of $s \in (10^{-3}, 1)$ and $K = 450$ (circles), 90 (upward triangles), 50 (downward triangles). Symbols are from stochastic simulations for $\epsilon = -0.2$ (light) and $\epsilon = 0.2$ (dark). Lines are from the constant- K BDCLV obtained with the same carrying capacity but a rescaled selection intensity $s(1 + \epsilon/2)$. (c) $T_F/\langle K \rangle$ vs. $s\langle K \rangle$ when K switches between $K_- = 50$ and $K_+ = 450$ with $s \in (10^{-3}, 1)$, and $\nu = 10$ (closed symbols) and $\nu = 0.001$ (open symbols). Symbols are from stochastic simulations obtained for $\epsilon = -0.2$; solid ($\nu = 10$) and dashed ($\nu = 0.001$) lines are from the switching- K BDCLV obtained with the same $K(t)$ but selection intensity $s(1 + \epsilon/2) = 0.9s$. (d) Same as in panel (c) with $\epsilon > 0$: Symbols are stochastic simulation results for $\epsilon = 0.2$; solid ($\nu = 10$) and dashed ($\nu = 0.001$) lines are results from the BDCLV with same switching carrying capacity and selection intensity $s \rightarrow s(1 + \epsilon/2) = 1.1s$. In all panels: $\vec{r} = (1, 1, 1)/3$ (black), $\vec{r} = (1, 5, 5)/11$ (green), $\vec{r} = (3, 2, 2)/5$ (blue); $\vec{r} = \vec{r}^{(1)}$ and $\vec{x}_0 = \vec{x}_c$. In panels (a) and (b): dark symbols and solid lines are for $\epsilon = 0.2$, light symbols and dashed lines are for $\epsilon = -0.2$.

References

- [1] C. Gardiner, *Handbook of Stochastic Methods* (Springer, New York, U.S.A., 2002).
- [2] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam, Netherland, 1992).
- [3] P. A. P. Moran, *The statistical processes of evolutionary theory*. (Clarendon, Oxford, 1962).
- [4] J. Hofbauer and K. Sigmund, K., *Evolutionary games and population dynamics* (Cambridge University Press, U.K., 1998).
- [5] J. Maynard Smith, *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, U.K.).
- [6] W. J. Ewens, *Mathematical Population Genetics* (Springer, New York, 2004).
- [7] R. A. Blythe and A. J. McKane, *J. Stat. Mech.* **P07018** (2007).
- [8] T. Antal and I. Scheuring, *Bull. Math. Biol.* **68**, 1923 (2006).

- [9] R. M. Nowak, *Evolutionary Dynamics* (Belknap Press, Cambridge, USA, 2006).
- [10] J.-C. Claussen and A. Traulsen, *Phys. Rev. Lett.* **100**, 058104 (2008).
- [11] M. Mobilia, *J. Theor. Biol.* **264**, 1 (2010).
- [12] T. Galla, *J. Theor. Biol.* **269**, 46 (2011).
- [13] T. Reichenbach, M. Mobilia and E. Frey, *Phys. Rev. E* **74**, 051907 (2006).
- [14] M. Berr, T. Reichenbach, M. Schottenloher and E. Frey, *Phys. Rev. Lett.* **102**, 048102 (2009).
- [15] R. West, M. Mobilia and A. M. Rucklidge, *Phys. Rev. E* **97**, 022406 (2018).
- [16] K. I. Tainaka, *Phys. Lett. A* **176**, 303 (1993).
- [17] M. Frean; E. D. Abraham, *Proc. R. Soc. Lond. B* **268**, 1323 (2001).
- [18] A. Dobrinevski and E. Frey, *Phys. Rev. E* **85**, 051903 (2012).
- [19] J. F. Crow and M. Kimura, *An Introduction to Population Genetics Theory* (Blackburn Press, New Jersey, 2009).
- [20] K. Wienand, E. Frey, and M. Mobilia, *Phys. Rev. Lett* **119**, 158301 (2017).
- [21] K. Wienand, E. Frey, and M. Mobilia, *J. R. Soc. Interface* **15**, 20180343 (2018).