



HIER

Harvard Institute of Economic Research

Discussion Paper Number 2177

Stochastic Stability in Large Populations

by

Drew Fudenberg and Daniel Hojman

July 2009

Harvard University
Cambridge, Massachusetts

This paper can be downloaded without charge from:
<http://www.economics.harvard.edu/journals/hier2009>

The Social Science Research Network Electronic Paper Collection:
<http://ssrn.com/abstract=1434592>

Stochastic Stability in Large Populations

Drew Fudenberg
Harvard University

Daniel Hojman
Harvard University

This Version: July 10, 2009

First Version: March 1, 2007

Abstract. Most work in evolutionary game theory analyzes a deterministic adjustment process on a continuum of agents. However, both the assumption of a continuum and that of no randomness are approximations, so it is important to study the behavior of adjustment processes on a large but finite population subject to small but persistent stochastic shocks. This paper characterizes the properties of the invariant distribution of birth-death processes in the double limit as the population becomes infinitely large and the perturbation vanishingly small. We show that the order of these limits does not change the conclusions for processes with “strong basins,” which is the case when the unperturbed process is deterministic. In contrast, the order of limits does matter for processes with “weak basins,” where the unperturbed process is stochastic except at a finite number of points.

Keywords: equilibrium selection, stochastic stability, large deviations

JEL Classification: C62, C73

We thank Mihai Manea, Nolan Miller, and William Sadholm for helpful comments and suggestions, NSF grant SES-0426199 for financial support, and Adam Clark-Joseph and Sungchun No for careful proof-reading.

1. Introduction

Most work in evolutionary game theory analyzes a deterministic adjustment process on a continuum of agents. However, both the assumption of a continuum and that of no randomness are approximations, so it is important to study the behavior of stochastic adjustment processes on a large but finite population. We study the case where the noise term makes the finite population process ergodic, and investigate when the limiting distribution is independent of the relative rates at which the population grows and the noise vanishes.

It is easy to see that the rates do sometimes matter, as noted by Binmore and Samuelson [1997]. This raises several questions. First, when does sending the population size to infinity for fixed noise and then sending the noise to zero yield the same selection predictions as the reverse iterated limit, where first the noise goes to zero and then the population grows?¹ Second, when the two iterated limits differ, can other forms of the double limit yield yet a third conclusion? Finally, what sorts of rates of convergence for noise and population yield the same results as the iterated limits?

Answering even the first question for a general class of adjustment processes is difficult, as for each order of limits described the available results require solving a potentially complex minimization problem.² This paper provides answers for the case of birth-death processes, which are one-dimensional finite-state processes that move at most one step per period.³ In this setting we can explicitly find the outcome for each of the iterated limits mentioned above, and more generally for all double limits where the noise term vanishes and the population size grows to infinity.

¹ Kandori, Mailath and Rob (KMR) [1993], Young [1993], and Ellison [1995] send the noise to zero. Ellison explicitly discusses the role of population size on the speed of convergence. Fudenberg and Imhof [2007] consider the iterated limit where first the noise goes to zero and then the population size grows to infinity, while Binmore and Samuelson [1997] and Benaïm and Weibull [2003] send the population size to infinity with a fixed level of noise.

² See Benaïm and Weibull [2003] for the calculus-of-variations minimization needed to characterize the large-population limit. Ellison [2000] gives a sufficient condition for a unique equilibrium to be selected as noise vanishes, and Fudenberg and Imhof [2007] characterize this no-noise limit for a class of “imitation processes.”

³ Birth-death processes have proved useful in a number of settings in biology, operations research, and economics. For example, the class of processes considered by KMR includes the birth-death case, and Binmore and Samuelson [1997], Binmore, Samuelson and Vaughn [1995], and Fudenberg et al [2006] consider only birth-death processes. In all of these papers agents choose among two possible actions, as is the case in our applications.

We show that the qualitative nature of the result depends on the nature of the “unperturbed” stochastic process obtained by sending the noise term to zero with the population held fixed. We say that the process has *strong basins* if, roughly, the unperturbed process is deterministic in a neighborhood of the steady states, as it is in Kandori, Mailath, and Rob [1993] (KMR). Note that in this case the “radius” of each basin of attraction (Ellison [2000]) of the finite-population no-noise process is proportional to the population size, so that standard “mutation-counting” arguments are typically sufficient to determine the stochastically stable set. We show that when the process has strong basins, all methods of sending the noise term to zero and the population size to infinity give the same result.

At the other extreme, we say that the process has *weak basins* if the unperturbed process has strictly positive probabilities of both upward and downward transitions except at a finite number of isolated points, as in Binmore-Samuelson [1997] and Fudenberg et al [2006].⁴ With weak basins, each steady state of the unperturbed process has radius 1, and mutation-counting is not sufficient to determine the stochastically stable set. In this case, we show that the order of limits matters. Specifically, the limit distribution computed by first sending the noise term to zero applies if the noise shrinks at a suitable exponential rate in the population size. If the noise shrinks more slowly than this, the result of the double limit is the same as in the iterated limit when the population size grows to infinity first. This implies that all of the double limits have the same outcome as one of the iterated ones.

Our setup is closely related to that of Benaim and Weibull [2003] (BW) who study the limits of a large class of population games as the population size N goes to infinity. For the case of irreducible processes, their results show that in this limit the system spends most of the time in a neighborhood of an attractor of the deterministic “mean flow” dynamics associated to the stochastic process. In this paper, we allow for processes that are not irreducible and incorporate persistent randomness in the form of an ε -perturbation which vanishes as $\varepsilon \rightarrow 0$ as in KMR. Thus, in contrast to BW, our selection results can be applied to imitation processes and best-response dynamics.

⁴ Note that these two cases are not exhaustive, but they do cover all of the birth-death processes in the literature.

Further, by specializing to the birth-death case we are able to provide an explicit calculation of the limit distribution of the process as $(N, \varepsilon) \rightarrow (\infty, 0)$.⁵ We show that in the limit as $(N, \varepsilon) \rightarrow (\infty, 0)$ the limit distribution puts probability one on a set which consists either of absorbing states of the “no-mutation” process (obtained by sending $\varepsilon \rightarrow 0$) or a subset of attractors of the mean flow dynamics associated to the process. Sandholm [2009] studies birth-death processes arising from population games; he shows that the order of limits does not matter for a class of “noisy best response” adjustment rules.

We analyze the case of strong basins in Section 5. Theorem 2 shows that the limit distribution puts weight on a subset of attractors contained in the stable set with the largest basin. We illustrate these results for a version of KMR and Auriol and Benaïm’s [2001] model of platform standardization.

Section 6 analyzes processes with weak basins, focusing on the case where transition probabilities are strictly positive everywhere in the interior (i.e. only the endpoints corresponding to all agents using the same action can be absorbing). For these processes, selection depends on the balance between two forces: the “mutation cost” associated with exiting an absorbing point, and the “viscosity cost,” which captures the likelihood of transitions between points in the interior. Theorem 3 shows that for any sequence $(N, \varepsilon) \rightarrow (0, \infty)$, the support of the limiting invariant distribution is either an absorbing state or an attractor with undominated viscosity cost. As a consequence all versions of the limit $(N, \varepsilon) \rightarrow (\infty, 0)$ give the same answer when the absorbing states of the unperturbed process are undominated attractors of the mean field. When the absorbing states of the unperturbed process are not attractors of the mean field, the limit distribution must depend on the relative rates at which $\varepsilon \rightarrow 0$ and $N \rightarrow \infty$, as sending $\varepsilon \rightarrow 0$ first concentrates probability mass on absorbing states of the no-mutation process. Theorem 4 shows that the key is whether $\varepsilon < e^{-NW}$ for some constant $W > 0$, and identifies which limits are obtained in each case. We apply the results to the process in Fudenberg et al [2006] and discuss how the conclusions generalize for a large class of evolutionary games.

⁵ BW provide sufficient conditions to characterize the support of the distribution in the limit as $N \rightarrow \infty$ that require solving a calculus of variations problem. See for example, Proposition 6 therein.

2. Motivating Examples

In each of the examples considered in this section there is a population of N individuals who have a binary choice between alternatives A and B , and the state of the system s is the fraction of the population using alternative A . Each period, a single agent is selected to update her choice. The “no mutation” or limit process is described by probability maps $p(s)$ and $q(s)$ that specify the probabilities that an agent chooses A respectively B at state s . In our examples, these probabilities are a function of the payoffs associated to each action, which are themselves a function of the state, i.e., the share of people who use A . We show how different modeling assumptions and comparative statics in the underlying primitives of an economic model give rise to processes of different nature (irreducible, weak or strong basins), thus highlighting the importance of providing “selection” results that cover all of these cases simultaneously.

2.1 Smooth Best-Response Dynamics in a 2x2 game

As in KMR and Young [1993] agents derive their payoffs from an underlying symmetric 2×2 game Γ with payoff matrix

	A	B
A	α, α	γ, δ
B	δ, γ	β, β

The payoff to each action if a share s of the population uses action A are $u_A(s) = \alpha s + \gamma(1-s)$ and $u_B(s) = \delta s + \beta(1-s)$. A large class of processes studied in evolutionary game theory considers transition maps such that, at each s , the ratio $p(s)/q(s)$ increases with the payoff difference $\Delta u(s) = u_A(s) - u_B(s)$ and $p(s) + q(s) = 1$. Intuitively, in the absence of noise, the choice of an agent who revises is based on a simple payoff comparison. This means that for some decreasing function λ

$$p(s) = \frac{1}{1 + \lambda(\Delta u(s))} \text{ and } q(s) = \frac{\lambda(\Delta u(s))}{1 + \lambda(\Delta u(s))}.$$

The logit case is $\lambda(\Delta u) = e^{-k\Delta u}$; this corresponds to the payoffs being subject to random shocks that have the extreme values distribution. We will consider a larger class of response functions, corresponding to more general distributions of shocks: We say that a rule is a ν -best response if there is $\nu > 0$, and a decreasing continuous map $\phi: (-\nu, \nu) \rightarrow \mathbb{R}$ such that

$$\lambda(\Delta u) = \begin{cases} \infty & \text{if } \Delta u \leq -\nu \\ \phi(\Delta u) & \text{if } \Delta u \in (-\nu, \nu) \\ 0 & \text{if } \Delta u \geq \nu \end{cases} \quad \text{so that } p(s) = \begin{cases} 0 & \text{if } \Delta u(s) \leq -\nu \\ \pi(\Delta u(s)) & \text{if } \Delta u(s) \in (-\nu, \nu) \\ 1 & \text{if } \Delta u(s) \geq \nu \end{cases}$$

where π is any c.d.f. with support $[-\nu, \nu]$.⁶

Since payoffs are linear in s , if Γ does not have a dominant strategy, there exists a threshold value $x^* \in (0, 1)$ such that $\Delta u(x^*) = 0$. For small enough ν , there exist x_ν^-, x_ν^+ with $x_\nu^- \leq x^* \leq x_\nu^+$ such that x_ν^- and x_ν^+ are the solutions to $|\Delta u(x)| = \nu$ and $\lim_{\nu \rightarrow 0} x_\nu^- = \lim_{\nu \rightarrow 0} x_\nu^+ = x^*$. The process “moves” either to the left for $s \in [0, x_\nu^-)$ ($p = 0$) and to the right for $s \in [x_\nu^+, 1]$ ($q = 0$), or else to the right for $s \in [0, x_\nu^-)$ ($q = 0$) and to the left for $s \in [x_\nu^+, 1]$ ($p = 0$). Thus, the “no mutation” process is not irreducible and one of the transition probability maps is zero on an interval of positive length.

The perturbed version of the model is characterized by a small mutation rate $\varepsilon > 0$, often interpreted as a probability of experimentation or some persistent noise unobservable to the modeler. The transition probabilities for this process are described by $P^\varepsilon(s) = (1 - \varepsilon^a)p(s) + \varepsilon^a$ and $Q^\varepsilon(s) = (1 - \varepsilon^b)q(s) + \varepsilon^b$, where $a, b > 0$. This process is irreducible for any p and q .

2.2 Standardization with Preferences for Conformity and Vanity

Our second example considers a process of “platform” standardization as in Auriol and Benaïm [2001]. The utility to each action $j \in \{A, B\}$ is given by $u_j(s_j) = U^c(s_j) + U^v(s_j)$ where s_j is the share using action j , U^c is a “taste for conformity” term that increases with the share using the platform, and U^v is a “vanity”

⁶ The logit model has $\nu = \infty$. Note that as $\nu \rightarrow 0$ the process approaches a best-response rule

term that decreases with the share using the platform.⁷ For concreteness, suppose that $U^c(s) = \log s$ and $U^v(s) = -\lambda s$ for some $\lambda > 0$. In contrast to the game considered in our first example the payoffs are unbounded. In particular, as illustrated in the figure, $\Delta u(s) = u_A(s_A) - u_B(s_B) = \log s/(1-s) - \lambda(2s-1)$ is “dominated” by the incentive to conform for states close to extremes (full standardization). At the same time, for intermediate states, the payoff comparison is “dominated” by vanity as the utility of the less popular standard exceeds the one of the most widely used. Hence, to the left of the equal share state $s = 1/2$ there is a higher probability of moving to the right than to the left and this order reverses to the right of this point. See Figure 1.

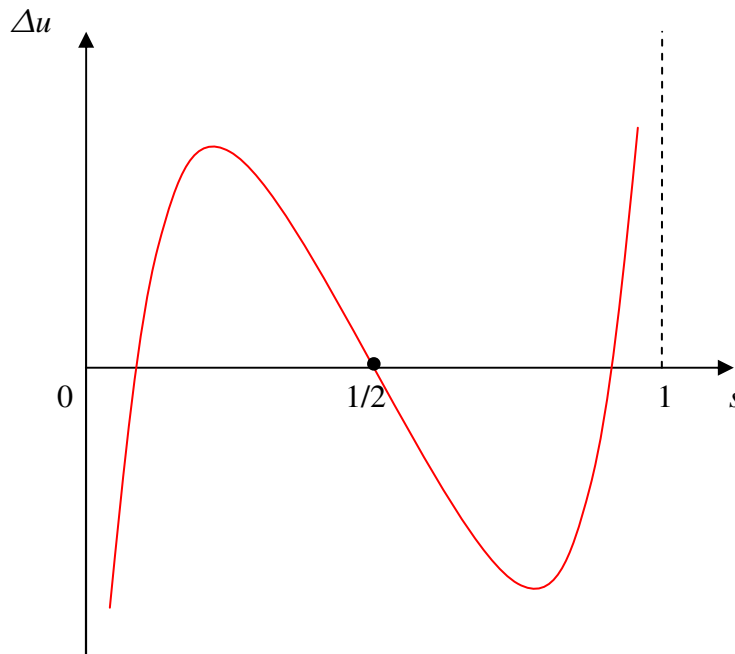


Figure 1

As in our first example, the transition probabilities follow a smooth best response. In contrast to our previous example, in the logit case the limit process has two absorbing points, the corners 0 and 1. At these two points one of the transition maps vanishes but

⁷ See Grilo, Thisse and Shy [2001] for a model of differentiation and pricing where consumers have preferences of this form. Note that these payoffs need not be linear in the population, so this is a population game as in Sandholm [2009].

transition probabilities are positive in the interior. As seen later, the state $s=1/2$ representing full diversity can also be “selected” as $(N, \varepsilon) \rightarrow (\infty, 0)$ depending on the value of λ .

2.3 Imitation Dynamics with mutations

We consider a version of the imitation dynamics in Fudenberg et al [2006]. Consider a 2×2 symmetric game Γ with strictly positive payoffs. We use the same notation as in the KMR example above. The fitness of an individual using strategy A is $u_A(s) = \alpha s + \gamma(1-s)$, which is the payoff of this strategy against the population *including himself*. Similarly, the fitness of individuals using strategy B is $u_B(s) = \delta s + \beta(1-s)$.⁸ The distribution of strategies in the population evolves as follows. Each period, each agent “reproduces” at a rate proportional to its fitness. One offspring is chosen at random to enter the population, so that the probability of adding an A offspring is $\frac{su_A(s)}{\bar{u}(s)}$, where $\bar{u}(s) = su_A(s) + (1-s)u_B(s)$ is the average fitness. Simultaneously, a randomly chosen agent is removed. Hence, the probability that the number of agents using A increases by one is $1-s$ -the probability of removing an agent using B - times the probability of adding an A offspring. Introducing mutations, this leads to the process with transition maps

$$P^\varepsilon(s) = (1-s) \left[\frac{su_A(s)}{\bar{u}(s)} (1-2\varepsilon) + \varepsilon \right] \text{ and } Q^\varepsilon(s) = s \left[\frac{(1-s)u_B(s)}{\bar{u}(s)} (1-2\varepsilon) + \varepsilon \right].$$

The transition maps of the limit process are $p(s) = \frac{(1-s)su_A(s)}{\bar{u}(s)}$ and $q(s) = \frac{(1-s)su_B(s)}{\bar{u}(s)}$.

Regardless of the payoff function, the corners are the only absorbing states of the “no mutation” process. The process is “positive in the interior”, i.e., it has strictly positive transition probabilities to the left and right for population shares that are not the extremes.

⁸ Fudenberg et al suppose that the fitness instead comes from playing each *other* individual exactly once, so that the fitness functions and thus probabilities of transition differ from the ones we analyze by a term proportional to $1/N$. Such processes are not covered by the assumptions in this paper, but they are covered by the extension of the results in a previous version.

3. Basic Assumptions and Preliminary Results

3.1 Basic Assumptions

We study a family of dynamic binary choice processes in a population of N agents, modeled as a finite-state Markov chain. In each period $t \in T = \{0, 1, 2, \dots\}$, each agent uses one of two alternatives A or B , and the state of the system $s \in S^N = \left\{0, \frac{1}{N}, \dots, \frac{N-1}{N}, 1\right\}$ represents the fraction of the population using action A . Each process $Z^{N, \varepsilon} = \{Z_t^{N, \varepsilon}\}_{t \in T}$ depends both on the size of the population N and on the size of a perturbation ε . The allowed parameters are integers N above some lower bound $\underline{N} > 0$, and perturbations ε less than some $\bar{\varepsilon} < 1$. We denote the set of parameters by Θ .

We consider time-homogeneous birth-death processes, where at most one individual changes their action in any period. For each $(N, \varepsilon) \in \Theta$ and each state $s \in S^N$, let $P^\varepsilon(s)$ be the transition probability from s to $s + 1/N$ and $Q^\varepsilon(s)$ be the probability of a transition from s to $s - 1/N$. These probabilities are assumed to be independent of the size of the population; this simplifies some of the analysis, but is not essential to our results.⁹ For a fixed population size the probabilities are only defined on the finite grid of possible states, but we assume that $P^\varepsilon(\cdot)$ and $Q^\varepsilon(\cdot)$ are defined on the whole $[0, 1]$ interval.¹⁰ We use x for a generic point in $[0, 1]$ and s for a generic state in S^N , and $s^N(x)$ for the closest state in S^N to x . To make sure the process is well defined we require $P^\varepsilon(0) = 1$ and $Q^\varepsilon(0) = 0$.

⁹ In a previous version we allow for the probabilities to depend on the population size in a “well-behaved” way. Benaïm and Weibull [2003] also assumed that switching probabilities are independent of population size; this is relaxed in Benaïm and Weibull [2009].

¹⁰ In our examples these continuous functions arise naturally because the transition probabilities depend on the payoffs associated to each of the actions at a given state.

Assumption (P1): The maps P and Q are continuous on $[0, \bar{\varepsilon}] \times (0, 1)$ and continuously differentiable in x . The partial derivative maps $D_x P = \frac{\partial P}{\partial x}$ and $D_x Q = \frac{\partial Q}{\partial x}$ are also continuous on $[0, \bar{\varepsilon}] \times (0, 1)$.

We are primarily interested in the behavior of these maps as the perturbation vanishes, i.e., $\varepsilon \rightarrow 0$. The Markov process $Z^{N,0}$ on state space S^N with transition probability maps $p = P^0$ and $q = Q^0$ is referred as the *limit* or *unperturbed process*.

Assumption (P1) ensures that P^ε , Q^ε , $\frac{dP^\varepsilon}{dx}$ and $\frac{dQ^\varepsilon}{dx}$ converge uniformly on $[0, 1]$ to p , q , $\frac{dp}{dx}$, and $\frac{dq}{dx}$ respectively as $\varepsilon \rightarrow 0$.

We will need to keep track of the subsets of the state space where some of the limit probabilities are 0. To do this, Let $L = \{x \in [0, 1] \mid p(x) = 0\}$ and $R = \{x \in [0, 1] \mid q(x) = 0\}$. For any state in L transitions the process can only go to the “left,” i.e. to lower states. Similarly, for states in R the process can only move to the “right,” i.e. to higher states. Let $X_0 \equiv L \cup R$ and $X_+ \equiv [0, 1] \setminus X_0$; the latter is the set of points where the limit transition probabilities are strictly positive.

Assumption (P2): For each $\varepsilon \in (0, \bar{\varepsilon}]$:

- (a) P^ε is strictly positive on $[0, 1]$. Further, there exist constants $a > 0$ and $k > 0$ such that $P^\varepsilon(x) = p(x) + O(\varepsilon^a)$ for all $x \in [0, 1]$ and $P^\varepsilon(x) \geq k\varepsilon^a$ for all $x \in L \setminus \{1\}$.
- (b) Q^ε is strictly positive on $(0, 1]$. Further, there exist constants $b > 0$ and $h > 0$ such that $Q^\varepsilon(x) = q(x) + O(\varepsilon^b)$ for all $x \in [0, 1]$ and $Q^\varepsilon(x) \geq h\varepsilon^b$ for all $x \in R \setminus \{0\}$.

Assumption (P2) has two main implications. First, it says the family of perturbed processes ($\varepsilon > 0$) are irreducible for each $(N, \varepsilon) \in \Theta$.¹¹ Hence, the process $Z^{N, \varepsilon}$ has a unique invariant distribution $\mu^{N, \varepsilon}$, which we view as a measure on the Borel subsets of the interval $[0, 1]$. Second, if we think of $Z^{N, 0}$ as a “base” process, the assumption says the transition probabilities of perturbed process $Z^{N, \varepsilon}$ differ from those of $Z^{N, 0}$ by a term of some order of ε . In regions where the unperturbed process has probabilities equal to zero the approximation is exact. Both of these restrictions are standard in the stochastic stability literature (See Ellison [2000]).¹²

We denote by $X_+ = [0, 1] \setminus (L \cup R)$ the set of points where the limit transition probabilities are strictly positive. Let $E = [x, y]$ be a closed interval with the convention that E is a singleton if $x = y$. The interval E is said to be *limit recurrent* if p and q are both positive in the interior of E ($\text{int}(E) \subseteq X_+$), the probability of a transition to the left at x is zero ($x \in R \cup \{1\}$), and the probability of a transition to the right at y is zero ($y \in L \cup \{0\}$).¹³ Observe that limit recurrent sets are the recurrent sets of the “limit process” on the continuum population where transitions are governed by p and q .

We will study two types of limit processes; these cases are not mutually exclusive but they cover all of the applications in the literature that we know of.

Definition 1. The limit process has *strong basins* if L and R can be expressed as a finite union of non-overlapping intervals and X_0 has positive Lebesgue measure.

The results in Section 5 assume that the limit process has strong basins.

Definition 2. The limit process has *weak basins* if X_0 is finite.

¹¹ The fact that the process is irreducible follows from the fact that the transition maps P^ε and Q^ε are strictly positive in $(0, 1)$, $P^\varepsilon(0) > 0$, and $Q^\varepsilon(1) > 0$. Note that this allows for $P^\varepsilon(1) = 0$ or $Q^\varepsilon(0) = 0$.

¹² In a previous version we show that our results hold if we allow for the more general case of state-dependent “mutations” as in Bergin and Lipman [1996].

¹³ Note that if $E = \{x\}$ is a singleton that is limit recurrent, it is either an interior point of $X = [0, 1]$ that belongs to $L \cap R$, or else $x = 0$ and $0 \in L$, or else $x = 1$ and $1 \in R$.

Processes with weak basins are analyzed in Section 6. Imitation processes in which the corners are absorbing points and all other points are transient fall in this category.

Remark: With either strong or weak basins, there is a finite number of limit recurrent sets; each limit recurrent set is either a closed interval or a point, and they do not intersect.¹⁴ It is straightforward to extend our results for case in which X_0 consists of locally isolated singleton and/or a union of intervals, i.e., a case that “combines” strong and weak basins. Processes with basins that are neither strong, nor weak nor a combination of both, can have some limit recurrent sets that are non-isolated singletons.

3.2 An Approximation Lemma and Preliminary Results

Lemma 1 gives an approximation to the invariant distribution that is central to our results. We start by introducing two “probability cost” functions that measure the relative likelihood of a transition between two states, one appropriate for the region where the limit process has strictly positive transition probabilities and one for the region where some of the probabilities are zero. Let $W \subseteq [0,1]$ and denote by χ_W the characteristic function for W . That is, $\chi_W(x) = 1$ if $x \in W$ and $\chi_W(x) = 0$ if $x \notin W$.

For each $\varepsilon \in (0, \bar{\varepsilon}]$, define $V^\varepsilon : [0,1] \times [0,1] \rightarrow \mathbb{R}$ by

$$V^\varepsilon(x, y) = - \int_x^y \chi_{X_+}(z) \log \frac{P^\varepsilon(z)}{Q^\varepsilon(z)} dz \quad (1)$$

If $x, y \in S^N$, this expression measures the likelihood or “probability cost” of a transition from x to y relative to a transition from y to x for the perturbed process $Z^{N,\varepsilon}$ in the region X_+ . Similarly, let

$$V(x, y) = - \int_x^y \chi_{X_+}(z) \log \frac{p(z)}{q(z)} dz. \quad (2)$$

¹⁴ The finite number of these sets follows from the observation that a limit recurrent set is delimited by points which are either in $L \cup \{1\}$ or $R \cup \{0\}$. In the case of weak basins, these are finite sets. In the case of strong basins if a delimiter of a limit recurrent set is contained in L or R , it must be an extreme points of one of the intervals, and by definition there is a finite number of them. Finally, the non-intersection comes from the requirement that the transition maps are strictly positive in the interior of a limit recurrent set.

In the Appendix we show that $V^\varepsilon \rightarrow V$ uniformly as $\varepsilon \rightarrow 0$ and that both of these functions are uniformly bounded. We refer to V as the *viscosity cost function*.

For each $x, y \in [0, 1]$, $x \leq y$ we define the *relative mutation cost from x to y* as

$$\Sigma^N(x, y) = \sum_{s \in S^N; x \leq s < y} a \chi_L(s) - \sum_{s \in S^N; x < s \leq y} b \chi_R(s). \quad (3)$$

If $x \geq y$, the relative mutation cost from x to y is $\Sigma^N(x, y) = -\Sigma^N(y, x)$. The function Σ^N estimates the resistance –the order of ε – of transition from x to y minus the resistance for a path from y to x . If the limit process $Z^{\infty, 0}$ is irreducible then $\Sigma^N \equiv 0$. With strong basins, $\|\Sigma^N\| = O(N)$ and with weak basins $\|\Sigma^N\| = O(1)$, where $\|\cdot\|$ denotes the sup norm.

Lemma 1. *The invariant distribution $\mu^{N, \varepsilon}$ of $Z^{N, \varepsilon}$ satisfies*

$$\frac{\mu^{N, \varepsilon}(s')}{\mu^{N, \varepsilon}(s)} = \exp\left(N(V^\varepsilon(s', s) + r^N(s, s')) - (\log \varepsilon)\Sigma^N(s', s) + O(1) |s' - s|\right)$$

for all $s, s' \in S^N$ with $s \leq s'$,

where $r^N(s, s') = 0$ if $(s, s') \subseteq X_+$ and $r^N(s, s') = O\left(\frac{\log N}{N}\right)$ otherwise.

The proof is in the Appendix.; it combines the detailed-balance property of the invariant distribution with our regularity assumptions about how the process behaves for large N and small ε . Observe that if the limit process is irreducible then $\Sigma^N(s, s') = r^N(s, s') \equiv 0$ for small ε , and the limiting behavior of the formula in the lemma is governed by the sign of $V^\varepsilon(s, s')$ as $\varepsilon \rightarrow 0$. In contrast, with strong basins, since V^ε is bounded, the limiting behavior is governed by $\Sigma^N(s', s)$ which is of order N . With weak basins, the limit of invariant distribution balances both the viscosity and relative mutation costs from one state to another.

4. Mean Flow Equation and Irreducible Processes

If a recurrent set E is an interval of positive length instead of a point, then its interior is contained in X_+ .

Definition 3. Let I be a closed interval such that $I \subseteq \text{int}(X_+)$. The viscosity of $z \in I$ in I is defined as $\phi_I(z) = \inf_{z' \in I} V(z, z')$. The maximum viscosity set in I is defined as $M_V(I) = \arg \max_{z \in I} \phi_I(z)$.

Lemma 1 implies that the following:

Theorem 1. Suppose the limit process $Z^{\infty,0}$ is irreducible ($X_+ = [0,1]$). Then, for any sequence $(N, \varepsilon) \rightarrow (\infty, 0)$, $\lim_{(N, \varepsilon) \rightarrow (\infty, 0)} \mu^{N, \varepsilon}(M_V[0,1]) = 1$.

The proof is omitted, the result follows from Theorem 3 in section 6. Note that this characterization is only useful when we are able to calculate $M_V[0,1]$. We do so by establishing the connection between this set and mean flow dynamics associated to the limit process. Formally, let $F(x) = p(x) - q(x)$ be the mean flow associated with the limit process. Note that at each state s the limit process moves to the right with probability $p(s)$ and to the left with probability $q(s)$. Hence, the mean flow measures the expected flow at each state. Our main results show that the support of the invariant $\mu^{N, \varepsilon}$ as $(N, \varepsilon) \rightarrow (\infty, 0)$ is either a subset of a recurrent set or a subset of the attractors¹⁵ of the deterministic mean flow equation

$$\dot{x} = \widehat{F}(x), \quad (4)$$

where $\widehat{F}(x) = F(x)$ for all $x \in (0,1)$, $\widehat{F}(0) = \max\{0, F(0)\}$, and $\widehat{F}(1) = \min\{0, F(1)\}$.

We assume that mean flow is well-behaved in the following sense.

¹⁵ The ω -limit set of any interval J is $\omega(J) = \bigcap_{t \geq 0} \{x(t) \mid x(0) \in J\}$. A compact set K is an attractor if there exists a neighborhood U such that $K = \omega(U)$.

Assumption (P3). The mean flow $F = p - q$ has regular attractors: each attractor of the differential equation (4) is a singleton, and there exists $d > 0$ such that for any attractor x^* in the interior of $[0,1]$ we have $F'(x^*) \leq -d$.

Under assumption (P3) an attractor of F is an asymptotically stable steady state of (4) and it is locally isolated from other attractors. The smoothness and continuity conditions in Assumption (P1) ensure that these nice properties extend to the “perturbed” mean flow $F^\varepsilon(x) = P^\varepsilon(x) - Q^\varepsilon(x)$.

The following Lemma establishes a connection between $M_V(E)$ and the set of attractors A_E of the mean flow $F = p - q$ when the differential equation is restricted to the interval E .

Lemma 2. Suppose that E is a recurrent set and assume it has a strictly positive length. Let $M_V(E)$ be the maximum viscosity set associated to V in E , A_E be the set of attractors of the unperturbed mean flow $F = p - q$ on E , and $A = A_{[0,1]}$.

- (i) $M_V(E) \subseteq A_E \subseteq A$.
- (ii) If $m \in M_V(E)$ then $V(m, y) \geq 0, \forall y \in E$, with equality only if $y \in M_V(E)$.

The proof is in the Appendix. The fact that the attractors of the mean flow F restricted to E are attractors of the same flow in $[0,1]$ uses that E is a recurrent set and does not hold for an arbitrary $I \subset X_+$. The Lemma provides an explicit way to compute $M_V(E)$: (i) find the set attractors A_E of the mean flow in E (e.g. interior attractors satisfy $F(x) = 0$ and $F'(x) < 0$); (ii) if $A_E = \{x_1, \dots, x_M\}$ is not a singleton, calculate $V(x_k, x_l)$ for all pairs of attractors in A_E .

Theorem 1 in combination with Lemma 2 provides a complete characterization of the limiting invariant distribution when the “base process” $Z^{\infty,0}$ is irreducible. For illustration, consider the smooth best-response dynamics introduced in section 2.1. Since, p is increasing in the payoff difference Δu between using A versus B , and $q = 1 - p$, the mean flow is also increasing in Δu . It is immediate that each steady state of the mean

flow corresponds to a symmetric Nash equilibrium of the game Γ . If the game has a unique Nash equilibrium then $M_V[0,1]$ is simply the associated state and no additional calculation is needed. The only case requiring additional analysis is a coordination game, which has three symmetric equilibria: coordination on A (associated to $x=1$), coordination on B (associated to $x=0$), and mixed strategy equilibrium (associated to x^* such that $\Delta u(x^*)=0$). For the logit model, the viscosity function is simply

$$V(x, y) = -\int_x^y \log \frac{1}{\lambda(\Delta u(z))} dz = \int_x^y \Delta u(z) dz,$$

where the integrand is linear. It is straightforward to check that if $x \in \{0,1\}$ then $V(x, x^*) > 0$, so only the corners can be selected. Furthermore, it is straightforward to check that $V(0,1) \geq 0$ if and only if action B is risk dominant.

5. Processes with Strong Basins

5.1 Analysis

Many applications of stochastic stability have been to processes with strong basins, where it is sufficient to count the number of mutations required to escape from each recurrent class. In this case, the limiting behavior of the formula in Lemma 1 is governed by the relative mutation cost term Σ^N which is of order N . To see this, fix N , consider two states $s, s' \in S^N$ with $s < s'$, and let $\varphi^N(s, s')$ denote the cost of mutations from s to s' . Note that this cost is zero unless the interval $[s, s']$ intersects L . For a process with strong basins, if this intersection is non-empty it is an interval or a disjoint union of intervals. For large N , the number of states in this intersection is roughly equal to N times the length of the intervals in this intersection. Thus, $\varphi^N(s', s) \approx Na \int_s^{s'} \chi_L(z) dz$.

Similarly, the cost of mutations from s' to s satisfies $\varphi^N(s, s') \approx Nb \int_s^{s'} \chi_R(z) dz$. Hence,

$$\text{for } N \text{ large and small } \varepsilon, \quad \frac{\Sigma^{N, \varepsilon}(s, s')}{N} = \frac{\varphi^N(s, s') - \varphi^N(s', s)}{N} \approx a \int_s^{s'} \chi_L(z) dz - b \int_s^{s'} \chi_R(z) dz.$$

This motivates the following definition.

Definition 4. Let a and b be the constants introduced in Assumption (P2). For each pair of points $x, y \in [0, 1], x \leq y$, the normalized relative mutation cost from x to y is

$$\sigma(x, y) = \int_x^y (a\chi_L(z) - b\chi_R(z))dz.$$

For $y \leq x$, the relative mutation cost from x to y is $\sigma(x, y) = -\sigma(y, x)$.

Note that if $(x, y) \subseteq X_+$ then $\sigma(x, y) = 0$. This holds if x and y are in the same limit recurrent set. Let $s = s^N(x)$ and $s' = s^N(y)$. From Lemma 1, in the limit, $\mu^{N,\varepsilon}(s) / \mu^{N,\varepsilon}(s') \approx \exp(N((- \log \varepsilon)\sigma(x, y) + V^\varepsilon(x, y)))$ is governed by the behavior of $\sigma(x, y)$. If $\sigma(x, y) > 0$, i.e., if the normalized relative mutation cost from x to y is strictly positive, then this expression converges to infinity as $(N, \varepsilon) \rightarrow (\infty, 0)$.

Definition 5. For each $x \in [0, 1]$, let $\omega(x) = \min_{y \in [0, 1]} \sigma(x, y)$. The gravity set $G_\sigma \subseteq [0, 1]$ associated to the limit process $Z^{\infty, 0}$ is $G_\sigma = \arg \max_{x \in [0, 1]} \omega(x)$.

The continuity of sigma, the principle of the maximum, and the compactness of $[0, 1]$ ensure that the gravity set is non-empty.

Lemma 3. Suppose the limit process $Z^{\infty, 0}$ has strong basins. Then:

(i) The maximal gravity set G_σ is a limit recurrent set or a union of limit recurrent sets.

Furthermore, for any $x \in G_\sigma$ we have that $\sigma(x, y) \geq 0$ with equality only if $y \in G_\sigma$.

(ii) Any limit recurrent set $E \subseteq G_\sigma$ contains an attractor of the mean flow equation (4).

Thus $M_V(E)$ consists of attractors of the mean flow equation in E .

Lemma 3(i) characterizes the maximal gravity set and (ii) says that the maximum viscosity cost points of this set are attractors. The proof is in the Appendix.

Let $G_\sigma = \cup_i E_i$, where the latter is a union of disjoint limit recurrent sets. With some abuse of notation we define $M_V(G_\sigma) = \cup_i M_V(E_i)$, i.e., the union of points that achieve maximum viscosity in each of the limit recurrent sets that compose G_σ , is referred as the set of maximum viscosity in G_σ . The set of points in $X = [0,1]$ that belong to state space S^N varies with N . To characterize the support of points in the limit distribution we use some additional notation: For any $H \subseteq [0,1]$ and $\eta > 0$ we denote by $M_V^\eta(H) = \left\{ x \in [0,1] / \inf_{y \in M_V(H)} |x - y| \leq \eta \right\}$ the set of points that are at most η away from some point in $M_V(H)$.

Theorem 2. Suppose the limit process $Z^{\infty,0}$ has strong basins. Let $\mu^{N,\varepsilon}$ be the unique invariant distribution of the birth and death process $Z^{N,\varepsilon}$, G_σ be the gravity set of the limit process $Z^{\infty,0}$, and $M_V(G_\sigma) \subseteq G_\sigma$ be the set of maximum viscosity in G_σ . Fix any $\eta > 0$. For any sequence $\{(N,\varepsilon)\} \subseteq \Theta$ with $(N,\varepsilon) \rightarrow (\infty,0)$ we have

$$\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon} \left(M_V^\eta(G_\sigma) \right) = 1.$$

Theorem 2 says that, if a process has strong basins, then as $(N,\varepsilon) \rightarrow (\infty,0)$ the mass of the invariant distribution is concentrated on those states that belong to the gravity set G_σ with maximum viscosity. The latter holds regardless of the specific trajectory in the parameter space Θ used to approach the limit. From Lemma 3(ii), any limit recurrent set $E_j \subseteq G_\sigma$ contains an attractor of the unperturbed mean flow and $M_V(E_j)$ consists of attractors of the mean flow equation (4).

Combined with Lemmas 2 and 3, Theorem 2 provides an explicit two-step procedure to identify $M_V(G_\sigma)$: First, identify the limit recurrent sets; then using Lemma 3, the maximal gravity set G_σ can be identified by computing $\sigma(x_i, x_j)$. Second, for each limit recurrent set E_j in the maximal set that is not a singleton, identify the

attractors A_{E_j} of the unperturbed mean flow $F = p - q$ restricted to E_j . By Lemma 3, if A_{E_j} is not a singleton, $M_V(E_j)$ is identified by computing $V(x_{jk}, x_{jk'})$ for any pair $x_{jk}, x_{jk'} \in A_{E_j}$.

The proof of Theorem 2 has two steps. We first show that $\lim_{(N, \varepsilon) \rightarrow (\infty, 0)} \mu^{N, \varepsilon}(G_\sigma) = 1$. This follows, the fact that $\mu^{N, \varepsilon}(s^N(x)) / \mu^{N, \varepsilon}(s^N(y))$ is roughly proportional to $\varepsilon^{N\sigma(x, y) + O(1)|x - y|}$ and, from Lemma 3(i), if $y \in G_\sigma$ then $\sigma(x, y) > 0$ unless $x \in G_\sigma$. If $\bar{\sigma} > 0$ then $N\varepsilon^{N\bar{\sigma}} \rightarrow 0$ for any sequence $(N, \varepsilon) \rightarrow (\infty, 0)$. It follows that the mass of the invariant distribution for states outside of G_σ converges to 0 as $(N, \varepsilon) \rightarrow (\infty, 0)$. From Lemma 3 we also know that the gravity set G_σ is a union of limit recurrent sets. Suppose that $G_\sigma = \bigcup_i E_i$. Now, if E_j is not a singleton, its interior is contained in X_+ and, as in Theorem 1, Assumptions (P1)-(P3) imply that if $y \in M_V(E_j)$ and $x \notin M_V(E_j)$, the ratio $\mu^{N, \varepsilon}(s^N(x)) / \mu^{N, \varepsilon}(s^N(y))$ converges to zero as $(N, \varepsilon) \rightarrow (\infty, 0)$.

Let δ_{x^*} be a point mass probability distribution on $x^* \in [0, 1]$. The following is an immediate Corollary of Theorem 2.

Corollary 1. *Suppose that G_σ consists of a single recurrent set and that $M_V(G_\sigma) = \{x^*\}$ is a singleton. There exist $\underline{\underline{N}}$ and $\bar{\bar{\varepsilon}}$ such that if $\underline{N} > \underline{\underline{N}}$ and $\bar{\varepsilon} < \bar{\bar{\varepsilon}}$ then any sequence of invariant distributions $\{\mu^{N, \varepsilon}\}_{(N, \varepsilon) \in \Theta}$ converges weakly to δ_{x^*} as $(N, \varepsilon) \rightarrow (\infty, 0)$.*

We conclude this section observing that Theorem 2 applies to sequences in Θ . This means that $\varepsilon > 0$ along any such sequence. The same techniques can be extended to study what happens for if we take $\lim_{\varepsilon \rightarrow 0}$ for some fixed N . If $G_\sigma \subseteq \{0, 1\}$ let $G_\sigma^N = G_\sigma$; otherwise let $G_\sigma^N = \left\{x \in S^N \mid \min_{y \in G_\sigma} |x - y| \leq 1/N\right\}$ be the set of states at most $1/N$ away from some point in G_σ .

Lemma 4. Suppose the limit process has strong basins. Let $\mu^{N,\varepsilon}$ be the unique invariant distribution of the birth and death process. There exist $\underline{N}$ and $\bar{\varepsilon}$ such that if $N > \underline{N}$ and $\varepsilon < \bar{\varepsilon}$ then $\lim_{\varepsilon \rightarrow 0} \mu^{N,\varepsilon}(G_\sigma^N) = 1$.

The proof is similar to that of Theorem 2 and is therefore omitted.

5.2 Applications

We revisit the examples introduced in section 2. In particular, we consider the case of a ν -best response dynamics for the evolutionary game introduced in 2.1 and the standardization model of 2.2.

5.2.1 KMR: Stochastic stability in 2×2 coordination games

We focus on the case in which Γ is a coordination game. In this case, the mixing probability for action A associated to equal payoffs for both actions ($\Delta u = 0$) is $x^* \equiv \beta - \gamma / (\alpha - \delta + \beta - \gamma) \in (0, 1)$. It is straightforward to check that $x_\nu^\pm = x^* \pm h_\Gamma \nu$ where $h_\Gamma = 1 / (\alpha - \delta + \beta - \gamma)$. Further, $R = [x_\nu^+, 1]$ and, for a fixed N any state in this set is the basin of attraction of the state $x = 1$ (all A). Similarly, $L = [0, x_\nu^-]$ contains the states in the basin of $x = 0$ (all B). The corners are also the only two limit-recurrent sets respectively denoted as E_A and E_B . As usual let $E_A^\eta = [1 - \eta, 1]$ and $E_B^\eta = [0, \eta]$.

Lemma 5. If $x^* < \frac{1}{1+a/b} + \frac{a-b}{a+b} h_\Gamma \nu$ then for any $\eta > 0$ $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(E_A^\eta) = 1$. If the opposite strict inequality holds then $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(E_B^\eta) = 1$.

Proof. From Lemma 3(i), to determine G_σ it suffices to compute

$\sigma(E_B, E_A) = \int_0^{x_\nu^-} a ds - \int_{x_\nu^+}^1 b ds = a x_\nu^- - b(1 - x_\nu^+)$. This number is positive if $a x_\nu^- - b(1 - x_\nu^+) > 0$. Replacing the values for x_ν^- and x_ν^+ the conclusion follows.

If $a = b$ the condition simplifies to $x^* < 1/2$ which is equivalent to saying A is risk dominant.

5.2.2 Standardization with conformity and vanity

For simplicity we take the case $a = b = 1$. In our standardization model, the limit process always has at least two limit recurrent points, the extremes corresponding to full standardization. In addition, if the “taste for vanity” parameter λ is sufficiently large, there exists an interval $E_{1/2}$ symmetric around the equal share point $x = 1/2$ that is also limit recurrent. The basins for each of these sets are illustrated in Figure 2.

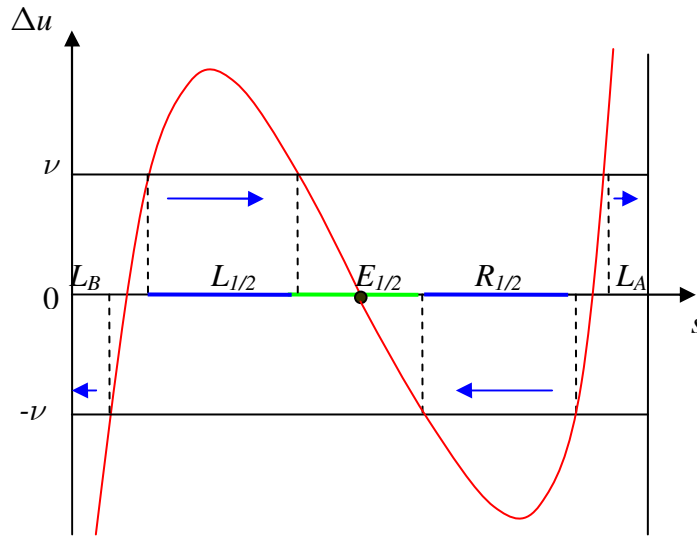


Figure 2

Intuitively, larger values of λ are associated with a stronger taste for vanity which dominates for intermediate states, so we expect the basin of $E_{1/2}$ to grow with λ .

Lemma 6. *For each ν there exists $\lambda(\nu)$ such that if $\lambda > \lambda(\nu)$ then $G_\sigma = E_{1/2}$ and, the equal share point $x = 1/2$ is stochastically stable in large populations. Instead, if $\lambda < \lambda(\nu)$ then $G_\sigma = \{0, 1\}$ and a full standardization state is stochastically stable in large populations.*

The maximizer and minimizer of $\Delta u(z)$ are monotonic in λ and so are the maximum and minimum values. As $\lambda \rightarrow \infty$ the maximizer converges to 0, the minimizer converges to 1, and the respective values diverge to ∞ and $-\infty$ respectively. Thus, for any ν there exists $\lambda(\nu)$ such that if $\lambda > \lambda(\nu)$ the length of $R_{1/2}$ ($L_{1/2}$) is strictly larger than the length of L_B (R_A), so $\sigma(E_{1/2}, 0) = \sigma(E_{1/2}, 1) > 0$ and the opposite holds if $\lambda < \lambda(\nu)$. It's easy to check that $x = 1/2$ is the only attractor of the mean flow on $E_{1/2}$. The result then follows from Corollary 1.

We observe that, in contrast to finite population models, stochastic stability in large populations selects a single point of $E_{1/2}$, an attractor of the mean flow equation in that interval.

6. Processes with Weak Basins

In this section we consider the case of limit processes with weak basins. In contrast to the case of strong basins, our main result is that selection will generally depend on the specific trajectory in the parameter space Θ used to approach the limit $(N, \varepsilon) \rightarrow (\infty, 0)$. We focus on the simplest case covering the applications discussed in this paper and, to our knowledge, the literature. In particular, we assume that $E \subseteq \{0, 1\}$, i.e., the recurrent sets of the limit process $Z^{\infty, 0}$ are the either one or both of extreme points 0 or 1. Further, we assume that if $0 \in E$ then $L = \{0\}$ and otherwise $L = \emptyset$. Similarly, if $1 \in E$ then $R = \{1\}$ and otherwise $R = \emptyset$. This means that the limit process has strictly positive transition probability maps in the interior of the $[0, 1]$ interval. For short, we say that the process is *positive in the interior*. This greatly simplifies the presentation and allows us to stress the main differences with the selection results of the previous sections. We discuss how the results extend for the more general case in Section 6.3.

6.1 Analysis

For a process with weak basins $\Sigma^N(x, y) = O(1)$. Further, our positive in the interior assumption implies that this quantity is independent of N , so we write

$\Sigma(x, y) \equiv \Sigma^N(x, y)$. Indeed, in this case, for any two interior points $x, y \in (0, 1)$ we have that $\Sigma(x, y) = 0$. Further, if $x \in E$ and $y \notin E$ then $\Sigma(x, y) > 0$, and $\Sigma(0, 1) = a - b$.

In the case of irreducible processes (Theorem 1) the support of the invariant distribution in the limit as $(N, \varepsilon) \rightarrow (0, \infty)$ is contained in the set of maximum viscosity points $M_V[0, 1]$, a subset of the mean flow attractors of the limit process ($x \in M_V[0, 1] \Leftrightarrow V(x, y) \geq 0$ for all y). In contrast, for strong basins the limiting distribution “selects” the maximal gravity set G_σ ($x \in G_\sigma \Leftrightarrow \sigma(x, y) \geq 0$ for all y). As seen shortly, processes with weak basins are a middle-ground case. The relative cost of mutations as measured by Σ matters because it determines the cost of reaching or leaving an absorbing point. However, since basins are weak and mutation costs are independent of N , the relative cost of transitions in “positive regions” as measured by the viscosity function V is also important. The characterization of the limiting distribution below illustrates how the balance between both of these transition costs determines which points are in the support for different trajectories in the (ε, N) space.

Definition 6. Given $x, y \in [0, 1]$, x is said to dominate y with respect to Σ and V if $\Sigma(x, y) \geq 0$ and $V(x, y) > 0$. The point x is said to be undominated with respect to Σ and V on $J \subseteq [0, 1]$ if there exists no $y \in J$ that dominates x .

Let $U_{\Sigma, V}(J) \subseteq J$ denote the set of points that are undominated with respect to Σ and V on $J \subseteq [0, 1]$. For $J = [0, 1]$ we use simply $U_{\Sigma, V}$ (i.e., $U_{\Sigma, V} \equiv U_{\Sigma, V}([0, 1])$). Note that absorbing points in E can only be dominated by other points in E (as $\Sigma(x, y) > 0$ if $x \in E$ and $y \notin E$). Since the process is positive in the interior and $\Sigma(x, y) = 0$ for any $x, y \notin E$, if $x \notin E$ and $x \notin M_V[0, 1]$ then x is dominated by all points in $M_V[0, 1]$. The next result follows:

Lemma 7. *If the limit process $Z^{\infty,0}$ is positive in the interior, $U_{\Sigma,V} \subseteq E \cup M_V[0,1]$. Furthermore, if $E \cap M_V[0,1] = \emptyset$ then $U_{\Sigma,V} = U_{\Sigma,V}(E) \cup M_V[0,1]$. Otherwise, if $E \cap M_V[0,1] \neq \emptyset$, then $U_{\Sigma,V} = U_{\Sigma,V}(E)$.*

The lemma provides a sharp characterization of the set of points that are undominated with respect to Σ and V . These are either absorbing points (that are themselves undominated by another absorbing point) or maximum viscosity points.¹⁶ Moreover, if there are absorbing points that are also of maximum viscosity, this set is simply the absorbing points undominated by another absorbing point. The next result shows that for any sequence $(N, \varepsilon) \rightarrow (0, \infty)$, in the limit, the support of the limiting invariant distribution is a subset of this set. Formally, for a fixed $\eta > 0$ let $U_{\Sigma,V}^\eta$ denote the set of points at most η away from some point in $U_{\Sigma,V}$.

Theorem 3. *If the limit process $Z^{\infty,0}$ is positive in the interior, then for any $\eta > 0$,*

$$\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(U_{\Sigma,V}^\eta) = 1.$$

The proof is in the Appendix, here is a brief intuition. Applying Lemma 1 to $s = s^N(x)$ and $s' = s^N(y)$, the limiting behavior of $\mu^{N,\varepsilon}(s') / \mu^{N,\varepsilon}(s) \approx e^{NV^\varepsilon(s',s) - (\log \varepsilon)\Sigma(s,s')}$. Thus, if x dominates y with respect to Σ and V ($\Sigma(x,y) \geq 0$ and $V(x,y) > 0$), regularity conditions (P1)-(P3) ensure that this ratio converges to zero for any sequence $(N, \varepsilon) \rightarrow (0, \infty)$. Hence, in the limit, only points contained in a small neighborhood of $U_{\Sigma,V}$ can be assigned strictly positive measure.

An immediate consequence of Theorem 3 and Lemma 7 is that as $(N, \varepsilon) \rightarrow (\infty, 0)$, the support of the invariant distribution is either a point in $M_V[0,1]$ (as is the case for irreducible processes) or an absorbing point in E (as is the case with strong basins). In particular, if either $E = \emptyset$ (irreducible limit process) or

¹⁶ By Lemma 2, if $M_V[0,1]$ is not a singleton and $x, y \in M_V[0,1]$ then $V(x, y) = 0$.

$U_{\Sigma,V}(E) \subseteq M_V[0,1]$ then for any sequence the support of the limiting distribution is contained in $M_V[0,1]$. This still leaves two unanswered questions: First, which sequences “select” points in E and which “select” points in $M_V[0,1]$ instead? Second, if $E = U_{\Sigma,V}(E) = \{0,1\}$ and we have a sequence that selects points in E , which of the two absorbing points is selected? Our next results answer these questions.

For any $x \in E$ and y such that $\Sigma(x, y) > 0$, define $\rho(x, y) = \left| \frac{V(x, y)}{\Sigma(x, y)} \right|$, the relative cost ratio. If $V(x, y) < 0$ this ratio is a measure of the disadvantage of a point $x \in E$ with respect to a point y in terms of the cost V relative to its advantage in terms of the mutation cost Σ . Note that for any two points $y, y' \in M_V[0,1]$ we have that $V(x, y) = V(x, y')$ for any x . We thus write $V(x, M_V)$ for this quantity. Note also that if $x \in U_{\Sigma,V}(E)$ and $x \notin M_V[0,1]$ then $\Sigma(x, y) > 0$ for each $y \in M_V[0,1]$.

The following theorem is the main result of this section. Let $m = \min\{a, b\}$ and $\bar{\rho} = \min\{\rho(x, y) \mid x \in U_{\Sigma,V}(E) \setminus M_V[0,1], y \in M_V[0,1]\}$.

Theorem 4. Suppose the limit process is positive in the interior, and Assumptions (P1)-(P3) hold. Let $\mu^{N,\varepsilon}$ be the unique invariant distribution of the birth and death process $Z^{N,\varepsilon}$. Assume that $U_{\Sigma,V}(E) \setminus M_V[0,1] \neq \emptyset$.

- (i) If $\varepsilon \leq e^{-WN}$ along the sequence with $W > \bar{\rho}$ then $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(E) = 1$.
- (ii) Let $\eta > 0$ and $\theta \leq O(\bar{\varepsilon}^m \mid \log \bar{\varepsilon} \mid)$. If $\varepsilon \geq e^{-WN}$ along the sequence with $W < \bar{\rho} - \theta$ then $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(M_V^\eta([0,1])) = 1$.

As is the case with our previous theorems, Theorem 4 in combination with our previous results provides an explicit procedure to determine which points are selected as $(N, \varepsilon) \rightarrow (\infty, 0)$. The main message of Theorem 4 is that in the case of weak basins selection is sensitive to the particular trajectory $(N, \varepsilon) \rightarrow (\infty, 0)$ considered. In particular,

reversing the order of the limits ($\lim_{\varepsilon \rightarrow 0} \lim_{N \rightarrow \infty}$ versus $\lim_{N \rightarrow \infty} \lim_{\varepsilon \rightarrow 0}$) yields different selection results if $M_V([0,1]) \not\subset E$. Our application below illustrates this point in the context of an evolutionary game. Further, Theorem 4 provides a tight condition to discriminate between sequences that converge to absorbing points and those that converge to a maximum viscosity attractor. Indeed, there is a threshold value $\bar{\rho}$ such that if the size of the perturbation ε goes to zero at a rate e^{-NW} or faster where W is above this threshold, the sequence selects absorbing points. Instead, if the perturbation goes to zero at rate slower than e^{-NW} where $W < \bar{\rho} - \theta$ then an attractor of maximum viscosity is selected, where $\theta \leq O(\bar{\varepsilon}^m |\log \bar{\varepsilon}|)$; since $\bar{\varepsilon}$ is uniform upper bound on ε , $\bar{\varepsilon} \ll 1$ implies that θ is near 0. In sum, ignoring cases near the threshold, the Theorem roughly says that the space of trajectories can be partitioned in two sets, with different selection results.

From Theorem 3 and Lemma 7, to establish theorem 4 it suffices to compare the probability mass that $\mu^{N,\varepsilon}$ assigns to points in E relative to the mass it assigns to points in a neighborhood of $M_V[0,1]$. To illustrate the result suppose that $E = \{0,1\}$ and $M_V([0,1]) = \{x_V\}$ where x_V is some point in the interior of $[0,1]$. Using Lemma 1, we have that the ratio $\mu^{N,\varepsilon}(0)/\mu^{N,\varepsilon}(s^N(x_V))$ is proportional to $e^{-NV^\varepsilon(0,x) - (\log \varepsilon)a}$. Our regularity conditions (Assumptions (P1)-(P3)) ensure that $V^\varepsilon(x,y) = V(x,y) + O(\varepsilon^m \log \varepsilon)$, where $m = \min\{a,b\}$. Note that $V(0,x_V) < 0$ by Lemma 2. Now, if $\varepsilon \leq e^{-NW}$ we have that

$$\frac{\mu^{N,\varepsilon}(0)}{\mu^{N,\varepsilon}(s^N(x_V))} \leq C e^{Na(W-V(0,x_V)/a + O(\varepsilon^m \log \varepsilon))} \leq C e^{Nhe^{-NW}} e^{Na(W-\rho(0,x_V))} \leq \tilde{C} e^{Na(W-\rho(0,x_V))}$$

for some $C, \tilde{C} > 0$. Similarly, in the limit, $\mu^{N,\varepsilon}(1)/\mu^{N,\varepsilon}(s^N(x_V)) \approx e^{-NV(1,x_V) - (\log \varepsilon)b}$. If $\varepsilon \leq e^{-NW}$ we have that

$$\frac{\mu^{N,\varepsilon}(1)}{\mu^{N,\varepsilon}(s^N(x_V))} \leq D e^{Nb(W-\rho(1,x_V))}$$

for some $D > 0$. Hence, if $W > \bar{\rho}$ we see that one of the above ratios converge to infinity as $N \rightarrow \infty$ and part (i) follows. Instead, if $W < \bar{\rho} - O(\varepsilon^m)$ the above ratios converge to zero as $N \rightarrow \infty$.

If E has a single element, the previous results are exhaustive with respect to identifying the points that are “selected” in the limit. We turn to the issue of identifying which particular absorbing point is selected if $E = \{0,1\}$. Let $y_{-x} \in E$ be the absorbing point other than x . Denote by $e_V = \arg \max_{x \in E} V(x, y_{-x})$ the absorbing point(s) with an “advantage” in terms of the relative cost of transitions as measured by V and let $e_\Sigma = \arg \max_{x \in E} \Sigma(x, y_{-x})$ be the absorbing point(s) with an advantage in terms of relative mutation costs. The following is an immediate corollary of Theorem 4.

Corollary 2. Suppose $E = \{0,1\}$ and consider a sequence such that $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(E) = 1$.

If $\Sigma(0,1) = 0$ ($a = b$) then $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(e_V) = 1$.

This corollary covers the case in which $U_{\Sigma,V}(E) = e_V$; its hypothesis that

$\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(E) = 1$ is satisfied for sequences satisfying the conditions of Theorem 4(i).

The complementary case is covered by the following result.

Corollary 3. Suppose $E = \{0,1\}$ and consider a sequence such that $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(E) = 1$.

Suppose in addition that $a \neq b$ and $V(0,1) \neq 0$ ($U_{\Sigma,V}(E) = E$).

(i) If either $\rho(e_\Sigma, M_V) < \rho(e_V, M_V)$ or $\bar{\rho} = \rho(e_V, M_V) < \rho(e_\Sigma, M_V)$ and $\varepsilon \leq e^{-NW}$ along the sequence with $W > \rho(e_V, e_\Sigma)$ then $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(e_\Sigma) = 1$; Otherwise,

(ii) If $\rho(e_V, M_V) < \rho(e_\Sigma, M_V)$ and $e^{-NW} \leq \varepsilon$ along the sequence with $W < \rho(e_V, e_\Sigma)$ then

$\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(e_\Sigma) = 1$.

The result is established by characterizing the limiting properties of the ratio

$\mu^{N,\varepsilon}(e_\Sigma)/\mu^{N,\varepsilon}(e_V)$; we omit the proof because it is similar to the proof of Theorem 2.

Corollary 3 says that if the mutation rate goes to zero sufficiently fast relative to the population growth (an exponential ratio higher than $\rho(e_V, e_\Sigma)$), $\mu^{N,\varepsilon}(e_\Sigma)/\mu^{N,\varepsilon}(e_V)$ goes to infinity and the absorbing point e_Σ with a mutation cost advantage is selected.

Conversely, if the mutation rate vanishes sufficiently slowly, this ratio converges to zero. However, this condition is only necessary to ensure that the absorbing point e_V . Indeed, if the mutation rate goes to zero sufficiently slow a point in $M_V[0,1]$ could be selected.

Thus, the exponential threshold $\rho(e_V, e_\Sigma)$ must be larger than the threshold

$\bar{\rho} = \min\{\rho(e_V, M_V), \rho(e_\Sigma, M_V)\}$ of Theorem 4, which holds if and only if

$\rho(e_V, M_V) < \rho(e_\Sigma, M_V)$.

6.2 Applications

We revisit the examples of section 2. For simplicity we fix $a = b = 1$.

6.2.1 Standardization

Consider the version of the standardization model that has logit probability maps. In this case the limit process has weak basins and the absorbing points are the corners corresponding to full standardization, i.e., $E = \{0,1\}$. Since the mean flow dynamics is payoff-monotonic, with weak basins, its attractors are the corners and the “equal share” point $x = 1/2$. To assess which state is selected as $(N, \varepsilon) \rightarrow (\infty, 0)$ we need to find $M_V[0,1]$. By symmetry, this set either contains both corners or none. In the latter case we have that $M_V[0,1] = \{1/2\}$. Observe that

$$V(0, 1/2) = \int_0^{1/2} \Delta u(z) dz = \int_0^{1/2} \log z dz - \int_0^{1/2} \log(1-z) dz - \lambda \int_0^{1/2} (2z-1) dz = 1 - \lambda.$$

Hence, if $\lambda < 1$ then $M_V[0,1] = E$ and, thus, $U_{\Sigma,V}(E) = E$ so $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(\{0,1\}) = 1$

(Theorem 3). For $\lambda > 1$, $M_V[0,1] = \{1/2\}$. In this case $\bar{\rho} = V(1/2, 0)/\Sigma(0, 1/2) = \lambda - 1$ and Theorem 2 gives the following:

Lemma 8. Suppose that $\lambda > 1$. Fix any small $\eta > 0$.

(i) If $\varepsilon \geq e^{-WN}$ along the sequence with $W < \lambda - 1$, $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon} \left(\left[\frac{1}{2} - \eta, \frac{1}{2} + \eta \right] \right) = 1$.

(Standards Coexistence)

(ii) Instead, if $\varepsilon \leq e^{-WN}$ along the sequence with $W > \lambda - 1 + \theta$

then $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(\{0,1\}) = 1$. (Full standardization)

6.2.2 Imitation with mutations and selection in 2×2 games

The mean flow of the limit process for the example in Section 2.3 is $F(x) = \frac{x(1-x)}{u(x)}(u_A(x) - u_B(x))$. Thus, the associated differential equation is payoff monotone.¹⁷ Regardless of the payoff function, the corners are steady states, as is typical in imitation processes. If the game has no dominant strategies, there is also an interior steady state at the point x_{AB} where both strategies yield the same payoff. The viscosity

function is $V(x, y) = \int_x^y \log \frac{u_A(z)}{u_B(z)} dz$. For a fixed game Γ , let

$$\varphi(\Gamma) \equiv \int_0^1 \log \frac{u_A(z)}{u_B(z)} dz = \int_0^1 \log(\gamma + (\alpha - \gamma)z) dz - \int_0^1 \log(\delta + (\beta - \delta)z) dz = \frac{\log(\alpha^\alpha / \gamma^\gamma)}{\alpha - \gamma} - \frac{\log(\beta^\beta / \delta^\delta)}{\beta - \delta}$$

Note that for a generic game $\varphi(\Gamma) \neq 0$, so $e_v = \{0\}$ if $\varphi(\Gamma) < 0$ and $e_v = \{1\}$ if $\varphi(\Gamma) > 0$.

Fudenberg et al [2006] study selection in the limit as $\varepsilon \rightarrow 0$ for a fixed N , and then let $N \rightarrow \infty$. From Corollary 2, it follows that their approach selects e_v . We compare these predictions with the ones obtained selection predictions for an arbitrary sequence $(N, \varepsilon) \rightarrow (\infty, 0)$, which depend on the game under consideration. A generic 2×2 symmetric game can be classified in terms of its set of symmetric Nash equilibria. In particular a game falls into one of the following three types:

¹⁷ Most of the imitation dynamics studied in evolutionary game theory are payoff monotonic. See Hofbauer and Sigmund [2003] for a discussion of these type of dynamics and, more generally, other evolutionary game dynamics.

Type I: Γ has a dominant strategy, thus a unique pure strategy equilibrium (e.g. Prisoner's Dilemma);

Type II: Γ has unique symmetric mixed strategy equilibrium (e.g. Hawk-Dove, Matching Pennies, anti-coordination);

Type III: Γ has three symmetric equilibria, two pure and one mixed (e.g., coordination games).

If the game is either of Type I or Type II the mean field dynamics has a unique attractor $\hat{x}$, which corresponds to the unique symmetric Nash equilibrium. Moreover, $M_V([0,1]) = \{\hat{x}\}$, i.e., the maximum viscosity set is exactly this point. In the case of a Type I game this unique Nash equilibrium is associated to the corner state in which the entire population plays the dominant strategy. Since this corner is a recurrent set, $M_V([0,1]) = e_V$. Instead, in the case of a Type II game, the unique mixed Nash equilibrium corresponds to the point $x_{AB} \in (0,1)$ where payoffs to each action are equal, i.e., $\hat{x} = x_{AB}$. Since $x_{AB} \notin \{0,1\}$, the selection prediction taking $N \rightarrow \infty$ first and $\varepsilon \rightarrow 0$ next differs from the other order of limits.

If the game is of Type III, the mean field dynamics has two asymptotically stable steady states, $x_A = 1$ and $x_B = 0$. Since the corners are the unique attractors of the dynamics, $M_V([0,1]) = e_V$ and (from Theorem 2 and Corollary 2) selection predictions are the same for any sequence $(N, \varepsilon) \rightarrow (\infty, 0)$. From the formula for $\varphi(\Gamma)$, it is clear that the selection criterion need not deliver the risk dominant equilibrium. In fact, the criterion depends both on risk and efficiency considerations. In particular, if action A is both risk and payoff dominant then $\varphi(\Gamma) > 0$ and $e_V = \{1\}$. The following result summarizes our previous discussion.

Lemma 9. *Let Γ be a generic 2x2 symmetric game and let $\varphi(\Gamma)$ be as defined above.*

(i) Suppose that Γ has pure Nash equilibria (Type I and Type III games). If $\varphi(\Gamma) < 0$ then $\lim_{(N, \varepsilon) \rightarrow (\infty, 0)} \mu^{N, \varepsilon}(\{0\}) = 1$. Otherwise, if $\varphi(\Gamma) > 0$ then $\lim_{(N, \varepsilon) \rightarrow (\infty, 0)} \mu^{N, \varepsilon}(\{1\}) = 1$.

(ii) Suppose that the unique symmetric equilibrium of Γ is mixed (Type II games). There exist $\underline{U}$, $\bar{U}$ with $\bar{U} \geq \underline{U} > 0$ such that: for any sequence $(N, \varepsilon) \rightarrow (\infty, 0)$ satisfying $\varepsilon \geq e^{-N\underline{U}}$ we have $\mu^{N, \varepsilon}(\{x_{AB}\}) \rightarrow 1$; and for any sequence $(N, \varepsilon) \rightarrow (\infty, 0)$ such that $\varepsilon < e^{-N\bar{U}}$, then $\mu^{N, \varepsilon}(\{0\}) \rightarrow 1$ is selected if $\varphi(\Gamma) < 0$ and $\mu^{N, \varepsilon}(\{1\}) \rightarrow 1$ if $\varphi(\Gamma) > 0$.

The characterization of the attractors of the mean field dynamics associated to the sequence processes in terms of the set of symmetric equilibria pursued above holds for any payoff monotonic mean field. It follows that for any birth and death process with an associated mean field of this class, a version of Lemma 9 holds. Of course, the map $\varphi(\Gamma)$ is specific to the process considered.

6.3 Extensions

6.3.1. General processes with weak basins

Two main issues arise once we relax the assumption that the limit process is positive in the interior. We illustrate these and how the results extend by means of a very simple example. Suppose that the limit process has a single absorbing point at $x = 1/2$, that is, $E = \{1/2\}$, so that $p(1/2) = q(1/2) = 0$. Assume that $L = R = \{1/2\}$ so that the limit process has transition maps that are positive in $[0, 1/2)$ and $(1/2, 1]$.

A first issue that arises is that, in contrast to the extremes of the interval 0 and 1, $x = 1/2$ is not always part of the state space S^N as $1/2 \notin S^N$ if N is odd.¹⁸ This has subtle implications in characterizing the limit of interest: Consider a sequence $(N, \varepsilon) \rightarrow (\infty, 0)$ where N takes odd values, so that $|s^N(1/2) - 1/2| = 1/2N$. If $Q^\varepsilon(\bullet)$ has a derivative bounded away from zero at $x = 1/2$ for all ε then

$$Q^\varepsilon(s^N(1/2)) \geq Q^\varepsilon(1/2) + h(s^N(1/2) - 1/2) \text{ for some } h > 0, \text{ so } Q^\varepsilon(s^N(1/2)) \geq h/2N$$

independently of ε . Thus, regardless of how fast we send $\varepsilon \rightarrow 0$ the transition probabilities on the points of the grid S^N closest to the absorbing point are bounded below by a term of order $1/N$. Hence if $x = 1/2$ is not a point of maximum viscosity in

¹⁸ Note that had we chosen x to be an irrational number, x would not be in the state space for any N .

$[0, 1/2]$ and x_v is any point in this interval such that $V(1/2, x_v) < 0$, Lemma 1 implies

that $\frac{\mu^{N,\varepsilon}(s^N(1/2))}{\mu^{N,\varepsilon}(s^N(x_v))} \leq k e^{NV^\varepsilon(1/2, x_v) + h \log N} = C N^h e^{NV^\varepsilon(1/2, x_v)}$ for some $C > 0$, which converges

to zero as $N \rightarrow \infty$ independently of ε . In sum, for these “odd N ” subsequences, no matter how fast we send the mutation rate to zero relative to the population size, the support of the invariant distribution does not contain the absorbing point $x = 1/2$. Thus, Theorem 4 (ii) no longer holds. Of course, for subsequences where N is even, we have that $s^N(1/2) = 1/2$ and $Q^\varepsilon(1/2) = O(\varepsilon^b)$ so the conclusion of Theorem 4 (ii) remains unchanged.

In contrast, it is easy to show that the conclusion of Theorem 4(i) can be generalized. In particular, there exists a threshold value $\bar{W}$ such that for any trajectory satisfying $\varepsilon \geq e^{-NW}$ with $W < \bar{W}$ we have that in the limit $\mu^{N,\varepsilon}$ is concentrated in a subset of $M_v([0, 1])$. Our previous example (N even) also shows that, as before the support of a limiting distribution may also contain absorbing points in E . However, in contrast to the “positive in the interior” case, it is also possible for a point that is not in E nor in $M_v([0, 1])$ to be in the support of a limiting distribution. Specifically, consider the subinterval $[0, 1/2]$. The limit transition probability maps in the interior of this interval are strictly positive, and there are no mutation costs for transitions within this interval. Hence, the exact same arguments used to analyze the limiting invariant probability ratios for processes with positive interior in $[0, 1]$ can be applied to this subinterval. In particular, in this subinterval, the only candidates to be in the support of the limiting distribution are points contained in $M_v[0, 1/2]$ the set of maximum viscosity points in the subinterval: the same argument used before shows that the ratio $\mu^{N,\varepsilon}(s^N(x)) / \mu^{N,\varepsilon}(M_v^\eta[0, 1/2])$ converges to zero if $x \notin M_v^\eta[0, 1/2]$. The same reasoning for the $[1/2, 1]$ subinterval, gives points in $M_v[1/2, 1]$ as potential candidates too. In general, it’s easy to check that either $M_v[0, 1] = M_v[0, 1/2]$, $M_v[0, 1] = M_v[1/2, 1]$ or $M_v[0, 1] = M_v[0, 1/2] \cup M_v[1/2, 1]$. It is not difficult to construct examples such that even if $M_v[0, 1] = M_v[1/2, 1]$ there are sequences such that

$\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(M_V^\eta[0,1/2]) = 1$. This requires a positive relative mutation cost from points in $M_V[0,1/2]$ to points in $M_V[1/2,1]$.

7. Conclusion

This paper provides a characterization of the long-run stability properties of noisy evolutionary processes for a general version of the double limit as the population size N becomes infinitely large and the noise ε vanishingly small. We thus extend the analyses of Benaïm and Weibull [2003], who focus on the iterated limit where N goes to infinity before the noise term goes to zero, and Fudenberg et al [2006] and Fudenberg and Imhof [2007], who study the limit where the noise term goes to zero before the population size grows.

We identify the processes for which trajectories to the limit in the (N,ε) space matter and provide an explicit characterization of the limit distribution. If the process has *strong basins* all trajectories yield the same answer and the support of the invariant distribution can be characterized by a two-step procedure: First, identify the basin with the “largest basin” –defined in a suitable way. Second, if this set is not a singleton, identify the subset of attractors of the mean flow dynamics of the process restricted to this set. In processes with *weak basins* the trajectory matters. We show that if the noise term decreases faster than a suitable rate exponential in the size of population, the support is included in a recurrent set of the noiseless process, the same answer we obtain by considering the iterated limit in which the noise goes to zero first and the population grows. If the opposite holds, the support is a subset of attractors just as in iterated limit in which the population grows to infinity first and the noise vanishes next.

Extending the conclusions for multi-dimensional birth-death processes remains a subject of future research. Our results also suggest that models in which the mutation rate is endogenously determined by the population size may also provide valuable insight on the nature of selection.

APPENDIX A: Proofs of Lemmas and Theorems

Lemmas A1 and A2 below are used to establish our main results.

Lemma A1. Suppose that Assumptions (P1) and (P2) hold. For each $\varepsilon \in (0, \bar{\varepsilon}]$ and each $x \in [0, 1]$ let $a^\varepsilon(x) = \frac{\log P^\varepsilon(x)}{\log(\varepsilon)}$ and $b^\varepsilon(x) = \frac{\log Q^\varepsilon(x)}{\log(\varepsilon)}$. Then:

(i) There exists C such that $\|a^\varepsilon\| \leq C$ and $\|b^\varepsilon\| \leq C$.

(ii) For any sequence $\varepsilon \rightarrow 0$, $\{a^\varepsilon\}$ converges pointwise to $\hat{a} = \chi_L a$.

Further, convergence is uniform on any closed interval $J \subset [0, 1]$ that does not contain a point in the boundary of L , and $a^\varepsilon(x) = \hat{a}(x) + O(|\log \varepsilon|^{-1})$ for all $x \in J$.

(iii) For any sequence $\varepsilon \rightarrow 0$, $\{b^\varepsilon\}$ converges pointwise to $\hat{b} = \chi_R b$. Further, convergence is uniform on any closed interval $K \subset [0, 1]$ that does not contain a point in the boundary of R and $b^\varepsilon(x) = \hat{b}(x) + O(|\log \varepsilon|^{-1})$ for all $x \in K$.

Proof. (i) From Assumption (P2), $|a^\varepsilon(x)| \leq \left| \frac{\log(k\varepsilon^a)}{\log(\varepsilon)} \right| \leq a + \left| \frac{\log(k)}{\log(\varepsilon)} \right|$. Similarly,

$|b^\varepsilon| \leq b + \left| \frac{\log(h)}{\log(\varepsilon)} \right|$. Hence, $C = \max\{a, b\} + \left| \frac{\log(h)}{\log(\varepsilon)} \right|$ satisfies the statement.

(ii) If $x \in L$ then Assumption (P2) implies immediately that $P^\varepsilon(x) = O(\varepsilon^a)$ from which $\{a^\varepsilon\}$ converges uniformly to $\chi_L a$ on any subset of L . If $x \notin L$ then $P^\varepsilon(x)$ is bounded below for any $\varepsilon \in (0, \bar{\varepsilon}]$ and $\{a^\varepsilon(x)\}$ converges to zero as $\varepsilon \rightarrow 0$, so $a^\varepsilon \rightarrow 0$ pointwise on $[0, 1] \setminus L$. For a closed interval $J \subset [0, 1] \setminus L$, define $\alpha_J = \sup_{x \in J} \sup_{\varepsilon \in [0, \bar{\varepsilon}]} |\log P^\varepsilon(x)| > 0$ and

$|a^\varepsilon(x)| \leq \frac{\alpha_J}{|\log \varepsilon|}$ for all $x \in J$. The right hand converges to zero as $\varepsilon \rightarrow 0$ (independently of x).

(iii) Analogous to (ii). $\square$

Lemma A2. Let $V, V^\varepsilon : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ be defined by (1) and (2) in the main text.

(i) V is uniformly bounded on $[0, 1] \times [0, 1]$, and so is V^ε (independently of ε);

(ii) $V^\varepsilon \rightarrow V$ uniformly on $[0, 1] \times [0, 1]$ as $\varepsilon \rightarrow 0$. Further, $\|V^\varepsilon - V\| = O(\varepsilon^m \log \varepsilon)$ where $m = \min\{a, b\}$;

(iii) If $z \in X_0$ is in the boundary of X_+ and $\{z^N\}$ is a sequence that converges to z with $z^N \in X_+$ then $V^\varepsilon(z, z^N) = O(|z^N - z| \log |z^N - z|)$;

(iv) If $g^{N, \varepsilon}$ and $h^{N, \varepsilon}$ are two maps that converge uniformly to the identity as

$(N, \varepsilon) \rightarrow (0, \infty)$, and $W^{N, \varepsilon}(x, y) = V^\varepsilon(g^{N, \varepsilon}(x), h^{N, \varepsilon}(y))$ then $W^{N, \varepsilon} \rightarrow V$ uniformly.

Proof. The set X_+ can be expressed as a finite disjoint union $\cup_j I_j$ where each $I_j = (\underline{x}_j, \bar{x}_j) \subseteq X_+$ and either $\underline{x}_j, \bar{x}_j \in X_0$. Thus, $V(x, y) = \sum_j -\int_x^y \chi_{I_j}(z) \log \frac{p(z)}{q(z)} dz$ and it suffices to establish the Lemma for V and V^ε restricted to $[\underline{x}_j, \bar{x}_j] \times [\underline{x}_j, \bar{x}_j]$ for each j . We omit the subindex j in what follows.

(i) Suppose first that $\underline{x} \in L$ ($p(\underline{x}) = 0$). Since $(\underline{x}, \bar{x}) \in X_+$, there exists $\delta > 0$ such that for any $z \in J \equiv [\underline{x}, \underline{x} + \delta]$ we have that $p'(z) > 0$. Let $H(\pi) = \pi \log \pi - \pi$ and observe that $\frac{\partial H}{\partial \pi}(\pi) = \log \pi$. Note also that $\lim_{\pi \rightarrow 0} H(\pi) = 0$ and that $H(\pi)$ is uniformly bounded on $[0, 1]$. Using this notation we have that

$$\int_{\underline{x}}^y \log p(z) dz = \frac{1}{l_p(y)} \int_{\underline{x}}^y \frac{\partial H}{\partial p}(p(z)) p'(z) dz = \frac{H(p(y)) - \lim_{z \rightarrow \underline{x}} H(p(z))}{l_p(y)} = \frac{H(p(y))}{l_p(y)} \quad (\text{A2.1})$$

where $l_p(y) = \int_{\underline{x}}^y p'(z) \log p(z) dz / \int_{\underline{x}}^y \log p(z) dz$. Since $l_p(y)$ is strictly positive for each

$y \in J$ (and thus bounded away from zero), then $\left| \int_{\underline{x}}^y \log p(z) dz \right|$ is bounded for all $y \in J$.

Similarly, if $\underline{x} \in R$ we have that for $J' \equiv [\underline{x}, \underline{x} + \delta']$ and $y \in J'$, we can define $l_q(y)$

analogously and conclude that $\left| \int_{\underline{x}}^y \log q(z) dz \right| = |H(q(y)) / l_q(y)|$ is bounded as well.

Finally, if V is uniformly bounded, then (ii) implies V^ε is as well and that this bound is independent of ε .

(ii) We are interested in estimating $\|V^\varepsilon - V\|$ as $\varepsilon \rightarrow 0$. Let $J = [x, y] \subseteq [\underline{x}, \bar{x}]$. Suppose first that $x, y \in X_+$. Let $\underline{p} > 0$ and $\underline{q} > 0$ be lower bounds for p and q on J . Note that,

$$|V^\varepsilon(x, y) - V(x, y)| = \left| \int_x^y \log \frac{P^\varepsilon(z)}{p(z)} - \log \frac{Q^\varepsilon(z)}{q(z)} dz \right| \leq \left| \int_x^y \log \frac{P^\varepsilon(z)}{p(z)} dz \right| + \left| \int_x^y \log \frac{Q^\varepsilon(z)}{q(z)} dz \right|. \quad (\text{A2.2})$$

By Assumption (P2), for any $z \in J$ we have that $P^\varepsilon(z) / p(z) \leq 1 + k\varepsilon^a / \underline{p}$ and

$Q^\varepsilon(z) / q(z) \leq 1 + h\varepsilon^b / \underline{q}$. Thus, if $m = \min\{a, b\}$ and observing that $\log(1+z) \leq z$,

$$|V^\varepsilon(x, y) - V(x, y)| \leq |x - y| \left(\frac{k\varepsilon^a}{\underline{p}} + \frac{h\varepsilon^b}{\underline{q}} \right) \leq \frac{2}{\min\{\underline{p}/k, \underline{q}/h\}} \varepsilon^m.$$

In sum, if $x, y \in X_+$ then $|V^\varepsilon(x, y) - V(x, y)| = O(\varepsilon^m)$.

Suppose instead that $x = \underline{x}$, i.e., $x \in X_0$. For concreteness, suppose that $x = \underline{x} \in L$ so that $p(\underline{x}) = 0$ (the case $x \in R$ is analogous). We will show that $|V^\varepsilon(\underline{x}, r) - V(\underline{x}, r)| = O(\varepsilon^m |\log \varepsilon|)$.

Let $J = [\underline{x}, \underline{x} + \delta]$ as in part (i) and note that δ can always be chosen so that $r \equiv \underline{x} + \delta < y$ (thus $r \in X_0$), and that $p'(y) > 0$ is bounded away from zero for each $y \in J$. Let $\bar{P}^\varepsilon(z) = p(z) + \Delta^\varepsilon$ where $\Delta^\varepsilon = \|P^\varepsilon - p\|$. Using (A2.2) and the fact that Q^ε and q are strictly positive on J we have that

$$|V^\varepsilon(\underline{x}, r) - V(\underline{x}, r)| \leq \left| \int_{\underline{x}}^r \log \frac{P^\varepsilon(z)}{p(z)} dz \right| + O(\varepsilon^m) \leq \left| \int_{\underline{x}}^r \log \frac{\bar{P}^\varepsilon(z)}{p(z)} dz \right| + O(\varepsilon^m). \quad (\text{A2.3})$$

Observe that

$$\int_{\underline{x}}^r \log \frac{\bar{P}^\varepsilon(z)}{p(z)} dz = \int_{\underline{x}}^r \frac{\partial}{\partial \pi} (H(p(z) + \Delta^\varepsilon) - H(p(z))) \frac{p'(z)}{p'(z)} dz = \frac{H(p(r) + \Delta^\varepsilon) - H(p(r)) - H(\Delta^\varepsilon)}{p'(z_r)} \quad (\text{A2.4})$$

where $z_r \in [\underline{x}, r]$ (from the intermediate value theorem and $\lim_{z \rightarrow \underline{x}} H(p(z)) = 0$). Note also

that $p'(z_r)$ is positive and bounded away from zero,

$$|H(p(r) + \Delta^\varepsilon) - H(p(r))| = \Delta^\varepsilon |\log(p(r))| + O(\Delta^{\varepsilon^2}),$$

and $|\log(p(r))| > 0$ ($r \in X_+$), and $H(\Delta^\varepsilon) = \Delta^\varepsilon \log \Delta^\varepsilon - \Delta^\varepsilon$. By assumption (P2),

$$\Delta^\varepsilon = O(\varepsilon^m), \text{ and combining (A2.3) and (A2.4) shows } |V^\varepsilon(r, \underline{x}) - V(r, \underline{x})| \leq O(\varepsilon^m \log \varepsilon).$$

Similarly, if $y = \bar{x}$ then for some $r' < \bar{x}$ with $r' \in X_+$ then

$$|V^\varepsilon(\bar{x}, r) - V(\bar{x}, r)| \leq O(\varepsilon^m \log \varepsilon).$$

Finally, observe that $V(x, y) = V(x, r) + V(r, r') + V(r', y)$ and similarly for V^ε . From

$$\text{above, } |V^\varepsilon(x, r) - V(x, r)| \text{ and } |V^\varepsilon(r', y) - V(r', y)| \text{ are } O(\varepsilon^m \log \varepsilon), \text{ and}$$

$$|V^\varepsilon(r, r') - V(r, r')| = O(\varepsilon^m). \text{ We conclude that, for any } x, y \in [0, 1],$$

$$|V^\varepsilon(x, y) - V(x, y)| \leq O(\varepsilon^m \log \varepsilon) \text{ as desired.}$$

(iii) The argument is very similar to the second part of (ii) observing that, if $x \in L$ then $P^\varepsilon(z^N) \approx P^\varepsilon(\underline{x}) + k^\varepsilon |z^N - \underline{x}|$ where $k^\varepsilon = P^{\varepsilon'}(\underline{x}) > 0$ is bounded away from zero for all ε . (If $x \in R$ then $Q^\varepsilon(z^N) \approx Q^\varepsilon(\underline{x}) + k^\varepsilon |z^N - \underline{x}|$.) The details are omitted.

$$(iv) |W^{N,\varepsilon}(x, y) - V^\varepsilon(x, y)| \leq C |g^{N,\varepsilon}(x) - x| + C |h^{N,\varepsilon}(y) - y| \leq C (\|g^{N,\varepsilon} - I\| + \|h^{N,\varepsilon} - I\|),$$

where I is the identity. Hence, $\|W^{N,\varepsilon} - V\| \leq C (\|g^{N,\varepsilon} - I\| + \|h^{N,\varepsilon} - I\|) + \|V^\varepsilon - V\|$ and

$$\|V^\varepsilon - V\| \rightarrow 0 \text{ from the uniform convergence of } V^\varepsilon.$$

□

Lemma 1. The invariant distribution $\mu^{N,\varepsilon}$ of $Z^{N,\varepsilon}$ satisfies

$$\frac{\mu^{N,\varepsilon}(s')}{\mu^{N,\varepsilon}(s)} = \exp(N(V^\varepsilon(s', s) + r^N(s, s')) - \log(\varepsilon) \Sigma^N(s', s) + O(1) |s' - s|)$$

for all $s, s' \in S^N$ with $s \leq s'$, where $r^N(s, s') = 0$ if $(s, s') \subseteq X_+$ and $r^N(s, s') = O\left(\frac{\log N}{N}\right)$ otherwise.

Proof of Lemma 1.

Let $s, s' \in S^N$ with $s' \geq s$. The invariant distribution $\mu^{N, \varepsilon}$ satisfies the detailed balance condition $P^\varepsilon(s) \mu^{N, \varepsilon}(s) = Q^\varepsilon(s + \frac{1}{N}) \mu^{N, \varepsilon}(s + \frac{1}{N})$ for $s \leq (N-1)/N$. Letting

$K(s, s') = N|s' - s|$, $\Delta_N = 1/N$, and $s_k^N = s + k\Delta_N$, we have

$$\frac{\mu^{N, \varepsilon}(s')}{\mu^{N, \varepsilon}(s)} = \prod_{k=1}^{K(s, s')} \frac{P^\varepsilon(s + (k-1)\Delta_N)}{Q^\varepsilon(s + k\Delta_N)} = \prod_{k=1}^{K(s, s')} \frac{P^\varepsilon(s_{k-1}^N)}{Q^\varepsilon(s_k^N)}.$$

It follows that

$$\log \frac{\mu^{N, \varepsilon}(s')}{\mu^{N, \varepsilon}(s)} = \sum_{k=0}^{K(s, s')-1} \log P^\varepsilon(s_k^N) - \sum_{k=1}^{K(s, s')} \log Q^\varepsilon(s_k^N) = \sum_{z \in S^N: s \leq z < s'} \log P^\varepsilon(z) - \sum_{z \in S^N: s < z \leq s'} \log Q^\varepsilon(z).$$

Recall that $X_+ = X \setminus (L \cup R)$ and $X_0 = L \cup R$. We decompose the above sum into two terms H_+ and H_0 so that $\log \frac{\mu^{N, \varepsilon}(s')}{\mu^{N, \varepsilon}(s)} = H_+ + H_0$ where

$$H_+ = \sum_{z \in S^N: s \leq z < s'} \chi_{X_+}(z) \log P^\varepsilon(z) - \sum_{z \in S^N: s < z \leq s'} \chi_{X_+}(z) \log Q^\varepsilon(z) \text{ and}$$

$$H_0 = \sum_{z \in S^N: s \leq z < s'} \chi_{X_0}(z) \log P^\varepsilon(z) - \sum_{z \in S^N: s < z \leq s'} \chi_{X_0}(z) \log Q^\varepsilon(z).$$

We first estimate H_0 and show that $H_0 = (\log \varepsilon) \Sigma^N(s, s') + O(1) |s' - s|$. Next, we show that $H_+ = NV^\varepsilon(s, s') + C_1 |s' - s| + C_2 \log ND_{(s, s')}$, where C_1 is a constant and $D_{(s, s')} = 0$ if $(s, s') \subseteq X_0$ and, otherwise D counts the number of discontinuities of χ_{X_+} on (s, s') ; combining these two facts will yield the result. Towards this end, note that

$$H_0 = \log \varepsilon \left(\sum_{z \in S^N: s \leq z < s'} \chi_{X_0}(z) a^\varepsilon(z) - \sum_{z \in S^N: s < z \leq s'} \chi_{X_0}(z) b^\varepsilon(z) \right).$$

By Lemma A1, $a^\varepsilon(z) = a + O(|\log \varepsilon|^{-1})$ if $z \in L$ and $a^\varepsilon(z) = 0 + O(|\log \varepsilon|^{-1})$ if $z \in R \setminus L$. Similarly, $b^\varepsilon(z) = b + O(|\log \varepsilon|^{-1})$ if $z \in R$ and $b^\varepsilon(z) = 0 + O(|\log \varepsilon|^{-1})$ if $z \in L \setminus R$. Hence,

$$H_0 = \log \varepsilon \left(\sum_{z \in S^N: s \leq z < s'} \chi_L(z) a + \chi_{R \setminus L}(z) O(|\log \varepsilon|^{-1}) - \sum_{z \in S^N: s < z \leq s'} \chi_R(z) b + \chi_{L \setminus R}(z) O(|\log \varepsilon|^{-1}) \right)$$

$$= \log \varepsilon \left(a \sum_{z \in S^N: s \leq z < s'} \chi_L(z) - b \sum_{z \in S^N: s < z \leq s'} \chi_R(z) \right) + O(1) |s' - s| = (\log \varepsilon) \Sigma^N(s, s') + O(1) |s' - s|.$$

Next, observe that if $s \in X_0$ then $\chi_{X_+}(s) = 0$ and, otherwise, if $s \in X_+$ then $Q^\varepsilon(s)$ is uniformly bounded away from zero for all ε . Thus, $\log Q^\varepsilon(s) = O(1)$. Similarly, if $s' \in X_0$ then $\chi_{X_+}(s') = 0$ and, otherwise, if $s' \in X_+$ then $\log P^\varepsilon(s') = O(1)$. Hence,

$$H_+ = \sum_{z \in S^N: s \leq z \leq s'} \chi_{X_+}(z) \log \frac{P^\varepsilon(z)}{Q^\varepsilon(z)} + O(1) |s - s'| = N \left(\sum_{k=0}^{K(s,s')} \chi_{X_+}(s) \log \frac{P^\varepsilon(z)}{Q^\varepsilon(z)} \Delta_N \right) + O(1) |s - s'|.$$

Observe that the term in parenthesis is a Riemann sum for the function $h^\varepsilon(z) = \chi_{X_+}(z) \log \frac{P^\varepsilon(z)}{Q^\varepsilon(z)}$. If $(s, s') \subseteq X_0$, P^ε and Q^ε are continuous and strictly

positive in (s, s') and $h^\varepsilon(z)$ is continuous and bounded in this interval (thus Riemann-integrable). We then have that

$$\left(\sum_{k=0}^{K(s,s')} \chi_{X_+}(s) \log \frac{P^\varepsilon(z)}{Q^\varepsilon(z)} \Delta_N \right) = \int_s^{s'} \log \frac{P^\varepsilon(z)}{Q^\varepsilon(z)} dz + O(\Delta_N) |s - s'| = V^\varepsilon(s', s) + O(\Delta_N) |s - s'|.$$

Suppose instead that (s, s') contains points in X_0 . Since L and R are a finite union of intervals, the set of points in $(s, s') \cap X_+$ (containing all the points where $\chi_{X_+} \neq 0$ in the above sum) can be expressed as a finite union of intervals $I_j = [x_j^N, y_j^N] \subseteq X_+$ such that (i) $x_j^N, y_j^N \in S^N$, (ii) either x_j^N is the smallest state in X_+ closest to the point $\underline{z}_j \in X_0$ or y_j^N is the largest state in X_+ closest to the point $\bar{z}_j \in X_0$ ($\bar{z}_j < y_j^N$ in the closure of or X_+), and (iii) we have that

$$\sum_{k=0}^{K(s,s')} \chi_{X_+}(s) \log \frac{P^\varepsilon(z)}{Q^\varepsilon(z)} \Delta_N = - \sum_j V^\varepsilon(x_j^N, y_j^N) + O(\Delta_N) |s - s'|.$$

Observe that $O(|x_j^N - \underline{z}_j|) = O(|y_j^N - \bar{z}_j|) = O(1/N)$. Hence, from Lemma A2(iii) we have that $V^\varepsilon(x_j^N, y_j^N) = V^\varepsilon(\underline{z}_j, \bar{z}_j) + O(1/N) \log(1/N)$. Combining this with the previous we get

$$\begin{aligned} H_+ &= N \left(- \sum_j (V^\varepsilon(\underline{z}_j, \bar{z}_j) + O(1/N) \log(1/N)) + O(\Delta_N) |s - s'| \right) + O(1) |s - s'| \\ &= N \left(-V^\varepsilon(s, s') + O(1/N) \log(1/N) \right) + O(\Delta_N) |s - s'| + O(1) |s - s'|. \end{aligned}$$

□

Proof of Lemma 2. The result follows from Lemma B2 in Appendix B. □

We use D_I to denote the number of discontinuities of χ_{X_+} on an interval $I \subset [0, 1]$.

Lemma A3. Suppose that the limit process has strong basins and let

$$\sigma^N(x, x') \equiv \frac{\Sigma^N(x, x')}{N}. \text{ Then } \sigma^N(x, x') = \sigma(x, x') + O(1/N)(\gamma D_{[x, x']} + |x' - x|) \text{ for some}$$

$\kappa \geq 0$ (thus, σ^N converges uniformly to σ on $[0, 1] \times [0, 1]$).

Proof. Let $\Delta_N = 1/N$ and $f(z) = \chi_L(z)a - \chi_R(z)b$, so that $\sigma(x, y) = \int_x^y f(z)dz$. Note that both L and R are a finite union of intervals and, thus, f is Riemann-integrable. Further, the discontinuities of f coincide with those of χ_{X_+} if L and R do not overlap. Thus, $\left| \sigma(x, x') - \sum_{z \in S^N: x \leq z \leq x'} f(z)\Delta_N \right| \leq \gamma D_{[x, x']} \Delta_N$ for some $\gamma \geq 0$ (any upper bound of the oscillation of f on intervals of length Δ_N times the number of discontinuities). Since $\Sigma^N(x, x') = \sum_{z \in S^N: x \leq z \leq x'} f(z) + O(1)|x' - x|$, the result follows. $\square$

Lemma 3. Suppose the limit process $Z^{\infty, 0}$ has strong basins. Then:

- (i) The maximal gravity set G_σ is a recurrent set or a union of recurrent sets of the limit process $Z^{\infty, 0}$. Furthermore, for any $x \in G_\sigma$ we have that $\sigma(x, y) \geq 0$ with equality only if $y \in G_\sigma$.
- (ii) A recurrent set $E \subseteq G_\sigma$ contains an attractor of the mean flow equation (4). Thus $M_V(E)$ consists of attractors of the mean flow equation in E .

Proof of Lemma 3.

We start by noting that $x \in G_\sigma \Leftrightarrow \sigma(x, y) \geq 0$ for all $y \in [0, 1]$. This follows from the fact that for any $z, z', y \in [0, 1]$, $\sigma(z, z') = \sigma(z, y) + \sigma(y, z')$, from which $\omega(z) = \min_{z' \in [0, 1]} \sigma(z, z')$ satisfies $\omega(z) = \sigma(z, y) + \omega(y)$. Hence, $\omega(x) \geq \omega(y)$ for all $y \Leftrightarrow \sigma(x, y) \geq 0$ for all y .

If a limit recurrent set E is not a singleton then $\text{int}(E) \subset X_+$. Since the integrand in σ is zero in X_+ we just need to show that for some limit recurrent set E and $x \in E$ we have that $\sigma(x, y) \geq 0$. We next show that if y is not in a recurrent set this inequality is strict.

As in Lemma A3, let $\sigma^N(x, x') = \Sigma^N(x, x')/N$. Now, for a fixed N and any $s, s' \in S^N$, $\Sigma^N(s, s')$ is the resistance or mutation cost from state s to state s' minus the cost from s to s' for the process $Z^{N, 0}$ on S^N . For each $s \in S^N$, define $w^N(s) = \min_{s' \in S^N} \sigma^N(s, s')$ and let $G^N = \arg \max_{s \in S^N} w^N(s)$. The σ^N is proportional to Σ^N and states in G^N are thus stochastically stable for the Markov process $Z^{N, 0}$. It is well known that G^N must be a recurrent set or a union of recurrent sets of this process. Since the set of transition maps of $Z^{N, 0}$ are p and q for any N , any recurrent set $E^N \subseteq S^N$ of $Z^{N, 0}$ can be written as $E^N = E \cap S^N$, where E is a limit recurrent set. In particular, any element of G^N is contained in a limit recurrent set. Note also that, the exact same argument above implies that if $x^N \in G^N$ then $\sigma^N(x^N, s) \geq 0$ for all $s \in S^N$.

Let $\{x^N\}$ be a converging sequence $x^N \rightarrow x$ such that $x^N \in G^N$. Let E be a limit recurrent set or a union of limit recurrent sets such that $G^N \subseteq E$. Thus, $x^N \in E$ and,

since a limit recurrent set is closed, then the limit lies in E , i.e. $x \in E$. We claim that $x \in G_\sigma$. Indeed, let $x' \neq x$ be an arbitrary point. We know that $\sigma^N(x^N, s^N(x')) \geq 0$. We claim that $\lim_N \sigma^N(x^N, s^N(x')) = \sigma(x, x')$. Indeed,

$$|\sigma^N(x^N, s^N(x')) - \sigma(x, x')| \leq |\sigma^N(x^N, s^N(x')) - \sigma(x^N, s^N(x'))| + |\sigma(x^N, s^N(x')) - \sigma(x, x')|$$

The right hand side converges to zero as $\sigma^N \rightarrow \sigma$ uniformly on $[0,1] \times [0,1]$ (Lemma A3) and σ is absolutely continuous. Thus, we have found $x \in E$ such that $\sigma(x, y) \geq 0$ for all y .

To complete the proof, let y be a point which is not contained in a recurrent set. We have two cases. Suppose first that $y \in X_0 \setminus \cup_j E_j$. Assume that $y \in L$ (the case $y \in R$ is analogous). This means that y is in the basin of attraction of a recurrent set $\tilde{E}$, “to the left”. Since $y \in L$, there exists an interval between y and E such that $\chi_L(x) = 1$ and $\chi_R(x) = 0$ for x in that interval. Thus, $\sigma(y, z) < 0$ for $z \in \tilde{E}$. Hence, since $\sigma(x, z) \geq 0$, $\sigma(x, y) = \sigma(x, z) + \sigma(z, y) > 0$. Suppose instead that $y \in X_+$. This means that y is contained in an interval $Y_+ \subseteq X_+$ which is adjacent to a subset of X_0 -call it Y_0 . It is immediate that there exists a $z \in Y_0$ such that $\sigma(y, z) < 0$. Again, $\sigma(x, y) = \sigma(x, z) + \sigma(z, y)$. If z belongs to a limit recurrent set and $\sigma(x, z) \geq 0$, as shown earlier. Otherwise, $z \in X_0 \setminus \cup_j E_j$ and we just showed that $\sigma(x, z) > 0$ in that case. It follows that $\sigma(x, y) > 0$. The proof is complete. $\square$

(ii) If $E \subseteq G_\sigma$ is a positive length interval the conclusion follows from Lemma B1 (Appendix B). Suppose that $E = \{x\}$ is a singleton. If x is an extreme, say $x = 0$, the fact E is limit recurrent implies that $0 \in L$. With strong basins, this means there exists a positive length interval to the right of x contained in L , from which it immediately follows that x is an attractor of the mean flow. Suppose instead that $x \in (0,1)$. This means that $x \in L \cap R$. Further, x must be the unique point at which some interval $L' \subseteq L$ and another interval $R' \subseteq R$ intersect as, by assumption $L \cap R$ contains isolated points. We argue that L' is to the right of x and R' is to its left, which means that x is an attractor. Indeed, if L' were to the left of x it would mean that there exists a limit recurrent set E' to right of x such that $\sigma(E', E) \geq \int_{L'} \sigma ds > 0$. From (i), this contradicts

$$E \subseteq G_\sigma.$$

$\square$

The following Lemma, used to prove Theorem 2, requires some extra notation. Let $G_\sigma = \cup_j E_j$ where each E_j is a recurrent set. We write $E_j = [\underline{x}_j, \bar{x}_j]$ with the convention $\underline{x}_j = \bar{x}_j$ if E_j is a point and larger values of j associated to “higher” recurrent sets. Note that $\underline{x}_j \in R \cup \{0\}$ while $\bar{x}_j \in L \cup \{1\}$. Let $s_-^N(x) = \max\{s \in S^N : s \leq x\}$ and

$s_+^N(x) = \max\{s \in S^N : s \geq x\}$, as in the proof of Theorem 2. Finally, let $E_j^N = [s_-^N(\underline{x}_j), s_+^N(\bar{x}_j)]$ and $G_\sigma^N = \cup_j E_j^N$.

Lemma A4. Let S_j^- respectively S_j^+ be the set of states smaller than $s_-^N(\underline{x}_j)$ respectively larger than $s_+^N(\bar{x}_j)$ which are closer to E_j than to any other recurrent set in G_σ and let $\sigma^N(x, x') \equiv \frac{\Sigma^N(x, x')}{N}$. There exist $\gamma > 0$, $\underline{N}$, and $\bar{\varepsilon} > 0$ such that if $N > \underline{N}$ and $\varepsilon < \bar{\varepsilon}$ then

$$\sigma^N(x_j^N, s) > \gamma |x^N - s|,$$

where $x_j^N = s_-^N(\underline{x}_j)$ and $s \in S_j^-$ or $x_j^N = s_+^N(\bar{x}_j)$ and $s \in S_j^+$.

Proof. We use the same notation as in the proof of Lemma A3 and, for ease of notation, omit the subindex j in the statement. Consider the case $x^N = s_-^N(\underline{x})$ and $s \in S^-$. The case $x^N = s_+^N(\bar{x})$ and $s \in S^+$ is analogous. If $\underline{x} = s_-^N(\underline{x}) = 0$ there is no $y < s_-^N(\underline{x})$ so the statement is vacuously true. Suppose that $\underline{x} > 0$. This means that $\underline{x} \in R$.

From Lemma A3,

$$\sigma^N(s_-^N(\underline{x}), s) = \sigma(s_-^N(\underline{x}), s) + O(\Delta_N)(|s_-^N(\underline{x}) - s| + \gamma D_{[s, s_-^N(\underline{x})]}). \quad (\text{A4})$$

Now, $\sigma(\underline{x}, s) > 0$ as $\underline{x} \in G_\sigma$ (Lemma 3(i)). Fix any $\eta_0 > 0$. If $|\underline{x} - s| \geq \eta_0$, since σ is uniformly continuous on $[0, 1] \times [0, 1]$ there exists $\underline{N}_1$ independent of s such that $\sigma(s_-^N(\underline{x}), s) \geq \beta(\eta_0) > 0$ for all $N > \underline{N}_1$ and the result follows immediately. Hence, we must show that there exists $\eta_0 > 0$ such that the result holds for $|\underline{x} - s| \leq \eta_0$.

Indeed, since $\underline{x} \in R$ the strong basins assumption implies that there exists a positive length interval $R = [z, \underline{x}]$ to the left of $\underline{x}$ contained in R . Let $\eta_0 = |z - \underline{x}|/2$. Observe

that for $\underline{N} > \frac{1}{\eta_0}$ and $N > \underline{N}$, $s_-^N(\underline{x}) \in R$ and there exists an interval of positive length

$R^N = [z, s_-^N(\underline{x})]$ to the left of $s_-^N(\underline{x})$ contained in R . It follows that $D_{[s, s_-^N(\underline{x})]} = 0$ for any

$s \in R^N$. Using this, from (A4) it suffices to show that, for a suitable choice of $\bar{N}$,

$h(y) = \frac{\sigma(s_-^N(\underline{x}), y)}{|s_-^N(\underline{x}) - y|}$ is positive and bounded away from zero on R^N . Indeed, this follows

immediately from the fact that $\sigma(s_-^N(\underline{x}), y)$ is differentiable with respect to y on R^N and

$$\frac{\partial \sigma(s_-^N(\underline{x}), y)}{\partial y} = b > 0 \quad \square$$

Proof of Theorem 2.

The proof has two steps. Step 1 shows that $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(G_\sigma) = 1$. Step 2 establishes that $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(M_V(G_\sigma)) = 1$.

Step 1: We show that there is a uniform lower bound on N and a uniform upper bound on ε such that for any sequence with elements in Θ satisfying $(N,\varepsilon) \rightarrow (\infty,0)$ the sequence of invariant distributions $\{\mu^{N,\varepsilon}\}_{N,\varepsilon}$ is such that $\mu^{N,\varepsilon}(J) \rightarrow 0$ for any closed set J that does not intersect G_σ . Let $\Gamma^{N,\varepsilon}$ be the cumulative distribution function corresponding to the invariant measure $\mu^{N,\varepsilon}$.

From Lemma 3(i), G_σ is a union of recurrent sets of the limit process. Assume that it consists of a single recurrent set. (The argument for the case in which G_σ is a union of two or more recurrent sets is sketched below.) Let $\underline{x} = \min\{x \in G_\sigma\}$ and $\bar{x} = \max\{x \in G_\sigma\}$. Suppose that $\underline{x} > 0$; the case $\underline{x} = 0$ is covered by the second part of the argument.

Let $G_\sigma^N = [s_-^N(\underline{x}), s_+^N(\bar{x})]$. Lemma 1 applied to $s' = s_+^N(\bar{x})$ and $s < s'$ implies

$$\mu^{N,\varepsilon}(s) \leq \mu^{N,\varepsilon}(G_\sigma^N) \exp\left(-N \left((-\log \varepsilon) \frac{\Sigma^N(s_-^N(\underline{x}), s)}{N} + V^\varepsilon(s_-^N(\underline{x}), s) + O(1/N) |s_-^N(\underline{x}) - s| \right)\right).$$

By Lemma A2(i), V^ε is uniformly bounded, and Lemma A4 implies $\Sigma^N(s_-^N(\underline{x}), s)/N \geq \gamma |s_-^N(\underline{x}) - s|$ for some $\gamma > 0$ and $N > \underline{N}$. Hence,

$$(-\log \varepsilon) \frac{\Sigma^N(s_-^N(\underline{x}), s)}{N} + V^\varepsilon(s_-^N(\underline{x}), s) + O(1/N) |s_-^N(\underline{x}) - s| > [(-\log \varepsilon)\gamma + O(1)] |s_-^N(\underline{x}) - s|$$

for some $c > 0$. We conclude that for some $\bar{\varepsilon} < \bar{\varepsilon}$ we have that $\mu^{N,\varepsilon}(s) \leq e^{-N(-\log \varepsilon)c}$ for some $c > 0$. We conclude that for any $y < s_-^N(\underline{x})$

$$\Gamma^{N,\varepsilon}(y) = \sum_{s \in S^N, s < s_-^N(\underline{x})} \mu^{N,\varepsilon}(s) \leq W N e^{-N(-\log \varepsilon)c}.$$

The right side converges to zero for any sequence $(N,\varepsilon) \rightarrow (\infty,0)$. Further, the above inequality is still valid replacing y with y^N where the latter is any sequence satisfying $y^N < s_-^N(\underline{x})$. Thus, $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \Gamma^{N,\varepsilon}(y) = 0$ for any $y < \underline{x}$. Let $H^{N,\varepsilon}(y) = 1 - \Gamma^{N,\varepsilon}(y)$. Using the same argument above it can be shown that $H^{N,\varepsilon}(y) \rightarrow 0$ as $(N,\varepsilon) \rightarrow (\infty,0)$ for any $y > \bar{x}$.

Now we sketch the case in which $G_\sigma = \bigcup_{j=1}^m E_j = [\underline{x}_j, \bar{x}_j]$ contains $m > 1$ recurrent sets.

Without loss of generality, assume $\bar{x}_j < \underline{x}_{j+1}$. Define $\Gamma^{N,\varepsilon}$ as before. In this case, we need to show that in the limit as $(N,\varepsilon) \rightarrow (\infty,0)$ the distribution has zero mass on intervals $[x, y]$ such that $\bar{x}_j < x < y < \underline{x}_{j+1}$ where $\bar{x}_0 = 0$ and $\underline{x}_{m+1} = 1$. The argument above implies that $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \Gamma^{N,\varepsilon}(y) = 0$ for any $y < \underline{x}_1$ and also that $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \Gamma^{N,\varepsilon}(y) = 1$ for any

$y > \bar{x}_m$. Let $y \in (\bar{x}_j, \bar{x}_{j+1})$ for some $j \in \{1, \dots, m-1\}$ (i.e., y is a point between two recurrent sets in G_σ) and let $H_j^{N,\varepsilon}(y) = \Gamma^{N,\varepsilon}(\bar{x}_{j+1}) - \Gamma^{N,\varepsilon}(y)$. Repeating the above argument shows that $\lim_{N \rightarrow \infty} H_j^{N,\varepsilon}(y) = 0$, which implies that $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(G_\sigma) = 1$. This concludes Step 1.

Step 2: Suppose that G_σ has more than one element ($\bar{x} \neq \bar{x}$), otherwise the statement is trivial. Any sequence $(N, \varepsilon) \rightarrow (\infty, 0)$ can be associated to sequences of the form $\{x^\varepsilon\}$ where each x^ε belongs to $M_{V^\varepsilon}(G_\sigma)$. Observe that $M_{V^\varepsilon}(G_\sigma) \subseteq A_{G_\sigma}^\varepsilon$ (Lemma B2(ii)). Consider any such sequence and suppose it converges to some point x . By Lemma B4(i) this limit is contained in $M_V(G_\sigma) \subseteq A_{G_\sigma}$. We extend the argument for the more general case in which the sequence need not converge later. Let $x_-^\varepsilon = \inf_{\varepsilon' \leq \varepsilon} x^{\varepsilon'}$, $x_+^\varepsilon = \sup_{N' \geq N, \varepsilon' \leq \varepsilon} x^{\varepsilon'}$, and $I^\varepsilon = [x_-^\varepsilon, x_+^\varepsilon]$.

Suppose that $x > \bar{x}$ (for $x = \bar{x}$, see the second part of the argument in Step 1). Consider any sequence $\{y^{N,\varepsilon}\} \subseteq G_\sigma \cap S^N$ such that $y^{N,\varepsilon} \rightarrow y$ as $(N, \varepsilon) \rightarrow (\infty, 0)$, $y < x$, and satisfying $y^{N,\varepsilon} < s_-^N(x_-^\varepsilon)$. Lemma B4(ii) implies that $\alpha_y = \sup_{N,\varepsilon} V^\varepsilon(s_-^N(x_-^\varepsilon), y^{N,\varepsilon}) > 0$.

Now, since transition probabilities are strictly positive independently of N and ε in the interior of G_σ (which is limit recurrent), from Lemma 1 we have that

$$\mu^{N,\varepsilon}(y^{N,\varepsilon}) \leq W \mu^{N,\varepsilon}(I^\varepsilon) e^{-NV^\varepsilon(s_-^N(x_-^\varepsilon), y^{N,\varepsilon})} \leq W e^{-N\alpha_y}$$

where W is a strictly positive constant. Therefore, $\Gamma^{N,\varepsilon}(y^{N,\varepsilon}) \leq W N e^{-N\alpha_y}$.

The left hand side is independent of ε and converges to zero for any sequence $(N, \varepsilon) \rightarrow (\infty, 0)$. It follows that $\Gamma^{N,\varepsilon}$ converges to a distribution which is equal to zero for any $y < x$. Again, mimicking the strategy in Step 1 and using the same argument above, one can show that $H^{N,\varepsilon}(y) = 1 - \Gamma^{N,\varepsilon}(y)$ converges to zero as $(N, \varepsilon) \rightarrow (\infty, 0)$ for any $y > x$. We conclude that $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(M_V(G_\sigma)) = 1$.

To extend the argument for a more general sequence $\{x^\varepsilon\}$ with $x^\varepsilon \in M_{V^\varepsilon}(G_\sigma)$ we invoke Lemma B3(iii). Indeed, since $M_{V^\varepsilon}(G_\sigma) \subseteq A_{G_\sigma}^\varepsilon$ there exists a set $M \subseteq M_V(G_\sigma)$ such that the sequence can be decomposed into converging subsequences $x^{\varepsilon_a} \rightarrow x_a$ for each $x_a \in M$. To each of these subsequences we associate the corresponding subsequences of invariant measures $\{\mu^{N_a, \varepsilon_a}\}$ and distribution functions $\{\Gamma^{N_a, \varepsilon_a}\}$. Applying the above argument to each of these converging subsequences we conclude that $\lim_{(N_a, \varepsilon_a) \rightarrow (\infty, 0)} \mu^{N_a, \varepsilon_a}(M_V(G_\sigma)) = 1$.

The proof is complete. $\square$

Theorem 3. Fix any $\eta > 0$. Then $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(U_{\Sigma,V}^\eta) = 1$.

Proof of Theorem 3.

Let $y \notin U_{\Sigma, V}$. By Lemma 7, since $E \subseteq \{0, 1\}$, there are two cases: (I) $y \notin E$ is an interior point of $[0, 1]$ which is not in $M_V[0, 1]$ or (II) $y \in E$ and there exists another $x \in E$ such that $\Sigma(x, y) \geq 0$ and $V(x, y) > 0$. The latter case requires $E = \{0, 1\}$, so if the process is irreducible or E is a singleton only (I) is possible.

We start with case (I). Let $\Gamma^{N, \varepsilon}$ be the distribution function associated to the invariant measure $\mu^{N, \varepsilon}$. For any sequence $(N, \varepsilon) \rightarrow (\infty, 0)$, let $\{x^\varepsilon\}$ be any sequence where each x^ε belongs to $M_{V^\varepsilon}([0, 1])$. Suppose that this sequence converges to some x . By Lemma B4(i), $x \in M_V([0, 1])$. Note that by Lemma B2(i) each $x^\varepsilon \in M_{V^\varepsilon}([0, 1])$ in the sequence is an attractor of the of the perturbed mean flow, i.e., $x^\varepsilon \in A_{[0, 1]}^\varepsilon$. Thus, the argument can be extended for a more general sequence of x^ε 's (i.e., a sequence that does not necessarily converge) invoking Lemma B3(iii) as in the last part of the proof of Theorem 2, we omit the details.

Mimicking the proof of Theorem 2 (see Step 2) and using the same notation introduced there, suppose that $x > 0$. Fix $\eta > 0$. Consider a sequence $\{y^{N, \varepsilon}\} \subseteq S^N$ such that

$\lim_{(N, \varepsilon) \rightarrow (\infty, 0)} y^{N, \varepsilon} = y < x - \eta$ and satisfying $y^{N, \varepsilon} < s_-^N(x^\varepsilon) - \eta$. Note that each $y^{N, \varepsilon}$ is in the interior of $[0, 1]$ for all (N, ε) and in the limit.

From Lemma 1, it follows that

$$\mu^{N, \varepsilon}(y^{N, \varepsilon}) \leq C \mu^{N, \varepsilon}(I^{N, \varepsilon}) e^{-NV^\varepsilon(s_-^N(x^\varepsilon), y^{N, \varepsilon}) + (\log \varepsilon) \Sigma^N(x^\varepsilon, y^{N, \varepsilon})} \quad (*)$$

for some $C > 0$. With weak basins $\Sigma^N(x^\varepsilon, y^{N, \varepsilon}) = \Sigma(s_-^N(x^\varepsilon), y^{N, \varepsilon})$. Further, since $y^{N, \varepsilon}$ is in the interior of $[0, 1]$ and the limit process is positive in the interior, $\Sigma(s_-^N(x^\varepsilon), y^{N, \varepsilon}) \geq 0$. On the other hand, Lemma B4(ii) implies $V_y = \sup_{N, \varepsilon} V^{N, \varepsilon}(s_-^N(x^\varepsilon), y^{N, \varepsilon}) > 0$. Combining the previous with the inequality (*) we get

$$\mu^{N, \varepsilon}(s^N(y)) \leq C e^{-NV_\eta}.$$

Letting $V_\eta^- = \min_{x \leq y \leq x - \eta} V_y$. It follows that

$$\Gamma^{N, \varepsilon}(y) - \Gamma^{N, \varepsilon}(0) \leq C N e^{-NV_\eta^-}. \quad (**)$$

Note that the left hand side is independent of ε and converges to zero for any sequence $(N, \varepsilon) \rightarrow (\infty, 0)$. Hence, $\Gamma^{N, \varepsilon}(y) - \Gamma^{N, \varepsilon}(0) \rightarrow 0$ for any $y \leq x - \eta$.

For $y \geq x + \eta$, set $H^{N, \varepsilon}(y) = 1 - \Gamma^{N, \varepsilon}(y)$ as in the proof of Theorem 2. Defining $V_\eta^+ = \min_{x + \eta \leq y \leq x} V_y$ it can be shown that an inequality analogue to (**) holds for $y \geq x + \eta$ replacing V_η^- with V_η^+ . Using the same argument above it can be shown that $H^{N, \varepsilon}(y) - H^{N, \varepsilon}(1) \rightarrow 0$ as $(N, \varepsilon) \rightarrow (\infty, 0)$ for any $y \geq x + \eta$. We conclude that in the

limit the support of the invariant distribution does not contain points in the interior of $[0,1]$ that are outside of $M_V^\eta[0,1]$.

Consider now case (II). The only relevant case is when $E = \{0,1\}$ and either 0 dominates 1 or vice versa. Without loss of generality, assume that 1 dominates 0 (the other case is symmetric). This means that $\Sigma(1,0) \geq 0$ and $V(1,0) > 0$. Observing that $s^N(0) = 0$ and $s^N(1) = 1$, applying Lemma 1 gives

$$\mu^{N,\varepsilon}(0) \leq C \mu^{N,\varepsilon}(1) e^{-NV^\varepsilon(1,0) + (\log \varepsilon)\Sigma(1,0)} \leq C e^{-NV^\varepsilon(1,0)}.$$

By Lemma A2, $V^\varepsilon(1,0) = V(1,0) + O(r(\varepsilon))$ where $\lim_{\varepsilon \rightarrow 0} r(\varepsilon) = 0$. Since $V(1,0) > 0$, for a suitable choice of $\bar{\varepsilon}$ we have that $V^\varepsilon(1,0)$ is bounded away from zero for all ε . It follows that $e^{-NV^\varepsilon(1,0)}$ converges to zero as $N \rightarrow \infty$ independently of ε . This completes the proof. $\square$

Theorem 4. Suppose the limit process has weak basins, is positive in the interior, and Assumptions (P1)-(P3) hold. Let $\mu^{N,\varepsilon}$ be the unique invariant distribution of the birth and death process $Z^{N,\varepsilon}$. Assume that $U_{\Sigma,V}(E) \setminus M_V[0,1] \neq \emptyset$.

- (i) If $\varepsilon \leq e^{-WN}$ along the sequence with $W > \bar{\rho}$ then $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(E) = 1$.
- (ii) Let $\eta > 0$ and $\theta > O(\bar{\varepsilon}^m |\log \bar{\varepsilon}|)$. If $\varepsilon \geq e^{-WN}$ along the sequence with $W < \bar{\rho} - \theta$ then $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(M_V^\eta([0,1])) = 1$.

Proof of Theorem 4.

Throughout the proof we assume that $U_{\Sigma,V}(E) \not\subset M_V[0,1]$. Otherwise the Theorem is trivial given Theorem 3 and Lemma 7.

(i) Let $x \in U_{\Sigma,V}(E)$. This means that $\Sigma(x,y) > 0$ for each $y \in M_V[0,1]$. We show that for any sequence such that $\varepsilon \leq e^{-WN}$ along the sequence with $W > \rho(x, M_V)$ then $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(E) = 1$. Without loss of generality, suppose that $x = 0$ so that $x < y$ for any $y \in M_V[0,1]$ (the case in which $x = 1$ is symmetric). Observe that $s^N(x) = x$ and, since the process has weak basins, $\Sigma^N(x,y) = \Sigma(x,y) > 0$. Applying Lemma 1 yields

$$\mu^{N,\varepsilon}(s^N(y)) \leq C \mu^{N,\varepsilon}(x) e^{NV^\varepsilon(y,x) + (\log \varepsilon)\Sigma(x,y)} \leq C e^{NV^\varepsilon(y,x) + (\log \varepsilon)\Sigma(x,y)}.$$

Now, by Lemma A2(iii), $V^\varepsilon(y,x) = V(y,x) + O(\varepsilon^m \log \varepsilon) |s^N(y) - s^N(x)|$ where $m = \min\{a,b\} > 0$. Hence, $V^\varepsilon(y,x) \geq V(y,x) + h\varepsilon^m |\log \varepsilon|$ for some constant h . If $\varepsilon \leq e^{-WN}$ we have that $N\varepsilon^m |\log \varepsilon| \leq N^2 W e^{-NWm}$ which converges to 0 as $N \rightarrow \infty$. Combing this with the previous we get

$$\mu^{N,\varepsilon}(s^N(y)) \leq \bar{C} e^{-N(\Sigma(x,y)W - V(y,x))} = \bar{C} e^{-N\Sigma(x,y)(W - \rho(x, M_V))}$$

where $\bar{C}$ is any upper bound to $C e^{-hN^2 W e^{-mN}}$. It follows that the probability mass over all states in a small neighborhood of any point in $M_V[0,1]$ is bounded by

$\bar{C}Ne^{-N\Sigma(x,y)(W-\rho(x,M_V))}$. For $W > \rho(x, M_V)$, this expression converges to 0 as $N \rightarrow \infty$. By Theorem 3 and Lemma 7, since $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(M_V^\eta[0,1]) = 0$, we must have that $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(E) = 0$.

(ii) Let $x \in M_V[0,1]$ and $y \in U_{\Sigma,V}(E) \setminus M_V[0,1]$. Recall that $V(x,y) > 0$ and $\Sigma(y,x) \geq 0$. We also have that $s^N(y) = y$ as $E \subseteq \{0,1\}$. Applying Lemma 1 we have that

$$\mu^{N,\varepsilon}(y) \leq C\mu^{N,\varepsilon}(s^N(x))e^{-NV^\varepsilon(x,y) - (\log \varepsilon)\Sigma(y,x)}.$$

From Lemma A2(iii), we have that $V^\varepsilon(y,x) \geq V(y,x) - O(\varepsilon^m |\log \varepsilon|)$, with $m = \min\{a,b\} > 0$. Hence, if $\varepsilon \geq e^{-WN}$ we obtain

$$\mu^{N,\varepsilon}(y) \leq Ce^{-N(V(x,y) - O(\varepsilon^m |\log \varepsilon|) - \Sigma(x,y)W)} = Ce^{-NV(x,y) \left(1 - O(\varepsilon^m |\log \varepsilon|) - \frac{W}{\rho(x,y)}\right)}.$$

If, as in the hypothesis, we have that $0 < 1 - \theta/\bar{\rho} - W/\bar{\rho}$ then $\mu^{N,\varepsilon}(y) \leq Ce^{-NV(x,y)K}$ for some $K > 0$, which converges to zero as $N \rightarrow \infty$. It follows that $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(E) = 0$ and, from Theorem 3 and Lemma 7, we conclude that $\lim_{(N,\varepsilon) \rightarrow (\infty,0)} \mu^{N,\varepsilon}(M_V^\eta([0,1])) = 1$ for any $\eta > 0$.

□

APPENDIX B. Mean Flow Dynamics and Attractors

Let $\hat{P}$ and $\hat{Q}$ be transition maps satisfying assumptions (P1)-(P3) and let $\hat{F} = \hat{P} - \hat{Q}$ the associated mean flow. For a closed interval $I = [\underline{x}, \bar{x}]$ we denote by $\hat{A}_I$ the set of attractors of the mean field differential equation $\dot{x} = \hat{F}(x)$ restricted to values of $x \in I$. For $I = [0,1]$ the set is denoted by $\hat{A}$. For a general interval I , if an endpoint $\underline{x}$ or $\bar{x}$ belongs to $\hat{A}_I$ then these points need not be in $\hat{A}$.

Lemma B1: $\hat{A}_I$ is non-empty for any interval I , and it contains those attractors of $\hat{A}$ that lie in I and possibly one or both extremes of I . If I is a limit recurrent set with positive length ($\underline{x} \neq \bar{x}$) then $\hat{A}_I \subseteq \hat{A}$ and $\hat{P}$ and $\hat{Q}$ are bounded away from zero for any point in $\hat{A}_I$.

Proof. Consider the following two conditions: (i) $\hat{P}(\underline{x}) \leq \hat{Q}(\underline{x})$ and $\hat{P}(x) < \hat{Q}(x)$ for $x \in I$ in a neighborhood of $\underline{x}$; (ii) $\hat{P}(\bar{x}) \geq \hat{Q}(\bar{x})$ and $\hat{P}(x) > \hat{Q}(x)$ for $x \in I$ in a neighborhood of $\bar{x}$. If (i) is satisfied $\underline{x}$ is an attractor of the mean field dynamics restricted to I and if (ii) is satisfied then $\bar{x}$ is. Suppose that neither of these properties is satisfied. This means that $\hat{P}(\underline{x}) > \hat{Q}(\underline{x})$ and $\hat{P}(\bar{x}) < \hat{Q}(\bar{x})$. By the continuity of $\hat{P}$ and $\hat{Q}$ there exists a point x^* in the interior of $E = [\underline{x}, \bar{x}]$ such that $\hat{P}(x^*) = \hat{Q}(x^*)$ and for some $\delta > 0$, $\hat{P}(x) \geq \hat{Q}(x)$ if

$x \in (x^* - \delta, x^*)$ and $\hat{P}(x) \leq \hat{Q}(x)$ if $x \in (x^*, x^* + \delta)$. The inequality is strict for at least some points in these δ -neighborhoods. This means that $x^* \in \hat{A}$.

Suppose that I is a limit recurrent set and assume that (i) is satisfied so that $\underline{x} \in \hat{A}_I$. We claim that if (i) is satisfied then $\underline{x} = 0$. Note that this implies $\underline{x} \in \hat{A}$. Assume not, i.e., $\underline{x} \neq 0$. Since the process cannot move to the left at $\underline{x}$ and since $\underline{x} \neq 0$ it must be that $\hat{Q}(\underline{x}) = 0$. On the other hand, since I is a recurrent set, we must have that $\hat{P}(\underline{x}) > 0 = \hat{Q}(\underline{x})$, and (i) is violated. This proves the claim. Similarly, if (ii) is satisfied so that $\bar{x} \in \hat{A}_I$ it must be that $\bar{x} = 1$ and, thus, $\bar{x} \in \hat{A}$. $\square$

Let $I \subseteq [0, 1]$ be a closed interval such that transition maps $\hat{P}$ and $\hat{Q}$ are strictly positive and continuous in I . For any pair of points $x, y \in I$ define the cost function $\hat{V}(x, y) = -\int_x^y \log \frac{\hat{P}(s)}{\hat{Q}(s)} ds$ associated to $\hat{P}$ and $\hat{Q}$. Letting $\phi_I(x) = \min_{y \in I} \hat{V}(x, y)$, the maximum viscosity set associated to $\hat{V}$ in I is defined as $M_{\hat{V}}(I) = \arg \max_{x \in I} \phi(x)$.

Lemma B2: Let $M_{\hat{V}}(I)$ be the maximum viscosity set associated to $\hat{V}$ in I and let $\hat{A}_I$ be the set of attractors of the mean flow differential equation associated to the flow $F = \hat{P} - \hat{Q}$. Then:

(i) $M_{\hat{V}}(I) \subseteq \hat{A}_I$.

(ii) For $m \in M_{\hat{V}}(I)$ and $y \in I$ we have that $\hat{V}(m, y) \geq 0$, with equality only if $y \in M_{\hat{V}}(I)$.

Proof. (i) We start with some notation. For any $x \in I$ and $y \in I$, let $\phi_x(y) = \hat{V}(x, y)$ and

$Z(x) = \arg \min_{y \in I} \phi_x(y)$. For any $y \in \text{int}(I)$, let $l(y) = -\log \frac{\hat{P}(y)}{\hat{Q}(y)}$. Observe that $l(y)$ has the

opposite sign as $F(y)$ and also that ϕ_x is continuously differentiable and $\phi_x'(y) = -l(y)$.

Given $\eta > 0$ we denote by $B_\eta(y) = (y - \eta, y + \eta)$ the η -neighborhood of $y \in I$.

We claim that any element of $Z(x)$ is an attractor associated to the mean flow restricted to interval I . Assume that $Z(x)$ is in the interior of I . We extend the argument for corner maximizers of ϕ_x below. Any interior minimizer $z \in Z(x)$ is a steady state for the mean flow F with $\phi_x'(z) = l(z) = F(z) = 0$.

Consider first the case in which z is a minimizer that is locally isolated. That is, there exists $\Delta > 0$ such that $\phi_x(z) < \phi_x(y)$ for any $y \in B_\Delta(z)$. Since, at any point, $\text{sign}(F) = -\text{sign}(l)$, to establish the claim it suffices to show that there exists a sufficiently small $\eta > 0$ such that $l(y)$ is strictly negative for all $y \in (z - \eta, z)$ and strictly positive for $y \in (z, z + \eta)$. Let $h(y, \delta) = \phi_x(y) + l(y)\delta$, which defines the hyperplane tangent to ϕ_x at y . Note that $h(z, \delta) = \phi_x(z) < \phi_x(z + \delta)$ for any $0 < |\delta| < \Delta$.

The continuity of ϕ_x and l ensures that there exists an $\Delta > \eta > 0$ such that for any $y \in B_\eta(z)$ we have $h(y, \delta) < \phi_x(y + \delta)$ for any $0 < |\delta| < \Delta$. In particular, if $0 < |\delta| < \eta$ the latter holds for $y = z - \delta$, so that $h(y, \delta) < \phi_x(z)$ or, equivalently, $\phi_x(y) + l(y)\delta < \phi_x(z)$. Since $\phi_x(z) < \phi_x(y)$ we conclude that $l(y) < 0$ for $y \in (z - \eta, z)$ and $l(y) > 0$ if $y \in (z, z + \eta)$.

Note that if a minimizer z is locally isolated to the right, i.e., there exists an open interval $(z, z + \Delta)$ to the right of z containing no other minimizers, then the argument above can be modified to show that there exists an $\eta > 0$ such that $l(y)$ strictly positive for $y \in (z, z + \eta)$. Indeed, the same argument can be applied to an auxiliary function $\bar{\phi}_x$ defined to be equal to ϕ_x for $y \leq z$ and $\bar{\phi}_x(z + \delta) = \phi_x(z - \delta)$ for each $0 < \delta < \Delta$. Similarly, for a minimizer z that is locally isolated to the left, there exists an $\eta > 0$ such that $l(y)$ strictly positive for $y \in (z - \eta, z)$.

Suppose now that a minimizer z is not locally isolated from other minimizers in $Z(x)$. This means that there exists a closed interval $J = [z_1, z_2] \subset \text{int}(I)$ which contains a subset of minimizers that is dense in J (any open subinterval of J contains a minimizer), the extremes z_1 and z_2 are minimizers (as $Z(x)$ is closed), z_1 is locally isolated to the left, and z_2 is locally isolated to the right. As argued above this means there exists an $\eta > 0$ such that $l(y)$ strictly positive for $y \in (z_1 - \eta, z_1)$ and $l(y)$ strictly negative for $y \in (z_2, z_2 + \eta)$. We claim that ϕ_x is flat on J and therefore this entire interval is an attractor of the flow F on I . Indeed, each minimizer $z \in J$ is such that $\phi_x'(z) = l(z) = 0$. Since minimizers are dense in J and l is continuous on J it follows that for any $y \in J$ we must have that $\phi_x'(y) = 0$.

We extend the argument for corner minimizers of $\phi_x(y)$ in interval I . Suppose that z is a corner minimizer and, for concreteness, assume that z is the left extreme of interval I . (The case in which z is the right extreme of the interval is analogous and omitted.) There are two cases. If $\phi_x'(z) = 0$ the same argument established above for interior maximizers carries through. Suppose instead that $\phi_x'(z) = l(z) > 0$. By continuity, there exists an open interval $(z, z + \Delta)$, $\Delta > 0$, to the right of z such that $l(y) > 0$ for all $y \in (z, z + \Delta)$. Since F and l have the opposite sign, the previous implies that z is an locally isolated attractor of F .

In conclusion, for each $x \in I$, $\phi_x(y)$ is maximized by attractors of the mean flow differential equation for F restricted to interval I . Let $\hat{y} \in \hat{A}_I$ be one such maximizer. Since $\phi_x(\hat{y}) = \hat{V}(x, \hat{y}) = -\hat{V}(\hat{y}, x) = -\phi_{\hat{y}}(x)$, the same argument above implies that the $\arg \max_{x \in I} \phi_x(\hat{y})$ is a subset of $\hat{A}_I$. It follows that $M_{\hat{V}}(I) \subseteq \hat{A}_I$.

(ii) The same argument used in the first part of the proof of Lemma 3 for σ shows that $x^* \in M_{\hat{V}}(I) \Leftrightarrow \hat{V}(x^*, y) \geq 0$ for all $y \in I$ with strict inequality if $y \notin M_V(I)$. $\square$

Lemma B3: Let A_I be the set of attractors of unperturbed mean flow $F = p - q$ and A_I^ε be the set of attractors of the perturbed mean flow $F^\varepsilon = P^\varepsilon - Q^\varepsilon$, both restricted to I . There exists $\bar{\varepsilon} > 0$ such that if $\varepsilon < \bar{\varepsilon}$ then:

- (i) For any $x^* \in A_I$ there exists a unique nearby attractor of the perturbed mean flow $x^\varepsilon \in A_I^\varepsilon$ and $\lim_{\varepsilon \rightarrow 0} x^\varepsilon = x^*$.
- (ii) If sequence $\{x^\varepsilon\}$ of attractors $x^\varepsilon \in A_I^\varepsilon$ converges to some x as $\varepsilon \rightarrow 0$ then $x \in A_I$.
- (iii) Any sequence $\{x^\varepsilon\}$ of attractors $x^\varepsilon \in A_I^\varepsilon$ can be partitioned into subsequences $\{x^{\varepsilon_a}\}$ that converge to an attractor $x_a \in A_I$ as $\varepsilon \rightarrow 0$.

Proof. (i) By Assumption (P3) an interior attractor x^* of the unperturbed mean flow F satisfies $F(x^*) = 0$ and $F'(x^*) < -d$. Assumptions (P1) and (P3) ensure that we can apply the Implicit Function Theorem on the perturbed flow to derive the result.

(ii) Consider a sequence $\{x^\varepsilon\}$, $x^\varepsilon \in A_I^\varepsilon$, that converges to some x as $\varepsilon \rightarrow 0$. Since the mean flow is regular on I we can find $c, d \in I$ such that for N sufficiently large and small enough ε each $x^\varepsilon \in [c, d]$ and this is the only attractor of F^ε that satisfies this. For each $\varepsilon \in (0, \bar{\varepsilon}]$ define the auxiliary function $W^\varepsilon(z) = \int_c^z F^\varepsilon(s) ds$, $z \in [c, d]$ and let

$W^0(z) = \int_c^z F(s) ds$. Observe that x^ε is the unique maximizer of $W^\varepsilon(z)$. Since $W^\varepsilon \rightarrow W^0$

uniformly on $[c, d]$ for any sequence $\varepsilon \rightarrow 0$ ($F^\varepsilon \rightarrow F$ uniformly) we can apply the Theorem of the Maximum (Berge [1997]) to conclude that x^* is a maximizer of $W^0(z)$. It can be verified that any maximizer of $W^0(z)$ must be an attractor of the unperturbed flow.

(iii) Consider any sequence of attractors $\varphi = \{x^\theta\}$, $x^\theta \in A_I^\varepsilon$. For each $x_a \in A_I$ let $\varphi_a = \left\{ x \in \varphi \mid x_a \in \arg \min_{x' \in A_I} |x - x'| \right\}$ be the set of elements of φ that are closest to x_a than

to any other attractor of the unperturbed flow. Clearly, the φ_a 's are a partition of φ . Without loss of generality we can focus on the case in which φ_a has an infinite number of elements, so that $\varphi_a = \{x^{\varepsilon_a}\}$ is a subsequence of φ (if for some x_a the set φ_a is finite we redefine the partition by assigning these elements to any other φ_b having an infinite number of elements). From (i) and (ii) it can be shown this subsequence must converge to x_a as $\varepsilon_a \rightarrow 0$. $\square$

Lemma B4. Let $M_{V^\varepsilon}(I)$ respectively $M_V(I)$ denote the maximum viscosity set for V^ε respectively V restricted to I and $\{x^\varepsilon\}$ be a sequence such that $x^\varepsilon \in M_{V^\varepsilon}(I)$.

(i) If $x^\varepsilon \rightarrow x^*$ as $(N, \varepsilon) \rightarrow (\infty, 0)$ then $x^* \in M_V(I)$.

(ii) For any sequence $y^{N, \varepsilon} \rightarrow y$ as $(N, \varepsilon) \rightarrow (\infty, 0)$ such that $y^{N, \varepsilon} \in I$ and $|y^{N, \varepsilon} - x^\varepsilon| > 0$ for all $(N, \varepsilon) \in \Theta$ and any $x^\varepsilon \in M_{V^\varepsilon}(I)$ then, either $\inf_{(N, \varepsilon) \in \Theta} V^\varepsilon(x^\varepsilon, y^{N, \varepsilon}) > 0$ or else $y \in M_V(I)$.

Proof. (i) From Lemma A2, $V^\varepsilon \rightarrow V$. This enables us to invoke the theorem of the maximum as in the proof of Lemma B3, which establishes (i).

(ii) We provide a detailed argument for the case in which $I \subseteq S_+$. If $|y^{N, \varepsilon} - x^\varepsilon| > 0$ for any $(N, \varepsilon) \in \Theta$ and any $x^\varepsilon \in M_{V^\varepsilon}(I)$ then $V^\varepsilon(x^\varepsilon, y^{N, \varepsilon}) > 0$ (Lemma B2(ii)). Hence, if $\inf_{(N, \varepsilon) \in \Theta} V^\varepsilon(x^\varepsilon, y^{N, \varepsilon}) = 0$ it must be the case that for some sequence $\{x^\varepsilon\}$, where $x^\varepsilon \in M_{V^\varepsilon}(I)$, $V^\varepsilon(x^\varepsilon, y^{N, \varepsilon}) \rightarrow 0$ as $(N, \varepsilon) \rightarrow (\infty, 0)$. From Lemma B3(iii), we can restrict to sequences that converge to some point $x \in M_V$. From Lemma A2 we have that $V^\varepsilon \rightarrow V$ uniformly from which $V^\varepsilon(x^\varepsilon, y^{N, \varepsilon}) \rightarrow V(x, y)$. This implies $V(x, y) = 0$. By Lemma B2 (ii) it follows that $y \in M_V(I)$. $\square$

REFERENCES

- Auriol, E. and Benaïm, M. [2001] “Standardization in Decentralized Economies,” *American Economic Review* 90: 550-570.
- Benaïm, M. and J. Weibull [2003] ““Deterministic approximation of stochastic evolution in games”, *Econometrica* 71, 873-903.
- Benaïm, M. and J. Weibull [2009] “Mean Field Approximations of Stochastic Population Processes in Games,” mimeo.
- Berge, C. [1997] “Topological Spaces: Including a Treatment of Multi-Valued Functions,” *Vector Spaces and Convexity*, Dover, New York.
- Bergin, J. and Lipman, B.L. [1996] “Evolution with state-dependent mutations,” *Econometrica*, 64: 943-956.
- Binmore, K., L. Samuelson, and R. Vaughn [1995] “Musical Chairs: Modeling Noisy Evolution, *Games and Economic Behavior* 11: 1-35.
- Binmore, K. and Samuelson, L. [1997] “Muddling Through: Deterministic Approximation of Stochastic Evolution in Games, *Journal of Economic Theory* 74: 235-265.
- Duflo, M. [1996] *Algorithmes Stochastiques*, Springer, New York.
- Ellison, G. [2000] “Basins of attraction, long run equilibria, and the speed of step-by-step evolution,” *Review of Economic Studies* 67(1): 17-45.
- Fudenberg, D. and L. Imhof [2006] “Imitation Processes with Small Mutations,” (with L. Imhof), *Journal of Economic Theory*, 131: 251–262
- Fudenberg, D., L. Imhof, M. Nowak and C. Taylor [2006] “Stochastic Evolution in Finite Populations,” *Theoretical Population Biology* 70: 352-363.
- Fudenberg, D., and L. Imhof [2007] “Monotone Imitation Dynamics in Large Populations,” mimeo.
- Grilo, I., O. Shy, and J. Thisse, [2001] "Price Competition When Consumer Behavior is Characterized by Conformity or Vanity, *Journal of Public Economics* 8: 385-408.
- Hofbauer, J. and Sigmund, K. [2003] “Evolutionary Games Dynamics,” *Bulletin (New Series) of the American Mathematical Society* 40: 479-519.

- Kandori, M., Mailath, G. and Rob, R. [1993] "Learning, mutation and long run equilibria in games," *Econometrica* 61: 29-56.
- Sandholm, W. [2009] "Orders of Limits for Stationary Distributions, Stochastic Dominance, and Stochastic Stability," mimeo.
- Traulsen, A. J.C. Claussen, C. Hauert [2005] "Coevolutionary Dynamics from Finite to Infinite Populations," *Physics Review Letters* 95.
- Young, P. [1993] "The Evolution of Conventions," *Econometrica* 61: 57-84.