The critical dependence of a population’s long-term behaviour on its initial conditions—the founder effect—is an important concept in population genetics and classical example of non-linear dynamics. It has been used both to argue for the relevance of population fracture in speciation, and to explain novel behaviours at the invasion front of range expansions. By modifying an established model of population dynamics, we now demonstrate that founder effects can be driven by growth in well-mixed populations. Such effects arise due to intrinsic fluctuations that decrease with the inverse square root of the population size, $N$. Whilst this guarantees that a population fixates on its deterministic attractors as $t \to \infty$, the likelihood of fixation can still depend on a population’s founders, both by number and demographic mix. This dependence can moreover be non-trivial, since decreasing $O(1/\sqrt{N})$ fluctuations couple dynamically to the geometry of the underlying attractors, giving rise to distinct regimes of demographic behaviour that change abruptly at critical population sizes. By changing the character and relative duration of such dynamical demographic phases, fluctuation-mediated founder effects can also depend on the rates of growth and mutation.

Established in the 1950s, the term “founder effect” was coined to describe how population fracture can promote speciation and/or reduced genetic variation. In the canonical example, a daughter population of nearly homogenous genetic mix is split from an otherwise genetically diverse parent. The action of random genetic drift then leads to divergent outcomes for the two populations: the daughter population is driven towards homogeneity, whilst the parent remains heterogeneous.

Such behaviour is an early articulation of how non-linear effects can manifest in population dynamics, with long term fixation depending critically on a population’s initial conditions [3].

More recently, a spatial analogue of this effect—termed the serial founder effect [4]—has also been shown to be of relevance, characterising the invasion front of range expansions [5] where there is a limited (or zero) rate of spatial mixing. The notion underpins a variety of phenomena, including: the ‘surfing’ of rare mutants into newly occupied territories [6], the structuring of newly colonized areas into distinct sectors [7], and the introgression of local genes into the genome of an invading species [8].

Employing a growing variant of an otherwise well-studied evolutionary game—the Iterated Prisoner’s Dilemma (IPD) with replicator-mutator dynamics [9]—we now demonstrate that decreasing intrinsic fluctuations in growing populations can lead to an additional dependence of fixation probabilities on initial conditions.

These effects occur in populations that are well-mixed, where spatial extent does not play a role, and are over and above those associated with non-linearities in the deterministic dynamics.

Mechanistically, we say that such effects are fluctuation-mediated, since ergodicity is broken by an intrinsic noise whose standard deviation decreases with the inverse square root of the growing population size, $N$. This implies that a population’s demographics will eventually fixate upon the stable attractors associated with its deterministic, $N \to \infty$ limit. However, time-time correlations can be still be non-vanishing, even as $t \to \infty$ [11], and therefore fixation probabilities can depend on initial conditions.

The manner by which this manifests can be shown to be non-trivial, since decreasing $O(1/\sqrt{N})$ fluctuations couple dynamically to the geometry of the aforementioned deterministic attractors, which results in dynamical demographic phases: a series of distinct regimes of demographic behaviour characterised by abrupt transitions at critical population sizes. The relative duration of such phases moreover depends on both the rates of population growth and mutation. The former by an overall rescaling and the latter as a result of a hitherto unrecognised antagonism between exogenous noise due to mutation, and intrinsic noise due to finite size fluctuations.

**RESULTS**

**Growing frequency-dependent IPD**

The frequency-dependent IPD is a well-established evolutionary game. It abstracts the key facets of biolog-
FIG. 1. Decreasing intrinsic fluctuations in the growing Iterated Prisoner’s Dilemma. Projecting the demographics of the growing IPD (panel a) onto the unit simplex (panel b, inset) demonstrates $O(1/\sqrt{N})$ intrinsic fluctuations that decrease as the population grows in time. Since overall population growth is exponential (with rate $b − d$) the system converges on canonical deterministic dynamics as $t \to \infty$ (panel b, colours represent the direction of deterministic flows). This results in only three different evolutionary outcomes. Either, the population goes extinct (panel a in black, panel c), or the system fixates on one of the two deterministic attractors: the mixed-strategy limit cycle (panels a & b in blue, panel d) or the AllD fixed point (panels a & b in red, panel e). The statistics of these two latter outcomes are dictated by fluctuations at early times, which determine the likelihood of crossing the separatrix that marks the boundary between the two basins of the stable attractors (dashed, panel b). Lower-case font represents population fractions, e.g., $n^\text{AllC} = N^\text{AllC}/N$.

Choosing $b > d$ gives rise to unbounded exponential (i.e., Malthusian) population growth (Fig. 1b). This rapidly suppresses $O(1/\sqrt{N})$ fluctuations and converges to well-studied deterministic behaviour (represented on the unit simplex in Fig. 1a, inset). In particular, for $10^{-7.5} \leq \mu \leq 10^{-2.5}$, there are two stable deterministic attractors: a stable ‘AllD’ fixed point, where a small fraction of TFT players (who mutually cooperate) are exploited by a large population of AllD defectors, and; a stable limit cycle around an unstable ‘mixed-strategy’ fixed point [9] (Fig. 1b & Appendix B). The latter is characterised by a three-phase cycle whose handedness is anti-clockwise in the traditional presentation of the state-space simplex; players of TFT can out-compete those playing AllD due to their capacity for mutual cooperation, however: they are then susceptible to invasion by players of AllC due to the complexity cost, whereby; AllC players can be exploited by those playing AllD, completing the cycle.

As a consequence, only three outcomes are possible as $t \to \infty$. Either i) the population goes extinct in the early stages due to finite size fluctuations (Figs. 1b-black & c) or, its demographic mix converges on ii) the mixed-strategy limit cycle (Figs. 1b-blue & d) or iii) the AllD fixed point (Figs. 1b-red & e).
FIG. 2. Fixation probabilities depend non-trivially on initial conditions, as well as the rates of growth and mutation. For a range of initial conditions, growth and mutation rates, we used repeated Gillespie-Itô simulations to compute $p_{\text{AID}}$, the probability that a growing population will fixate on the AllD fixed point. Choosing the initial demographic mix as the mixed-strategy fixed point, we see that $p_{\text{AID}}$ depends on both the rate of population growth and the rate of mutation (panel a). Conversely, fixing the rates of growth and mutation, $p_{\text{AID}}$ also depends on both the initial demographic mix $n_0$ (panels b, c, e & f) and the initial population size, $N_0$ (panel d & g). Behaviour can be qualitatively classified into two regions separated by the red dashed line (panel a). In the upper region, behaviour agrees with expectations: higher rates of growth more rapidly suppress $O(1/\sqrt{N})$ intrinsic fluctuations and therefore increase the likelihood of fixation within the basin of attraction that the system started (panel a). Increasing the initial population size reduces the likelihood of large fluctuations at early times, therefore exacerbating this effect (panels b-d). In the lower region, behaviour is more complex: there are high likelihoods of an AllD outcome at both low ($b - d \leq 0.01$) and intermediate ($0.03 \leq b - d \leq 0.07$) rates of growth (panel a). This results in a non-monotonic dependence of $p_{\text{AID}}$ on $b - d$, which implies that population growth can actually increase the likelihood of crossing the separatrix between the two basins of attraction, despite more rapidly suppressing $O(1/\sqrt{N})$ fluctuations. There is a similarly non-trivial structure to the effects of initial demographic mix on $p_{\text{AID}}$ where, for certain $n_0$, an increase in $N_0$ actually causes $p_{\text{AID}}$ to increase rather than decrease (representative initial states $s''$ and $s'$, respectively, panels e-g).

Founder effects

Since the growing IPD is both non-linear and time-inhomogeneous, it resists most standard approaches to probabilistic analysis. Computing the statistics of the aforementioned $t \to \infty$ outcomes—i.e., the fixation probabilities—therefore involves using a high performance computing facility [12] to perform stochastic simulations.

Specifically, we employ a hybrid Gillespie-Itô approach (Appendix H) to approximate the following fixation probability:

$$p_{\text{AID}} = \lim_{t \to \infty} \Pr \{ n_t = \text{AllD-fp} \mid N_0, N_t > 0 \forall t' \leq t \},$$

where $n_t, N_t = \{ N_t^{\text{AID}}, N_t^{N_t} \}$. This is the limiting probability, after long times, that the system converges to the AllD fixed point, given specified initial conditions, and conditioned on populations that do not become extinct.

The results demonstrate several interesting features. For instance, choosing $n_0$ to be the unstable mixed-strategy fixed point, we see that $p_{\text{AID}}$ depends on both the rate of population growth and the rate of mutation (Fig. 2b). Similarly, fixing the rates of population growth and mutation reveals founder effects, where $p_{\text{AID}}$ depends on both initial demographic mix, $n_0$, and initial population size, $N_0$ (Figs. 2b-g).

The parameter space of mutation and growth rates can moreover be divided into two qualitative regions (panel a, red dashed line).

In the upper region, behaviour agrees with the expectation that faster growth rates reduce $O(1/\sqrt{N})$ intrinsic fluctuations more rapidly and therefore increase the likelihood of fixation within the same basin of attraction that the system started (panel a). Similarly, as the initial population size, $N_0$, increases, the system experiences fewer large fluctuations at early times and hence this also increases the likelihood of fixation within the starting basin of attraction (Figs. 2b-d).

In the lower region, however, behaviour is more complex. There are high likelihoods of an AllD outcome at both low ($b - d \leq 0.01$) and intermediate ($0.03 \leq b - d \leq 0.07$) rates of growth (panel a). This results in a non-monotonic dependence of $p_{\text{AID}}$ as a function of population growth. In other words, despite more rapidly reducing the fluctuations that are ostensibly required to cross the separatrix between mixed-strategy
and AllD basins, growth can actually increase the likelihood of fixing on the AllD fixed point. Similarly, rather than decreasing the likelihood of crossing the separatrix and fixing there, increasing \( N_0 \) can actually increase this likelihood for certain initial states, confounding expectations regarding the role of fluctuations at early times when populations remain small (Figs. 2-6).

That is, not only do the observed founder effects depend the rates of growth and mutation, but they can be in addition to the naive coupling between deterministic non-linearities and the reduction of fluctuations due to population growth.

Dynamical phases

To understand the non-trivial nature of Fig. 2, we repeatedly simulate the fixed population IPD (Appendix [H]), computing a so-called empirical distribution [13]. Specifically, we calculate the mean fraction of time spent in the AllD basin,

\[
F = \left\langle \frac{1}{T} \int_{0}^{T} 1_{n \in \text{AllD.b}} \, dt \right\rangle_N,
\]

where 1 is the indicator function, \( N \) denotes the size of the ensemble over which the average is taken, and the integration is understood in the Itô sense. The time \( T = 2 \times 10^4 \) is chosen to be larger than the maximum duration of our simulations of growing populations, which we stop when population sizes exceed \( 10^{10} \) (see Appendix [H]).

Our results (Fig. 5a) are consistent with a large deviation principle, such that \( p(F|N) \propto \exp \left(-N I_\mu(F_N)\right) \), where \( I_\mu \) is a convex rate function. Although determining the precise functional form of \( I_\mu \) is considered out-of-scope for this article, our data suggests that it has only three zeros, despite varying \( N \) over ten orders of magnitude:

\[
\arg \min I_\mu(F) \approx \begin{cases} 
0.5, & \forall \ N \leq 10^{2.7} \\
0, & 10^{2.7} < \ N < 1/\mu \\
1, & \forall \ N > 1/\mu
\end{cases}.
\]

In other words, depending on the population size, there are three statistically distinct types of demographic behaviour. As a population grows, and the characteristic scale of its intrinsic fluctuations decreases, it exhibits these dynamical phases sequentially, with transitions occurring abruptly at critical population sizes.

Stochastically-induced phase— Demographic trajectories at small population sizes are characterised by large intrinsic fluctuations and an intermediate value of \( F \) (Fig. 3b-c). Fluctuations are both correlated and state-dependent: features that are captured by the symmetric \( 3 \times 3 \) correlation matrix, \( B_{ij}^\prime \), that can be obtained by performing a Van Kampen system-size expansion [14] and projecting the results onto the unit simplex using Itô’s lemma (Appendix [I]). Fluctuations at the centre of the simplex are large and isotropic, whilst the components normal to the boundaries decrease rapidly as the edges and corners are approached. This gives rise to stochastically-induced effects [15-16], where fluctuation gradients bias stochastic trajectories, driving them towards the simplex edges and corners, on average (Fig. 3b-b and Appendices [C-F]). Despite such overall behaviour, \( B_{ij}^\prime \) is not symmetric under the interchange of \( n^{\text{AllC}}, n^{\text{AllD}} \) and \( n^{\text{TFT}} \), and stochastic trajectories retain characteristics encoded by the payoff matrix; including a bias for anti-clockwise motion, and a comparatively low likelihood of reaching the AllC corner (when compared to AllD and TFT corners). These behaviours are crucial to understanding the precise \( \mu \)-dependent mechanisms that underpin the value \( F \approx 0.5 \), which are discussed in the following section.

Asymmetric phase— As the population increases in size, the magnitude of fluctuations decrease, and the relative geometry of the underlying attractors becomes increasingly important. In particular, the system enters an asymmetric regime at populations above \( N \approx 10^{2.7} \), for which \( F \approx 1 \). Here, state-dependent fluctuations permit the system to cross from the mixed-strategy limit cycle to the AllD basin, but not from the AllD fixed point to the basin of the limit cycle (Fig. 3b). In other words, once the separatrix has been crossed, trajectories are extremely unlikely to come back.

Locked-in phase— Once \( N \) is sufficiently large, fluctuations are small and demographic trajectories are effectively locked into the mixed-strategy basin, implying \( F \approx 0 \) (Fig. 3f).

Role of mutation

The aforementioned dynamical phases couple to mutation, which alters both the character of the stochastically-induced phase and the duration of the asymmetric phase.

In the stochastically-induced phase, the standard deviation of fluctuations in the direction normal to the boundaries is \( O(\sqrt{\mu}) \) (Appendix [G]). Evolutionary trajectories therefore become increasingly confined to the boundaries as \( \mu \) decreases (Figs. 3b & c). This not only exacerbates stochastically-induced effects, but also increases the mean residence times associated with the corners. In particular, at comparatively high levels of mutation, residence times are less than \( T \), which results in a stochastic cycling between TFT and AllD corners (recall that there is a lower likelihood of finding the AllC corner) (Fig. 3b). Since these two corners have comparable mean rates of escape to the opposite basin of attraction (Appendix [I], \( F \) takes a value of approximately 0.5. By contrast, lower rates of mutation imply dwell times greater than \( T \) (Fig. 3). On average, therefore, the system will find either the AllD or TFT corner and then remain there. Here, the value of \( \approx 0.5 \) results from the position of the mixed strategy fixed point, and the correspondingly equal probability that trajectories are ex-
FIG. 3. **Three distinct $N$-dependent regimes characterise dynamical demographic phases.** When characterised by the fraction of time, $F$, that trajectories spend in the basin of the AllD fixed point, repeated simulations of the IPD at fixed population sizes demonstrate three distinct regimes of demographic behaviour (panel a). When population sizes are small, fluctuations are dominant, and the separatrix can be crossed in both directions (circles, panels b & c). Eigenvalues of the projected correlation matrix, $B^\dagger_{ij}$ indicate that the magnitude and bias of the fluctuations depend on demographic mix (orange crosses, panels b & c). Generally, this gives rise to fluctuation gradients that drive the system towards the simplex edges and corners. Although, due to the structure of the payoff matrix there is also an anti-clockwise bias and a lower probability of finding the AllC corner. At the simplex boundaries, normal fluctuations are proportional to $\sqrt{\mu}$ (panels b & c, magnified inset). Characteristic trajectories (beige) are therefore increasingly confined to the boundaries and corners as $\mu$ decreases. Since the residence times associated with the corners are $O(1/\mu)$, trajectories spend a disproportionate fraction of their time in the AllD and TFT corners (bars, panels b & c). Beyond this regime, behaviour becomes increasingly deterministic (green half-arrows, panels e-d). For intermediate population sizes, behaviour is asymmetric: demographic trajectories can cross the separatrix from the mixed-strategy limit cycle, but not from the AllD fixed point (panel d). The upper critical population size of this regime scales as $\sim 1/\mu$ (white, panel a). For large enough $N$, behaviour becomes increasingly deterministic, and trajectories remain locked-in to the mixed-strategy limit cycle (beige, panel e). At low $\mu$, the separatrix is typically crossed where it intersects either the TFT or AllD edges, which is where the magnitude of stochastic effects, $\|B^1\|_F$, are largest relative to deterministic flows, $|A_i|$, with $\|\cdot\|_F$ and $|\cdot|$ denoting Frobenius- and $\ell^2$-norms, respectively (panels c, d & f). Representative trajectories and population fractions (bars) only show a fraction of the total time simulated (see bar legend) in order to aid visualisation. Crossing statistics are taken from single simulations lasting $10^6$ s.

Despite the presence of these two $\mu$-dependent mechanisms in the fixed-$N$ ensemble, it should be stressed that, in a growing system, the duration of the stochastically-induced regime is sufficiently short that the latter behaviour is dominant for all but pathologically slow growth rates or very high $\mu$. As highlighted in later sections, this explains how the likelihood of being in either the AllD or mixed-strategy basins at the onset of the asymmetric regime depends on the population’s initial conditions.

For the asymmetric phase, mutation dictates the critical population size ($\sim 1/\mu$) at which the locked-in regime starts (Fig. 3b). The form of this dependence is subtle, and can be understood in terms of the stochastic ‘footprint’ of evolutionary trajectories in the locked-in regime—i.e., those that repeatedly (and stochastically) navigate the limit cycle (Fig. 3). In particular, due to the state-dependence of noise, the convex hull of this footprint is well-defined, reflecting the shape of the limit cycle at a different value of $\mu$ (Fig. 3a-d). The result is a family of locked-in trajectories, as depicted in panel e (beige).
of scaling relations, where the footprint of small populations with high levels of mutation (i.e., high noise, small limit cycle, Fig. 4c) is approximately equivalent to that of large populations with low levels of mutation (i.e., low noise, large limit cycle, Fig. 4d). The critical scaling that defines the onset of the regime occurs when the stochastic footprint fills the mixed-strategy basin (Figs. 4a & e) therefore facilitating the crossing of the separatrix. Here, $O(\sqrt{\mu/N})$ fluctuations must be equivalent to the $O(\mu)$ deterministic repulsion in the direction normal to the AllC-TFT edge (Appendix G), implying $N \sim 1/\mu$, which agrees with both ensemble statistics and convex hull analysis (cf. Figs. 3 & 4).

More generally, this behaviour can be recast as a size-dependent antagonistic relationship between mutation and intrinsic noise, where the first favours population heterogeneity and the second homogeneity. This is particularly important in growing populations, since the balance between the two factors changes over time.

### Equivalence classes

The asymmetric phase is especially important for determining fixation probabilities and hence founder effects. If the separatrix is crossed during this phase, then the system remains in the basin of the AllD fixed point until the onset of the locked-in phase, where it remains for all times.

A relevant quantity is therefore the conditional probability $P_{\text{AllD}}^{\text{asy}}(s) = \Pr\{n_{\text{lock}} \in \text{AllD-b} \mid n_{\text{sys}} = s\}$, where avoidance of extinction is now assumed implicitly. That is, the likelihood of being in the AllD basin at the onset of the locked-in phase, $t_{\text{lock}}$, given that the system was in a state $s$ at the onset of the asymmetric phase, $t_{\text{asy}}$. Computing this probability via stochastic simulation demonstrates the existence of three ‘equivalence classes’, $S_i$, such that $P_{\text{AllD}}^{\text{asy}}(s_i)$ is agnostic to the demographic mix $s_i \in S_i$, at the onset of the asymmetric phase (Fig. 5a and Fig. 10). These are: those states in the AllD basin, $S_1$; those along the AllD-TFT edge that stretch from the separatrix to the TFT corner, $S_2$; and; those along the TFT-AllC edge (excluding the TFT corner) on the mixed-strategy side of the separatrix, $S_3$. The remaining states of the mixed-strategy basin are labelled $S_4$. Whilst these do not form an equivalence class, we assume (and later show) that they only minimally contribute to overall fixation probabilities.

The existence of equivalence classes prompts the fol-
The dependence of fixation probabilities on rates of growth and mutation is reproduced by a decomposition based on dynamical phases and equivalence classes. Fixation probabilities can be decomposed in terms of conditional probabilities that are based on equivalence classes of the asymmetric phase (panel a). Within each class $S_i$, $i = 1, \ldots, 3$, the outcome of the asymmetric phase does not rely on the specific demographic mix at the onset of the phase. Moreover, the remaining states $S_4$ contribute very little to the overall fixation probabilities (panels e, i & m, and main text). This permits the approximation (4), graphically represented by panels b-n. The resulting reconstruction is in good agreement with the full stochastic simulations (panels n & o and Fig. 2a). The same decomposition is shown in Appendix L for additional initial conditions. The initial population size is $N_0 = 128$, whilst the times $t_{asy}$ and $t_{lock}$ are derived from the critical population sizes identified in Fig. 3.

Following approximation (Appendix I):

$$p_{\text{AID}} \approx \sum_{i=1}^{4} p_{\text{AID}}^{\text{sys}}(s_i) p_{S_i}^{\text{stoc}},$$

where $p_{S_i}^{\text{stoc}} = \Pr \{ n_{t_{\text{asy}}} \in S_i | N_0 \}$ (again, with avoidance of extinction assumed implicitly), and states $s_i$ can be chosen arbitrarily from $S_i$. This allows us to verify our heuristic understanding of how the rates of growth and mutation impact fixation probabilities by controlling the duration and stochastic character of dynamical phases and therefore the likelihood of (stochastic) behaviours, such as crossing the separatrix, that are crucial in dictating long term outcomes and hence founder effects.

In the stochastically-induced regime, fluctuation gradients ‘drive’ trajectories from the centre of the mixed-strategy basin towards the simplex edges and then the corners. As a result, the growth-dependent (average) duration of the regime, $\tau^{\text{stoc}} = -\ln (10^{2.7}/N_0)/(b-d)$, dictates in which class the trajectories are likely to start the asymmetric phase: rapid growth rates are required to confine trajectories that end in the mixed-strategy basin to the $S_4$ region, whilst intermediate and slow growth rates suffice for the $S_3$ and $S_2$ regions, respectively (Figs. 5b-d). Although the average time spent in the two basins by such trajectories during the stochastically-induced regime is independent of $\mu$ (Fig. 3a), the likelihood that the system is in $S_1$ at time $t_{\text{asy}}$ actually increases with decreasing $\mu$ (Fig. 5a). The reason is that decreasing $\mu$ changes the shape of the separatrix, therefore reducing the size of the $S_3$ region (Appendix K).

The asymmetric regime, by contrast, has a duration that is both growth- and $\mu$-dependent: $\tau^{\text{asy}} = -\ln (\mu 10^{2.7})/(b-d)$. Here, the probability of crossing the separatrix hinges, principally, on the likelihood of avoiding the TFT corner and its associated large confinement times (Figs. 5g-i). For example, $\tau^{\text{asy}}$ must be
FIG. 6. Founder effects can be reproduced by a decomposition based on dynamical phases and equivalence classes. For certain states—i.e., those below the red line, Fig. 2a—fixation probabilities can rely critically on initial conditions, inviting comparisons with the founder effect. That is, initial demographic mixes that are ‘close’ (and not necessarily near the separatrix) can give rise to dramatically different long term fixation probabilities. This is particularly striking at low levels of mutation ($\log \mu = -6.5$) and comparatively high initial population size ($N_0 = 250$) (cf. Fig. 2f). The reason for this is twofold. First, larger initial populations reduce the effect of the stochastically-induced phase, which is to expel populations to the simplex boundaries and the AllD and TFT corners (panels c & d). Second, the duration of the asymmetric phase is significant and cannot be neglected (middle row). This implies that trajectories finishing the stochastically-induced phase in classes $S_3$ or $S_4$ can still cross the separatrix and fixate on the AllD fixed point (panels g & h). Whilst this provides a heuristic understanding of the founder-like effect, the contribution to $p_{\text{AllD}}$ from the $S_4$ region violates one of the assumptions of (4), and therefore the error increases with $N_0$ for the stated values of $\mu$ and $b - d$ (panel j and Appendix M, Figs. 18 & 19).

extremely long in order to permit crossings from the $S_2$ region, since populations starting the asymmetric regime from this region encounter the TFT corner with almost certainty. Crossings from the $S_3$ region, however, occur at more modest $\tau^\text{asy}$—achieved by either low growth and high $\mu$, or modest growth and low $\mu$—reflecting the possibility that trajectories might avoid the TFT corner. Those from the $S_3$ region can happen at the smallest $\tau^\text{asy}$, since there is a high likelihood that trajectories will avoid the TFT corner and the ratio of the magnitude of stochastic effects to the magnitude of the deterministic flow is large in $S_3$ (see Fig. 3).

Combining these conditional probabilities using Eq. (4) satisfactorily reproduces the overall statistics of demographic outcomes (Fig. 3b): the difference between our approximation and the full simulations have a mean value of 0.027, when averaged over growth and mutation rates, and a maximum value of 0.099 (Fig. 3d). This also confirms our assertion that the trajectories that start the asymmetric phase from the $S_4$ region do not impact long term outcomes. The reason is that this only happens with significant likelihood when growth rates are high, which simultaneously ensures that such trajectories never cross the separatrix (Figs. 3i, i & m). The same decomposition is shown for initial demographic mixes other than the mixed-strategy fixed point in Appendix L.

Moreover, the approximation (4) further provides insight into the founder-like dependence of fixation probabilities on initial conditions (Fig. 6 and Appendix M, Figs. 15 & 17). At high values of $\mu$, the asymmetric phase has negligible duration and behaviour is trivial (Figs. 2d & c, Appendix M, Figs. 15 & 19). At low values of $\mu$, however, the asymmetric regime cannot be ignored and has a significant bearing on fixation probabilities.

In this case, if the initial population size is small (Fig. 2e, Appendix M, Fig. 17), then the stochastically-induced phase is sufficiently long as to expel trajectories to the AllD or TFT corners from the AllD or mixed-strategy basins, respectively. Only the former trajectories impact fixation probabilities, however, since for all but the slowest growth rates, trajectories stuck in the TFT corner have a residence time longer than the duration of the asymmetric regime.

By contrast, if the initial population size is large
(Fig. 2 and Fig. 6), then the duration of the stochastically-induced phase is not long enough to expel trajectories to the boundaries and/or corners, resulting in a non-zero probability of starting the asymmetric regime from either the $S_1$ or $S_3$ regions. For the former, there is a high probability of crossing to the AlID basin during the asymmetric regime, since the deterministic flows direct demographic trajectories towards the separatrix. For the latter, this probability is much lower, since trajectories are more likely to be entrained to the limit cycle. Nevertheless, the small contribution that results from the $S_3$ region breaks one of the assumptions on which (4) is based, which also explains why the error in our decomposition increases with $N_0$ for certain values of $\mu$ and $b - d$ (the mean value of the error in Fig. 6 is 0.079, while the maximum value is 0.285). Of note, the values shown in Fig. 6 are the worst case of those we have simulated (Appendix M, Figs. 18 & 19).

DISCUSSION

Using a growing variant of an iconic model of evolutionary game theory, we have demonstrated the existence of non-trivial growth- and mutation-dependent founder effects in growing populations. Our model is well-mixed, and as such pertains to systems that combine short-range interactions with a mechanism for mixing that is fast on the timescales of the population dynamics, or those that otherwise have effectively long-range interactions, either explicitly or via the mutual interaction with public resource (although some of these assumptions have recently been brought into question [17]).

In this context, we follow several pioneering works that have characterised various effects of growth in well-mixed populations, including: driving novel demographic transients [18]; facilitating cooperation [19, 20], even when deterministic selection favours defection [21], and; altering the success of invading variants [22]. Although they do not explicitly examine the role of initial conditions, the behaviours appearing in two of these works [21, 22] are related to the those reported here. The same is true of several studies concerning growing systems of binary 'spins' [11, 23, 24].

The shared mechanism in all these cases is the presence of $O(1/\sqrt{N})$ intrinsic fluctuations that decrease as the system grows. In growing populations, this has two main ramifications: first, ergodicity is broken, implying that a population fixes even in the presence of non-zero mutation, and; second, there is a decreasing scale by which fluctuations can couple dynamically to the geometry of underlying deterministic attractors of the system. It is the combination of these effects that gives rise to our headline behaviour: fixation probabilities that depend non-trivially on initial conditions (or otherwise stochastic events at early times), even as $t \to \infty$.

Such fluctuation-mediated effects are further characterised by a dependence on the rates of growth and mutation, which we explain in the context of dynamical demographic phases (a direct corollary of the second point, above). The rate of growth controls the rate at which intrinsic fluctuations decrease, and therefore also the relative duration of dynamical phases. Mutation, by contrast, has two related effects. First, it changes the nature of the underlying deterministic attractors (the size of the mixed-strategy limit cycle, in our case) and hence the structure to which decreasing fluctuations couple. Second, it also changes the state-dependence of fluctuations and hence stochastically-induced effects. Whilst the former impacts the character of the initial stochastically-induced phase, it is the combination of both of these effects that sets the $\sim 1/\mu$ dependence of the critical transition between asymmetric and locked-in phases.

The latter behaviour is an example of a potentially interesting and unexpected antagonistic relationship between an exogenous source of stochasticity, the mutation, and the intrinsic effects of finite sized populations. In our model and elsewhere, mutation promotes heterogeneity, whilst intrinsic fluctuations typically drive the system towards homogeneity [23]. This is particularly important in the context of population growth, since the effects of mutation do not depend on the population size, whereas the stochastically-induced forcing due to intrinsic noise decreases with increasing population size. As a result, we speculate the other growing systems may also exhibit mutation-dependent critical transitions, where the effects of intrinsic fluctuations are balanced by those of mutation.

More generally, for growing well-mixed populations with all but the simplest of deterministic attractors—i.e., fixed points, limit cycles and stable manifolds etc.—the implication is therefore that an interplay between growth and mutation, via dynamical phases, is an important consideration for evolutionary outcomes, and population dynamics more generally, even in the absence of a direct coupling.

Against this backdrop, we put forward that such ideas may be examined within the context of directed evolution [25, 26]. Here, another, albeit direct, interplay between growth and mutation has already been demonstrated: mutations occurring at the genetic loci associated with growth-control promote so-called genetic instabilities [27]. Indeed, our results also appear relevant to state-of-the-art in silico representations of directed evolution, where the role of intrinsic fluctuations during growth has so far been overlooked [28]. Other settings might also be considered, beyond the notion of point mutations, including post-translational effects or other sources of exogenous stochasticity that couple to replication. We therefore speculate that there may be a wider role for the behaviours outlined here, with embryogenesis [29] and cancerous tumours [30, 31] two such examples.

Nevertheless, bridging the gap between the abstract setting of the present work and the aforementioned applications will undoubtedly involve significant work. The
extent to which this will be possible remains an open question, and may hinge on features that are not included in our model, such as spatial structure and/or other physical constraints [32–34]. Exploring how the ideas set out here translate across a wider class of systems is therefore an important avenue of future research, and we welcome further work in the area.

AUTHOR CONTRIBUTIONS

EC and JP performed simulations and analysis, under the guidance of ST and RM. EC and RM wrote the manuscript, with help from JP and ST. The project initially arose from discussions between JP, ST and RM.

ACKNOWLEDGMENTS

The authors would like to thank V. Guttal for helpful discussions at the project’s outset. RGM and EC acknowledge EMBL Australia for funding. We further acknowledge support from the Simons Foundation (Grant No. 287975 to ST), the Max Planck Society through a Max-Planck-Partner-Group at NCBS-TIFR (ST) and the Department of Atomic Energy, Government of India, under projects RTI4001 and RTI4006. This research includes computations using the computational cluster Katana supported by Research Technology Services at UNSW Sydney.

Appendix A: Growing Frequency-Dependent IPD

The model we use in the main text is a straightforward extension of a Moran process (see Fig. 7) where birth and death are uncoupled. We use the canonical m-round IPD payoff matrix [9]:

\[
\begin{pmatrix}
\text{AllC} & \text{AllD} & \text{TFT} \\
\text{AllC} & Rm & Sm & Rm \\
\text{AllD} & Tm & Pm & T + P(m-1) \\
\text{TFT} & Rm - c & S + P(m-1) - c & Rm - c
\end{pmatrix}
\]

such that \( T > R > P > S \) and \( R > (T + S)/2 \). The specific values used were \( T = 5, R = 3, P = 1, S = 0.1 \) and \( c = 0.2 \). The average fitness of a given strategy, when played against the whole population (including individuals of the same strategy), is just

\[
f_i = \frac{\sum_{j=1}^{3} a_{ij} N^j - a_{ii}}{N - 1}, \tag{A1}
\]

FIG. 7. Decoupling birth and death. Individuals are born into a strategy \( i \) with a specific (per individual) birth rate \( b_i \), which is weighted by that strategy’s relative fitness, \( f_i/\varphi \). Individuals die with a specific rate \( d \), irrespective of their strategy. Mutation at birth is captured statistically; a fraction \( \mu \) of individuals have their strategy switched at random, with a rate proportional to their current strategy’s frequency-dependent birth rate.
where latin indices denote the three different strategies, such that AllC → 1, AllD → 2 and TFT → 3. The $a_{ij}$ are then the components of the above payoff matrix, and the $N^i$ the number of individuals playing each strategy. In the following, but not in simulations, we use the simplification that

$$f_i = \frac{\sum_{j=1}^{3} a_{ij} N^j}{N},$$

which, although including self-interactions, still retains all the relevant features associated with the IPD [10]. In either case, the mean fitness is just

$$\varphi = \frac{\sum_{i=1}^{3} f_i N^i}{N}.$$  

(A3)

**Appendix B: The deterministic limit**

In the $N \to \infty$ limit, the behaviour of the Growing Iterated Prisoner Dilemma described in the main text is equivalent to that of the following continuous, deterministic system:

$$\frac{dn^i}{dt} = \frac{b\mu}{\varphi} \sum_{j \neq i} (f_j n^j - f_i n^i) + \left(\frac{bf_i}{\varphi} - d\right) n^i,$$

(B1)

where $n^i = N^i/N$. These equations can be easily derived from Fig. 7.

As shown in Fig. 8 (see also [9]), the mutation rate has a clear effect on the dynamics of the continuous system. For $10^{-7.5} \leq \mu \leq 10^{-2.5}$, the system exhibits two stable attractors: a fixed point characterised by a population of almost entirely AllD players and a stable mixed-strategy limit cycle. The most notable effect of varying $\mu$ within this range is the resizing of the limit cycle: the smaller the mutation rate, the larger the limit cycle. For large values of $\mu$ (approximately $10^{-3.1} \leq \mu \leq 10^{-2.5}$), the limit cycle collapses into a point, although the separatrix still exists (see panel d and h). Varying $\mu$ also slightly changes the shape of the separatrix, especially in the proximity of the AllC-TFT edge of the simplex.

![figure 8](image)

**FIG. 8.** Population dynamics in the deterministic limit of the GIPD at different values of $\mu$. The top row (panels a-d) show the direction of the determinist flow, while the bottom row (panels e-h) show the magnitude (speed). For all panels, the growth rate is $b - d = 0.05$. 

$log_{10}\mu = -6.9$

$log_{10}\mu = -5.1$

$log_{10}\mu = -3.9$

$log_{10}\mu = -2.7$

$0$

$0.1$

$0.2$

$0$
Appendix C: System-Size Expansion and FPE

Following [23], we derive a set of coupled SDEs that approximate the dynamics of the underlying protocol when the system size is greater than some number, \( N_0 \). To do this, we trivially introduce the variables \( x_i = N_i/N_0 \) and \( s = N/N_0 \), such that \( \sum_i x_i = s \) and \( n_i = N_i/N = x_i/s \). In terms of step-operators, the resulting Master-equation has the form:

\[
\frac{dP(\vec{x};t)}{dt} = N_0 \sum_i \left( \frac{E_{x_i}}{N_0} - 1 \right) \frac{bf_i x_i}{\varphi} P(\vec{x};t)
+ N_0 \sum_i \left( \frac{E_{x_i}^{-1}}{N_0} - 1 \right) dx_i P(\vec{x};t)
+ N_0 \sum_i \sum_{j \neq i} \left( \frac{E_{x_i}^{-1} E_{x_j}^{-1}}{N_0} - 1 \right) \frac{b\mu f_i x_i}{\varphi} P(\vec{x};t).
\]  

(C1)

Expanding the step-operators in the usual fashion, and retaining only the leading and next-to-leading order terms, gives

\[
E_{x_i} = 1 \pm \frac{1}{N_0} \frac{\partial}{\partial x_i} + \frac{1}{2 N_0^2} \frac{\partial^2}{\partial x_i^2} + O \left( \frac{1}{N_0^3} \right).
\]  

(C2)

Substituting into (C1) results in an equation of the Fokker-Planck type [35]:

\[
\frac{\partial P(\vec{x};t)}{\partial t} = -\sum_i \frac{\partial}{\partial x_i} \left[ \alpha_i(\vec{x}) P(\vec{x};t) \right]
+ \frac{1}{2} \sum_i \sum_j \frac{\partial^2}{\partial x_i \partial x_j} \left[ \beta_{ij}(\vec{x}) P(\vec{x};t) \right].
\]  

(C3)

After some manipulation, it can be shown that

\[
\alpha_i = \frac{b(1 - 2\mu)}{\varphi} f_i x_i + \sum_{j \neq i} \frac{b\mu f_j x_j}{\varphi} - d x_i,
\]  

(C4)

and

\[
\beta_{ij} = \frac{1}{N_0} \times \left\{ \frac{b(1+2\mu)}{\varphi} f_i x_i + \sum_{j \neq i} \frac{b\mu f_j x_j}{\varphi} + d x_i \right. \text{ if } i = j,
\]  

\[
\left. -\frac{b\mu}{\varphi} (f_i x_i + f_j x_j) \right) \text{ if } i \neq j.
\]  

(C5)

Appendix D: SDE: additive correlated noise

Equation (C3) implies an SDE for the variables \( x_i \). Rescaling time by a constant factor of \( N_0 \) transforms back to the variables \( N^i \), and we write the result using traditional (compact) conventions [10]:

\[
\frac{dN^i}{dt} = \sum_j f_j N^j q_{ji} - d N^i + \frac{\zeta_i}{\sqrt{N_0}},
\]  

(D1)

where

\[
q_{ij} = \begin{cases} \frac{b(1-2\mu)}{\varphi} & \text{if } i = j, \\ \frac{b\mu}{\varphi} & \text{if } i \neq j. \end{cases}
\]  

(D2)

The deterministic part of (D1) is equivalent to (B1). The noise sources that appear in (D1) again have zero mean (i.e., \( \langle \zeta_i \rangle = 0 \)) and are correlated according to:

\[
\langle \zeta_i(t)\zeta_j(t') \rangle = B_{ij} \delta(t - t'),
\]  

(D3)
with

\[ B_{ij} = \begin{cases} \sum_k f_k N^k q^*_i \frac{d N^i}{d t} & \text{if } i = j \\ -\frac{b_i}{\varphi} \left(f_i N^i + f_j N^j\right) & \text{if } i \neq j \end{cases} \]

and

\[ q^*_ij = \begin{cases} b(1+2\mu) & \text{if } i = j \\ b\mu/\varphi & \text{if } i \neq j \end{cases} \]

which is distinct from (D2). Summing over index \( i \) in (D1), we have

\[ \frac{dN}{dt} = (b - d) N + \sqrt{\frac{(b + d) N}{N_0}} \xi, \]

where \( \xi \) is a single source of zero-mean Gaussian white noise.

**Appendix E: SDE: multiplicative delta-correlated noise**

Eqn. (D1) can be re-written in terms of delta-correlated noise and multiplicative pre-factors. To do this we must choose a matrix \( b \) that satisfies \( B = b^T \cdot b \). For \( b \) to be square, this requires a Cholesky decompostion, and ensures no more independent noise sources than there are variables in the system. However, a more natural approach is to decompose \( B \) according to the rules set out in Gillespie’s CLE approach [36]. This results is SDEs of the following form:

\[ \frac{dN^i}{dt} = \sum_j f_j N^j q_{ji} - d N^i + \frac{1}{\sqrt{N_0}} \sum_\alpha b_{i\alpha} \eta_\alpha, \]

where \( q_{ij} \) is given by (D2). The index \( \alpha \) (and all other Greek indices, unless otherwise specified) run from 1...12. The twelve independent noise sources each have mean zero, and are delta-correlated—i.e., \( \langle \eta_\alpha \rangle = 0 \) and \( \langle \eta_\alpha(t) \eta_\beta(t') \rangle = \delta_{\alpha\beta} \delta(t - t') \). The matrix \( b^T \) is given by

\[
\begin{pmatrix}
\sqrt{b f_1 N^1} / \varphi & 0 & 0 \\
0 & \sqrt{b f_2 N^2} / \varphi & 0 \\
-\sqrt{d N^1} & 0 & \sqrt{b f_3 N^3} / \varphi \\
0 & -\sqrt{d N^2} & 0 \\
0 & 0 & -\sqrt{d N^3} \\
-\sqrt{b f_1 N^1} / \varphi & -\sqrt{b f_2 N^2} / \varphi & 0 \\
\sqrt{b f_3 N^3} / \varphi & 0 & -\sqrt{b f_2 N^2} / \varphi \\
0 & \sqrt{b f_3 N^3} / \varphi & 0 \\
-\sqrt{b f_1 N^1} / \varphi & 0 & \sqrt{b f_3 N^3} / \varphi \\
\sqrt{b f_3 N^3} / \varphi & 0 & -\sqrt{b f_3 N^3} / \varphi \\
\end{pmatrix}
\]
Appendix F: Projected dynamics: Itô’s lemma

We wish to project the dynamics onto the unit simplex—i.e., in terms of variables \( n^i = N^i / \sum_i N^i \). For this, we require the multivariate form of Itô’s lemma \[35\]. For finite \( N \), we have

\[
\frac{dn^i}{dt} = \left( \sum_{k,j} f_j N^i q_{jk} - d N^k \right) \frac{\partial n^i}{\partial N^j} + \frac{1}{2N_0} \sum_{k,j} B_{kj} \frac{\partial n^i}{\partial N^k} + \frac{1}{\sqrt{N_0}} \sum_j \sum_{\alpha} b^i_{\alpha j} \left( \frac{\partial n^i}{\partial N^j} \right) \eta_{\alpha},
\]

where the shorthand \( \partial_i = \partial / \partial N^i \) has been used. First, we deal with the deterministic parts. Using

\[
\frac{\partial n^i}{\partial N^j} = \begin{cases} \frac{1-n^i}{N} & \text{if } i = j \\ -\frac{n^i}{N} & \text{if } i \neq j \end{cases},
\]

it can be shown that

\[
\sum_{k,j} f_j N^i q_{jk} \frac{\partial n^i}{\partial N^j} = \frac{b_i}{\varphi} \left( \sum_{j \neq i} f_j n^j - 2f_i n^i \right) + b \left( \frac{f_i}{\varphi} - 1 \right) n^i.
\]

Similarly, using

\[
\frac{\partial^2 n^i}{\partial N^j \partial N^k} = \begin{cases} \frac{2(n^i-1)}{N^2} & \text{if } i = j = k \\ \frac{2n^j}{N^2} & \text{if } j = i \neq k \text{ or } k = i \neq j \\ -\frac{2n^j}{N^2} & \text{if } j \neq i \text{ and } k \neq i \end{cases},
\]

gives

\[
\frac{1}{2N_0} \sum_{k,j} B_{kj} \frac{\partial n^i}{\partial N^k} = -b \left( \frac{f_i}{\varphi} - 1 \right) \frac{n^i}{N_0 N}.
\]

Notice that, when summed over \( i \), both \[[F3]\] and \[[F5]\] are zero by virtue of the fact that \( \sum_i f_i n^i / \varphi = 1 \). Also, trivially, \[[F5]\] goes to zero in the deterministic \( N \to \infty \) limit. For the stochastic part of \[[F1]\], define a new matrix \( b^i_{\alpha j} = \sum_j b_{\alpha j} \left( \partial_j n^i \right) \) such that \[[F1]\] can be re-cast in terms of correlated noise sources—i.e.,

\[
\frac{dn^i}{dt} = \frac{b_i}{\varphi} \left( \sum_{j \neq i} f_j n^j - 2f_i n^i \right) + b \left( \frac{f_i}{\varphi} - 1 \right) n^i \left( 1 - \frac{1}{N_0 N} \right) + \frac{\zeta_i}{\sqrt{N_0}},
\]

with \( \langle \zeta_i \rangle = 0 \) and \( \langle \zeta_i(t) \zeta_j(t') \rangle = B_{ij}^L \delta (t - t') \), where

\[
B_{ij}^L = \sum_\alpha b^i_{\alpha i} b^j_{\alpha j} \begin{cases} \frac{b_i n^i}{\varphi} \left( n^i + \frac{f_i}{\varphi} (1 - 2n^i) \right) + \frac{b_j n^j}{\varphi} \left( 1 + \frac{f_i n^i}{\varphi} \right) + \frac{d n^i}{\varphi} \left( 1 - n^i \right) & \text{if } i = j \\ -\frac{b_i n^i}{\varphi} \left( \frac{f_i}{\varphi} - 1 \right) - \frac{b_j n^j}{\varphi} (f_i n^i + f_j n^j) - \frac{d n^i n^j}{\varphi} & \text{if } i \neq j \end{cases}.
\]

To understand the impact of these correlations, \[[F7]\] can be computationally decomposed into an eigenbasis for different values of \( n^i \). This always results in one zero-eigenvalue eigenvector pointing perpendicular to the simplex. The remaining in-simplex eigenvalues reveal that populations towards the centre of the simplex experience large uncorrelated fluctuations whilst, closer to the simplex boundary, correlations suppress fluctuations in the direction normal to the boundary (Fig. 3b-c in the main text, orange crosses). Moreover, the magnitude of along-boundary fluctuations decrease as a corner is approached.
Appendix G: Boundary effects

Equation (G6) can be evaluated at the simplex edges, where we are particularly interested in the both deterministic drift and fluctuations in the direction of the bulk, which is captured by the dynamics of the strategy who’s concentration is zero along a given edge.

- AllD-AllC edge: setting $n^3 = 0$ and $n^2 = 1 - n^1$, gives $dn^3/dt = b\mu + \zeta_1/\sqrt{N_0}$, where

$$\langle \zeta_1 \zeta_j \rangle = \left( \begin{array}{c} \frac{b\mu n^1[R-S]}{N \{P(n^1)^2+n^1[R-S-T]+S+T]\}} \\ -\frac{b\mu n^1[R-S-T]}{N \{P(n^1)^2+n^1[R-S-T]+S+T]\}} \end{array} \right). \quad \text{(G1)}$$

- AllD-TFT edge: setting $n^2 = 0$ and $n^1 = 1 - n^3$, gives $dn^2/dt = b\mu + \zeta_2/\sqrt{N_0}$, where

$$\langle \zeta_2 \zeta_j \rangle = \left( \begin{array}{c} \frac{b\mu m(n^3-1)R}{N (mR-c+n^3)} \\ \frac{b\mu n^3 (mR-c)}{N (c-n^3-mR)} \end{array} \right). \quad \text{(G2)}$$

- TFT-AllD edge: setting $n^1 = 0$ and $n^3 = 1 - n^2$, gives $dn^3/dt = b\mu + \zeta_3/\sqrt{N_0}$, where

$$\langle \zeta_3 \zeta_j \rangle = \left( \begin{array}{c} \frac{b\mu n^2[P(m+n^2-1)+T(n^2-1)]}{N \{c(n^2-1)-(n^2)^2\} [P(m+n^2-1)+T(n^2-1)]} \\ -\frac{b\mu n^2[P(m+n^2-1)+T(n^2-1)]}{N \{c(n^2-1)-(n^2)^2\} [P(m+n^2-1)+T(n^2-1)]} \end{array} \right). \quad \text{(G3)}$$

In all three of the above cases, the deterministic repulsion from the edge is $O(\mu)$. The fluctuations (positive and negative) in the same direction are $O(1/\sqrt{N})$. The implication, in the context of the convex hull analysis of the main manuscript, is that, very close to the AllD-TFT edge, fluctuations can overcome the deterministic forces only if $N \sim 1/\mu$.

Appendix H: Gillespie-Itô simulations

The Gillespie algorithm \[36, 37\] allows the exact simulation of the stochastic dynamics of the growing IPD when $N$ is small, i.e., when such dynamics cannot be approximated by simply integrating (E1). The downside of this algorithm is that its computational time scales linearly with $N$, quickly becoming impractical as the population grows. We therefore adopted a hybrid approach: when $N$ is smaller than a chosen threshold the system’s dynamics are simulated using the Gillespie algorithm, and when the threshold is exceeded they are simulated by numerically integrating the SDEs (Euler-Maruyama). We remark that, since the latter method results in values of $N^1$ along the real line, rounding is required if stochastic fluctuations trigger a switch back to the Gillespie algorithm from numerical integration.

Our choice for the algorithm switch threshold is $N = 10^{10}$. This is motivated by the existence of the locked-in phase, which begins when $N \approx 10^{10}$ (see Fig. 3 of the main text). During this phase, the risk that the SDE approximation would lead to an ‘accidental’ crossing of the separatrix is extremely small. A minimum threshold of $N = 10,000$ is used for simulations with large $\mu$.

All results were obtained using the High Performance Computing Facility Katana \[12\]. The results in Fig. 2a in the main text (and Figs. 2i to 14, panel p), were obtained with 10,000 repetitions of the hybrid Gillespie-Itô simulations for each combination of $b$ and $\mu$ (fixing $d = 1$). Each repetition was carried out until $N > 10^{10}$. The results in Fig. 2b-g in the main text were similarly obtained with 1000 repetitions for each initial condition. The results in Fig. 5b-i in the main text (and Figs. 12 to 14, panels b-i) were obtained with 1000 repetitions for each combination of $\mu$ and $b$. 1000 repetitions were also used to obtain the results in Figs. 6a-d in the main text and Figs. 15 to 19 panels a-d.

The fixed size dynamics were also simulated using the hybrid Gillespie-Itô approach. The Gillespie algorithm was modified by setting $d = 0$ and ‘killing’ a randomly selected player at every birth. The Itô part of the algorithm...
involved replacing the matrix $b^T$ in Eq. [E2] with

\[
\begin{pmatrix}
(1 - n^1)\sqrt{\frac{bf_1 N^1}{\phi}} & -n^1 \sqrt{\frac{bf_2 N^2}{\phi}} & -n^1 \sqrt{\frac{bf_3 N^3}{\phi}} \\
-n^2 \sqrt{\frac{bf_1 N^1}{\phi}} & (1 - n^2)\sqrt{\frac{bf_2 N^2}{\phi}} & -n^2 \sqrt{\frac{bf_3 N^3}{\phi}} \\
-n^3 \sqrt{\frac{bf_1 N^1}{\phi}} & -n^3 \sqrt{\frac{bf_2 N^2}{\phi}} & (1 - n^3)\sqrt{\frac{bf_3 N^3}{\phi}} \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & -\sqrt{\frac{bf_2 N^2}{\phi}} & \sqrt{\frac{bf_2 N^2}{\phi}} \\
-\sqrt{\frac{bf_3 N^3}{\phi}} & 0 & \sqrt{\frac{bf_3 N^3}{\phi}} \\
\sqrt{\frac{bf_1 N^1}{\phi}} & 0 & -\sqrt{\frac{bf_1 N^1}{\phi}} \\
\end{pmatrix}.
\] (H1)

The results in Fig. 3 in the main text were obtained by running the fixed size Gillespie-Itô algorithm 1000 times for 20,000 seconds for each combination of $\mu$ and $N$. 
Appendix I: Escape rate from corners in the stochastically-induced regime

Here we show the statistics of the time $t^\dagger$ that is needed by the fixed-size system to cross the separatrix for the first time (Fig. 9). We compare the cases of the systems starting from the AllD corner (panels a and c) and the TFT corner (panels b and d). In both cases, $t^\dagger$ is proportional to $\mu$: in the order of thousands of seconds when $\log_{10} \mu = -3.5$ (panels a-b) and in the order of millions of seconds when $\log_{10} \mu = -6.5$ (panels c-d). However, the influence of the population size $N_0$ on $t^\dagger$ is different in the two cases. On the one hand, the crossing time starting from the AllD corner increases very rapidly when $N_0$ approaches the onset of the asymmetric regime at around $N_0 = 10^{2.7}$ (panels a and c), indeed showing that crossing the separatrix from the AllD basin to the mixed-strategy basin becomes extremely unlikely in the asymmetric regime. On the other hand, the crossing time starting from the TFT corner decreases with $N_0$.

**FIG. 9.** Statistics of separatrix crossing time in the stochastically induced regime. The figure shows the statistics of the time taken by the system to cross the separatrix for the first time using the fixed-size model with $N_0 = s_1$, i.e. the AllD corner (panels a for $\log_{10} \mu = -3.5$ and c for $\log_{10} \mu = -6.5$) and with $N_0 = s_2$, i.e. the TFT corner (panels b for $\log_{10} \mu = -3.5$ and d for $\log_{10} \mu = -6.5$). The statistics, for each of the four cases, are obtained from 10,000 repetitions. On each box, the central red mark indicates the median, and the bottom and top edges indicate the 25th and 75th percentiles, respectively; the whiskers extend to the most extreme data points not considered outliers, and the red dots are the outliers (a data value is considered an outlier if it is greater than $Q_3 + 1.5(Q_3 - Q_1)$ or less than $Q_1 - 1.5(Q_3 - Q_1)$, where $Q_1$ and $Q_3$ are the 25th and 75th percentiles, respectively). For an easier visualisation of the statistics, all data values exceeding a maximum threshold are collapsed into the horizontal dashed lines on the top part of each plot. The blue lines represent the average crossing time over all repetitions.
Appendix J: Equivalence classes decomposition

Consider a trajectory of the growing IPD that starts from an arbitrary state in the stochastically-induced phase (i.e., \( N_0 < 10^{27} \)) at time 0, enters the asymmetric phase at time \( t_{\text{asy}} \) and the locked-in phase at time \( t_{\text{lock}} \). Due to the nature of the locked-in phase, the evolutionary outcome at \( t \to \infty \) is already known at time \( t_{\text{lock}} \), and thus Eq. (1) in the main text can be simplified:

\[
    p_{\text{AllD}} = \Pr\{n_{\text{lock}} \in \text{AllD-b} \mid N_0\}, \tag{J1}
\]

where avoidance of extinction is implicitly assumed and \( \text{AllD-b} \) indicates the AllD basin of the simplex.

Moreover, since the GIPD is a Markovian process, \( p_{\text{AllD}} \) can be decomposed as follows:

\[
    p_{\text{AllD}} = \sum_{s \in U} \Pr\{n_{\text{lock}} \in \text{AllD-b} \mid n_{\text{asy}} = s\} \Pr\{n_{\text{asy}} = s \mid N_0\}, \tag{J2}
\]

where \( U \) is the set of all possible states of the system at time \( t_{\text{asy}} \).

Eq. (J2) is impractical, since it requires the numerical estimation of the probabilities \( \Pr\{n_{\text{lock}} \in \text{AllD-b} \mid n_{\text{asy}} = s\} \) and \( \Pr\{n_{\text{asy}} = s \mid N_0\} \) for all \( s \) in the very large set, \( U \). However, the approximation given in Eq. (4) in the main text can be made by inspecting \( \Pr\{n_{\text{lock}} \in \text{AllD-b} \mid n_{\text{asy}} = s\} \) in Fig. 10a, which was numerically estimated for a subset of all points \( s \) (for illustration purposes we only show the case of \( b \# d = 0.05 \) and \( \log_{10} \mu = -6.5 \)).

A first area of the simplex, \( S_1 \) (outlined in red), corresponding to the AllD basin, can be immediately identified as it is characterised by a homogeneous probability of AllD outcomes of approximately 1. If the system is anywhere within \( S_1 \) at time \( t_{\text{asy}} \), then it is expected never to move to the mixed-strategy basin, since the crossing of the separatrix in this direction is extremely unlikely. Any point \( s_1 \) would be good candidate for representing the entire area \( S_1 \), however, we chose the AllD corner (red dot in Fig. 10b) as the system is in the proximity of such point for the vast majority of the time spent in \( S_1 \).

A second area can similarly be identified: \( S_2 \) (orange), corresponding to the mixed-strategy part of the AllD-TFT edge (including the TFT corner) and characterised by a homogeneous probability of AllD outcomes of approximately 0. If the system is in this area at time \( t_{\text{asy}} \), then it is very unlikely to cross the separatrix because of the long time (inversely proportional to \( \mu \)) spent in the TFT corner while growing (i.e., as the fluctuations become smaller and smaller). Again, any point \( s_2 \) can represent the area \( S_2 \), but we chose the TFT corner (orange dot in Fig. 10b).

A third area, \( S_3 \) (yellow), corresponding to the mixed-strategy part of the AllC-TFT edge with exclusion of the TFT corner, also stands out: here we see the probability of AllD outcomes quickly increases with the fraction of AllC players. Due to the counter-clockwise dynamics, in \( S_3 \) the system directed towards the point where the separatrix meets the AllC-TFT edge, which is where a move from the mixed-strategy basin to the AllD basin is most likely (see Fig. 3f in the main text). Since the probability of AllD outcomes is less homogeneous, an approximation must be made for \( S_3 \). The dynamics along the AllC-TFT edges are slow, and they get even slower close to the corners. We observe that a good proxy for the average position of the system over time during a climb of the AllC-TFT edge from the TFT corner to the separatrix is around \( n_{\text{AllC}} = 0.25 \) and \( n_{\text{TFT}} = 0.75 \) (yellow dot in Fig. 10b).

Within the remainder of the simplex, \( S_4 \) (blue), \( \Pr\{n_{\text{lock}} \in \text{AllD-b} \mid n_{\text{asy}} = s\} \) can be very heterogeneous, however, as discussed in the main text, this contributes very little to the outcome statistics for small values of \( N_0 \). We chose \( s_4 = \text{MS-fp} \) to represent this area (blue dot in Fig. 10b).

FIG. 10. Equivalence classes of the asymmetric regime. \( \Pr\{n_{\text{lock}} \in \text{AllD-b} \mid n_{\text{asy}} = s\} \) was numerically estimated via Gillespi-It\^o simulations for a subset of all points \( s \) (panel a). The results motivates the decomposition of the simplex into four equivalence classes, \( S_i \), as well as the choice of their representative point \( s_i \) (panel b). The parameters used for this illustration are \( b \# d = 0.05 \) and \( \log_{10} \mu = -6.5 \).
Fig. 10a was obtained with $b - d = 0.05$ and $\log_{10} \mu = -6.5$. Different values of these parameters produce different values of $\Pr\{n_{\text{lock}} \in \text{AllD-b} \mid n_{\text{asy}} = s\}$; however, the four areas can always be identified. This decomposition is validated by the successful reconstruction (see Figs. 5-6 in the main text as well as Figs. 12-19 of the all statistics of the evolutionary outcomes obtained via full simulations (Fig. 2 in the main text). The conditional probability decomposition is always very accurate except for the case of low mutation rate (e.g., $\log_{10} \mu < -6$), growth rates around $b - d \approx 0.05$ and large initial populations (e.g., $N_0 > 200$), for which it begins to produce less accurate results (Fig. 6 panel o in the main text).

Appendix K: Statistics of the stochastically-induced regime in detail

In this section we describe the effect of the mutation rate on the statistics of the stochastically-induced regime outcomes more in detail. Fig. 11a shows the same probabilities $p_{\text{stoc}}$ shown in Figs. 5b-e in the main text, but for a single value of $b - d = 0.026$ (we remind that $n_0 = \text{MS-fp}$ and $N_0 = 128$). We can see that all probabilities depend on $\mu$ for high mutation rates (i.e., $\log_{10} \mu > -4$), but only $p_{\text{stoc}}^{\text{S1}}$ and $p_{\text{stoc}}^{\text{S3}}$ depend on $\mu$ for lower mutation rates.

For a more detailed analysis, we decompose $p_{\text{stoc}}^{\text{S1}}$ into $p_{\text{stoc}}^{\text{S1}_{\text{AllD-c}}}$, the probability of $n_{\text{asy}}$ being the AllD corner, $p_{\text{stoc}}^{\text{S1}_{\text{AllC-c}}}$, the probability of $n_{\text{asy}}$ being the AllC corner, $p_{\text{stoc}}^{\text{S1}_{\text{AllC-TFT}}}$, the probability of $n_{\text{asy}}$ being in the AllD basin side of the AllC-TFT edge excluding the AllC corner and $p_{\text{stoc}}^{\text{S1}_{\text{other}}}$, the probability of $n_{\text{asy}}$ being anywhere else within the AllD basin. We can now see that $p_{\text{stoc}}^{\text{S1}_{\text{AllC-TFT}}}$ decreases with $\mu$, in contrast with $p_{\text{stoc}}^{\text{S1}_{\text{other}}}$, which instead increases with $\mu$. This is explained by the shape of the separatrix changing with $\mu$: the point where the separatrix meets the AllC-TFT edge moves towards the TFT corner as $\mu$ decreases (see panels c & d).

At the same time, we also see that the probability of the system being in one of the two corners at time $t_{\text{asy}}$, $p_{\text{stoc}}^{\text{S1}_{\text{AllD-c}}}$ and $p_{\text{stoc}}^{\text{S1}_{\text{AllC-c}}}$, are independent on the mutation rate for $\log_{10} \mu < -5$. However, for higher mutation rates it becomes more and more likely for the system to be in the AllD corner rather than in the AllC corner (see also panels c & d).

Finally, for mutation rates $\log_{10} \mu > -5$ the probability of the system being away from the boundaries of the simplex $p_{\text{stoc}}^{\text{S1}_{\text{other}}}$ becomes higher (the same of course can be observed for $p_{\text{stoc}}^{\text{S3}}$).
FIG. 11. **Statistics of the stochastically induced regime outcome \( p_{S_i}^{stoc} \) in detail.** Panel a shows the probabilities \( p_{S_i}^{stoc} \) over \( \mu \) with \( b - d = 0.026 \), \( n_0 = \text{MS-fp} \) and \( N_0 = 128 \). Panel b decomposes \( p_{S_1}^{stoc} \) into cases of the AllD corner, the AllC corner, the AllD basin side of the AllC-TFT edge (excluding the AllC corner) and the rest of the AllD basin. Panels c and d show the location of \( n_{taxy} \) within the simplex for the same 1000 stochastic trajectories used for estimating the probabilities in a and b, for the cases of \( \log_{10} \mu = -6.5 \) and \( \log_{10} \mu = -3.5 \) respectively. The size of the circles is proportional to the occurrences of \( n_{taxy} \) in the location and the colours reflect the different equivalence classes \( S_i \) (see panel a).

**Appendix L: Outcome statistics over \( b - d \) and \( \mu \): additional cases of \( n_0 \)**

Here we show that our decomposition of the outcome statistics based on equivalence classes holds not only for \( n_0 = \text{MS-fp} \) but in general. We do this by considering other three cases that are very different: \( n_0 = \text{centre} \), i.e., \( n_{\text{AllC}} = n_{\text{AllD}} = n_{\text{TFT}} = 1/3 \) (Fig. 12); \( n_0 = \text{TFT-c} \), i.e., \( n_{\text{AllC}} = n_{\text{AllD}} = 0, n_{\text{TFT}} = 1 \) (Fig. 13); and \( n_0 = \text{AllD-c} \), i.e., \( n_{\text{AllC}} = n_{\text{TFT}} = 0, n_{\text{AllD}} = 1 \) (Fig. 14). In all figures, comparing the results obtained via decomposition (panels n) against those obtained via full simulation (panels p) yields a low error (panels o), demonstrates the validity of our method. The average error over all the considered values of growth (\( b - d \)) and mutation (\( \log_{10} \mu \)) is approximately 0.029 for the \( s_0 = \text{centre} \), 0.003 for \( s_0 = \text{TFT-c} \), and 0.001 for \( s_0 = \text{AllD-c} \).
FIG. 12. Decomposition of outcome statistics based on equivalence classes (\(n_0 = \text{centre}\)). The statistics of evolutionary outcomes can be decomposed in terms of conditional probabilities that are based on four equivalence classes of states (panels a-i). The landscape of \(p_{\text{AID}}(\text{centre})\) for different rates of growth \((b - d)\) and mutation \((\mu)\) can be reconstructed with good agreement (panels j-n and panel o). The initial population size is \(N_0 = 128\). \(t_{\text{asy}}\) and \(t_{\text{lock}}\) are derived from the critical population sizes identified in Fig. 3 in the main text.
FIG. 13. Decomposition of outcome statistics based on equivalence classes ($n_0 = \text{TFT-c}$). The statistics of evolutionary outcomes can be decomposed in terms of conditional probabilities that are based on four equivalence classes of states (panels a-i). The landscape of $P_{\text{AllD}}(TFT)$ for different rates of growth ($b - d$) and mutation ($\mu$) can be reconstructed with good agreement (panels j-n and panel o). The initial population size is $N_0 = 128$. $t_{\text{asy}}$ and $t_{\text{lock}}$ are derived from the critical population sizes identified in Fig. 3 in the main text.
FIG. 14. Decomposition of outcome statistics based on equivalence classes ($n_0 = \text{AllD-c}$). The statistics of evolutionary outcomes can be decomposed in terms of conditional probabilities that are based on four equivalence classes of states (panels a-i). The landscape of $p_{\text{AllD}}(\text{AllD-c})$ for different rates of growth ($b - d$) and mutation ($\mu$) can be reconstructed with good agreement (panels j-n and panel o). The initial population size is $N_0 = 128$. $t_{\text{asy}}$ and $t_{\text{lock}}$ are derived from the critical population sizes identified in Fig. 3 in the main text.
In this section we illustrate how the outcome statistics in Fig. 2 panels b, c, d, e and g in the main text can be reconstructed using our probability decomposition method.

At high mutation rate (Figs. 2b-c in the main text), the asymmetric phase is very short. This is reflected in the probability decomposition in Figs. 15 and 16 where the probabilities in panels f-h are all zeros, i.e., the statistics of the outcome are given exclusively by the fixation of the system into one of the two basins during the stochastically induced phase (panels a). The accuracy of the probability decomposition is very high for both small (Fig 15) and large (Fig 16) initial populations, as illustrated in panels j.

At low mutation rate the asymmetric regime is very long, however, such regime has little effect on the outcome statistics for small initial populations. This is explained by the probability decomposition in Fig. 17. For some starting points \( N_0 \) the system can be in the equivalence class \( S_2 \) at the end of the stochastically-induced regime (panel b), however, from this area of the simplex it is extremely unlikely for the system to cross the separatrix towards the AllD basin, even if the asymmetric regime is long (panel f). Moreover, it is very unlikely for the system to be in \( S_4 \) at the end of the stochastically induced regime, since with a small initial population such regime is longer and fluctuations are likely to drive the systems towards the edges and corners (panels d). Thus, even if the system can cross the separatrix from \( S_4 \) during the asymmetric regime, the combined probability in panel h is approximately zero for every initial condition. Finally, we can see that for some initial state \( s \) the system has a small chance of being in \( S_3 \) at the end of the stochastically-induced regime (panel e). From this area the system is very likely to cross the separatrix during the asymmetric regime, leading to the small outcome probability contribution in panel g.

Since the probabilities in panel h are very low, the accuracy of the probability decomposition is very high (panel j) also in this case. The accuracy gets worse only for middle values of growth rate (around \( b - d = 0.05 \)), very small mutation rates (e.g., \( \log_{10} \mu < -6 \)) and large initial populations (e.g., \( N_0 > 200 \)). This case is reported in Fig. 6 in the main text, which shows that the probability of AllD outcomes is underestimated for some initial starts around the mixed-strategy unstable fixed point.

The results in Fig. 2 panels d and g can be reconstructed in a similar way, as shown in Figs. 18 and 19.

![Fig. 15. Decomposition of outcome statistics over \( n_0 \): \( \log_{10} \mu = -3.5 \), \( N_0 = 50 \) and \( b - d = 0.05 \). Panels a-d show the probability of the system being in the four equivalence classes \( S_i \) at the end of the stochastically-induced regime. The probabilities of AllD outcomes given that the system is in the representative points \( s_i \) at time beginning of the asymmetric regime are reported in the second row. Panels e-i show how the probabilities are combined to approximate the outcome probability and panel j shows accuracy of such approximation (cf. Fig. 2b in the main text).](image-url)
FIG. 16. Decomposition of outcome statistics over $n_0$: log$_{10} \mu = -3.5$, $N_0 = 250$ and $b - d = 0.05$. Panels a-d show the probability of the system being in the four equivalence classes $S_i$ at the end of the stochastically-induced regime. The probabilities of AllD outcomes given that the system is in the representative points $s_i$ at time beginning of the asymmetric regime are reported in the second row. Panels e-i show how the probabilities are combined to approximate the outcome probability and panel j shows accuracy of such approximation (cf. Fig. 2c in the main text).

FIG. 17. Decomposition of outcome statistics over $n_0$: log$_{10} \mu = -6.5$, $N_0 = 50$ and $b - d = 0.05$. Panels a-d show the probability of the system being in the four equivalence classes $S_i$ at the end of the stochastically-induced regime. The probabilities of AllD outcomes given that the system is in the representative points $s_i$ at time beginning of the asymmetric regime are reported in the second row. Panels e-i show how the probabilities are combined to approximate the outcome probability and panel j shows accuracy of such approximation (cf. Fig. 2e in the main text).
FIG. 18. Decomposition of outcome statistics over $N_0$: $\log_{10} \mu = -3.5$ and $b - d = 0.05$. The results are shown for the two initial states $s'$ and $s''$ in Fig. 2 in the main text. Panels a-d show the probability of the system being in the four equivalence classes $S_i$ at the end of the stochastically-induced regime. The probabilities of AllD outcomes given that the system is in the representative points $s_i$ at time beginning of the asymmetric regime are reported in the second row. Panels e-i show how the probabilities are combined to approximate the outcome probability and panel j shows accuracy of such approximation (cf. Fig. 2d in the main text).

FIG. 19. Decomposition of outcome statistics over $N_0$: $\log_{10} \mu = -6.5$ and $b - d = 0.05$. The results are shown for the two initial states $s'$ and $s''$ in Fig. 2 in the main text. Panels a-d show the probability of the system being in the four equivalence classes $S_i$ at the end of the stochastically-induced regime. The probabilities of AllD outcomes given that the system is in the representative points $s_i$ at time beginning of the asymmetric regime are reported in the second row. Panels e-i show how the probabilities are combined to approximate the outcome probability and panel j shows accuracy of such approximation (cf. Fig. 2g in the main text).
P. K. & Chapman, S. J. Mathematical modeling of cell population dynamics in the colonic crypt and in colorectal cancer. *Proceedings of the National Academy of Sciences* **104**, 4008–4013 (2007).

[32] Lieberman, E., Hauert, C. & Nowak, M. A. Evolutionary dynamics on graphs. *Nature* **433**, 312–316 (2005).

[33] Hindersin, L. & Traulsen, A. Most undirected random graphs are amplifiers of selection for birth-death dynamics, but suppressors of selection for death-birth dynamics. *PLoS computational biology* **11**, e1004437 (2015).

[34] Marrec, L., Lamberti, I. & Bitbol, A.-F. Toward a universal model for spatially structured populations. *Physical Review Letters* **127**, 218102 (2021).

[35] Gardiner, C. W. *Handbook of stochastic methods* (Springer, 2003), 3rd edn.

[36] Gillespie, D. T. The chemical Langevin equation. *The Journal of Chemical Physics* **113**, 297–306 (2000). URL [http://aip.scitation.org/doi/10.1063/1.481811](http://aip.scitation.org/doi/10.1063/1.481811)

[37] Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry* **81**, 2340–2361 (1977).