Mathematical Biology

Long-time behavior of a PDE replicator equation for multilevel selection in group-structured populations

Daniel B. Cooney[1] · Yoichiro Mori[2]

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

In many biological systems, natural selection acts simultaneously on multiple levels of organization. This scenario typically presents an evolutionary conflict between the incentive of individuals to cheat and the collective incentive to establish cooperation within a group. Generalizing previous work on multilevel selection in evolutionary game theory, we consider a hyperbolic PDE model of a group-structured population, in which members within a single group compete with each other for individual-level replication; while the group also competes against other groups for group-level replication. We derive a threshold level of the relative strength of between-group competition such that defectors take over the population below the threshold while cooperation persists in the long-time population above the threshold. Under stronger assumptions on the initial distribution of group compositions, we further prove that the population converges to a steady state density supporting cooperation for between-group selection strength above the threshold. We further establish long-time bounds on the time-average of the collective payoff of the population, showing that the long-run population cannot outperform the payoff of a full-cooperator group even in the limit of infinitely-strong between-group competition. When the group replication rate is maximized by an intermediate level of within-group cooperation, individual-level selection casts a long shadow on the dynamics of multilevel selection: no level of

✉ Daniel B. Cooney
dbcooney@sas.upenn.edu

Yoichiro Mori
y1mori@sas.upenn.edu

[1] Department of Mathematics and Center for Mathematical Biology, University of Pennsylvania, Philadelphia, PA, USA

[2] Department of Mathematics, Department of Biology, and Center for Mathematical Biology, University of Pennsylvania, Philadelphia, PA, USA

between-group competition can erase the effects of the individual incentive to defect. We further extend our model to study the case of multiple types of groups, showing how the games that groups play can coevolve with the level of cooperation.

Keywords Multilevel Selection · Evolutionary Game Theory · Evolution of Cooperation · Replicator Equations · Hyperbolic PDEs

Mathematics Subject Classification 92D15 · 91A22

1 Introduction

Across a variety of biological and social systems, population structure often induces selective forces operating at multiple levels of organization. Of particular interest are hierarchical structures in which there is a tug-of-war between the interests at a one level of organization and the interests at a larger level. For problems of cooperation or collective behavior, the incentives of an individual to be a free-rider are often misaligned with the incentives of its group to produce a collective benefit for all of its members (Levin 2010). Considering the effects of selection at multiple levels of organization is particularly important in systems on the cusp of undergoing a transition to a higher order of complexity, establishing a collective that can compete or replicate as a single unit. Multilevel selection has been invoked to describe major evolutionary transitions, with examples ranging from the evolution of multicellularity (Rainey and Rainey 2003; Tarnita et al. 2013; Rose et al. 2020; Pichugin et al. 2015; Pichugin and Traulsen 2018; Staps et al. 2019), to the evolution of social group structure (Nowak et al. 2010; Fu et al. 2015). The transition to higher levels of biological complexity can be understood as a triumph of cooperative behavior via multilevel selection, in which groups can form a cooperative population structure overcoming individual competition within the group (Szathmáry and Smith 1995; Nowak 2006; Taylor and Nowak 2007).

Questions about multilevel selection have ranged widely across scales. On one end of the spectrum, ideas of conflict between individuality and collective behavior have been considered for the evolution of multicellularity (Tarnita et al. 2013; Pichugin and Traulsen 2018), replication control of plasmids (Paulsson 2002), and the evolution of mutualism in the microbiome (Van Vliet and Doebeli 2019). At the other end, the alignment of the individual-level and group-level incentives have been studied for problems ranging from collective hunting in animal groups (Boza and Számadó 2010) and the eusocial structure of insect colonies (Nowak et al. 2010; Fu et al. 2015) to the establishment of cooperative institutions for the management of common-pool resources (Ostrom 2010; Schlüter et al. 2016; Tavoni et al. 2012) and within-group cooperation coevolving with warfare (Henriques et al. 2019; Turchin 2010) in human societies. Experimental and field work has addressed problems ranging from the cooperative cofounding of ant colonies (Shaffer et al. 2016), to the establishment of multicellularity in biofilms (Rose et al. 2020) and the artificial selection for nonaggression in chickens (Muir 1996; Bijma et al. 2007b, a; Wade et al. 2010). A natural tension between selective forces at different levels of selection arises in the evolution of virulence in infectious disease dynamics, in which competition for pathogen replica-

tion within an individual host promotes selection for more virulent pathogens, while increased virulence can also harm the host and prevent onward transmission within the host population (Levin and Pimentel 1981; Dwyer et al. 1990; Gilchrist et al. 2004; Gilchrist and Coombs 2006; Blackstone et al. 2020; Boni et al. 2013).

One framework that has often been used to study multilevel selection is evolutionary game theory, which provides stylized models for the evolution of cooperative behavior in which individuals can maximize payoff by cheating, while groups achieve higher collective payoffs when at least some of their members cooperate. Traulsen and coauthors have studied the evolution of cooperation in the presence of multilevel selection, showing that group-level competition for replication could help to promote the fixation of cooperators over defectors in finite populations (Traulsen et al. 2005; Traulsen and Nowak 2006; Traulsen et al. 2008). The work of Simon and coauthors has further explored how more realistic mechanisms like group-level fission and fusion events and the possibility of non-constant group size can help to facilitate cooperation win out over the defection that is favored by individual-level selection (Simon 2010; Simon and Nielsen 2012; Simon et al. 2013; Simon and Pilosov 2016). Further work on stochastic multilevel selection models in evolutionary games have explored the role of group-level extinctions (Böttcher and Nagler 2016), the role of spatial structure on between-group competition (Akdeniz and van Veelen 2020), and asymptotic formulas for fixation probabilities in the limit of large population size (McLoone et al. 2018).

Luo (2014) introduced a stochastic model of two-level selection featuring two types of individuals: one with a constant reproductive advantage at the individual level (i.e. defectors), and the other that confers a selective advantage to its group (i.e. cooperators) (Luo 2014; van Veelen et al. 2014). In the limit of infinitely many groups of infinite size, Luo (2014) derived a non-local hyperbolic PDE describing the simultaneous competition within and between groups. Luo and Mattingly (2017) characterized the long-time behavior of this PDE based upon the relative strengths of selection at the two levels and the Hölder exponent of the initial condition near the full-cooperator group. They showed that there was a threshold level of between-group selection strength such that defectors would fix in the population when between-group competition below the threshold, while the population converges to a steady state density supporting positive levels of cooperation for between-group selection above the threshold. Further work on related nested birth-death models for multilevel selection has explored application to host-pathogen dynamics (Pokalyuk and Goerzer 2019; Pokalyuk and Wakolbinger 2019; Osorio and Winter 2020), as well as mathematical aspects of behavior in alternate infinite-population scaling limits, including fixation probabilities in stochastic Fleming-Viot models (Luo and Mattingly 2017; Dawson 2018; Meizis 2020) and quasi-stationary distributions in a Wright-Fisher diffusion equation with multilevel selection (Velleret 2019, 2020).

This model of two-level selection was later extended to include individual-level and group-level birth rates that depended on the personal and collective payoffs obtained from a two-strategy games played between members of the groups (Cooney 2019b). Results analogous to those of Luo and Mattingly were demonstrated for special cases of the Prisoners' Dilemma (PD) and Hawk-Dove (HD) games in which the within-group dynamics were exactly solvable, and further work explored the multilevel dynamics all two-player, two-strategy social dilemmas (Cooney 2020a) and in the presence

of within-group mechanisms of assortment or reciprocity (Cooney 2019a). For the PD and HD games, it was conjectured that, for sufficiently strong between-group competition, the population would converge to the unique steady state with the same Hölder exponent near $x = 1$ as that of the initial distribution. These steady state densities displayed a surprising property, called the "shadow of lower-level selection", in which the payoff of the modal group composition at steady state and the average payoff of the steady state population were limited by the payoff of the full-cooperator group. As a result, for games in which group payoff was maximized by intermediate levels of cooperation, the population always features less cooperation than optimal, even in the limit of infinitely strong between-group competition.

The PDE models of multilevel selection derived from two-level birth processes typically assume that between-group competition takes place through group-level birth events in which groups faithfully produce a copy of itself as an offspring group (Luo 2014; van Veelen et al. 2014; Cooney 2019b). In practical biology applications, competition between groups often manifests through differences in average growth rates (Chuang et al. 2009), fission events for cells or social groups with potentially uneven segregation of genes or behaviors (Szathmáry and Demeter 1987; Maynard Smith and Száthmary 1993; Van Vliet and Doebeli 2019; Simon and Pilosov 2016), differing rates of group-level extinction events (Böttcher and Nagler 2016), or explicit events of intergroup conflict (Boyd et al. 2003; Henriques et al. 2019). While the models with faithful group-level replication events provide a more stylized depiction of natural selection operating at multiple levels, the resulting PDE models provide a simple setting for exploring the tension between evolutionary forces favoring defection at the individual level and favoring positive levels of cooperation at the group level. The goal of this paper is to provide a complete analytical characterization of this class of PDE models for multilevel selection, allowing us to understand the ways in which our assumptions about individual-level and group-level competition impact the possibilities for long-time support for cooperation in group-structured populations. These results provide an analytical benchmark against which we can compare numerical study of more realistic models of multilevel selection that either feature stochastic, finite population effects (Traulsen and Nowak 2006; Luo 2014; Markvoort et al. 2014) or featuring group-level fission events (Szathmáry and Demeter 1987; Maynard Smith and Száthmary 1993; Simon and Pilosov 2016), and we hope that this analysis can be used as well to generate hypotheses for further empirical work on multilevel selection.

In this paper, we characterize the long-time behavior for a broad class of models for multilevel selection. While previous work had shown convergence to steady state densities for several one-parameter families of models for multilevel selection arising from special cases of PD games (Luo and Mattingly 2017; Cooney 2019b), the techniques used in those cases relied on the ability to obtain explicit solutions for the characteristic curves describing the within-group dynamics, making it difficult to extend the results to the more general situation. Here, we obtain careful estimates on the solutions along characteristics and extract a principal growth rate for the multilevel dynamics, allowing us to prove convergence to steady state for multilevel PDEs with continuously differentiable individual-level and group-level replication rates in which defectors have an individual-level advantage over cooperators and all-cooperator groups have a collective advantage over all-defector groups. This result confirms a previous conjecture for

the long-time dynamics multilevel replicator equations arising from any PD game, and extends the scope of that conjecture to include applications to a range of topics including protocell evolution and the origin of chromosomes (Cooney et al. 2021). Through this more general formulation of our model for multilevel selection, we are able to formalize previous intuition to understand how the possibility of achieving cooperation at steady state relies on the ability for the collective advantage of full-cooperator groups over full-defector groups to overcome the individual-level advantage of defecting in a group with many cooperators.

We also extend our analysis of our generalization of the multilevel Prisoners' Dilemma dynamics to study long-time behavior for initial populations beyond the class of measures with a well-defined Hölder exponent near full-cooperation that was considered in previous work (Luo and Mattingly 2017; Cooney 2019b, 2020a). By characterizing the tail behavior of the initial measure through quantities that we call the supremum and infimum Hölder exponent, we can obtain upper and lower bounds for the principal growth rate for solutions along characteristics for any initial measure. Using these estimates, we find that the population will not converge to a density steady for any initial measure without a well-defined Hölder exponent near full-cooperation. However, we show that, for any initial measure, defectors will take over the population when between-group competition is sufficiently weak, while cooperation will survive in the long-time limit in the sense of weak persistence when between-group selection exceeds a threshold value that depends on the supremum Hölder exponent of the initial measure. We also use our estimates on the principal growth rate to derive long-time upper and lower bounds on the time-average of the average group-level replication rate of the population, showing that the long-time collective outcome cannot exceed the group-level replication rate of the all-cooperator group. This observation serves as a dynamical analogue of the "shadow of lower-level selection", showing how this limitation on collective outcome stems from the tug-of-war between the collective incentive to cooperate and the individual incentive to defect.

We also characterize the multilevel dynamics for a generalization of the Prisoners' Delight game, in which cooperation is favored at both levels of selection, extracting a principal growth rate to show how full-cooperation is achieve via multilevel selection. Combining this with the results mentioned above for our generalization of the multilevel Prisoners' Dilemma dynamics, we fully characterize the dynamics of our two-level replicator equation for scenarios in which the within-group dynamics features no interior equilibria. This approach of extracting a principal growth rate for a population can be further applied to understand the a generalization of the multilevel dynamics to a case in which our group-structured population may feature different possible individual-level and group-level replication rates. We provide a sufficient condition for the long-time concentration of the population upon a single type of group under our multi-type two-level birth-death dynamics, showing that the long-time dynamics will favor the dominance of the group type feature the maximal principal growth rate. This result can be used to study the coevolution of group features and the strategic composition of groups, showing how multilevel competition can help to select the games played within groups. As one application of this framework with multiple group types, we show that this concentration result can be used to study the dynamics of generalized versions of Hawk-Dove and Stag-Hunt games, which each

feature an interior within-group equilibrium. The results for these cases confirm and extend existing conjectures on convergence to steady state for the multilevel dynamics of those games (Cooney 2020a).

In Sect. 1.1, we describe the mathematical formulation of our model of multilevel selection with comparisons to previous work on multilevel selection in evolutionary games. In Sect. 1.2, we summarize our main results for the long-time dynamics of our PDE model of multilevel selection. Section 1.3 provides an outline for the remainder of the paper.

1.1 Model of multilevel selection

For our model of multilevel selection, we consider a population with m groups that are each composed of n members. The within-group selection follow a frequency-dependent Moran process replacing a randomly chosen member of the same group. In a group with i cooperators, cooperators and defectors give birth with rates $1 + w_I \pi_C(\frac{i}{n})$ and $1 + w_I \pi_D(\frac{i}{n})$, respectively, where $\pi_C(\cdot)$ and $\pi_D(\cdot)$ are C^1 functions on [0, 1] and w_I is the intensity of selection for within-group competition. We can further consider the advantage of defectors over cooperators under within-group competition in an i-cooperator group through the quantity

$$\pi\left(\frac{i}{n}\right) := \pi_D\left(\frac{i}{n}\right) - \pi_C\left(\frac{i}{n}\right). \tag{1.1}$$

Between-group competition takes place through a group-level birth-death process in which a group with i cooperators produces a copy of itself and replaces a randomly chosen group with rate $\Lambda(1 + w_G G(\frac{i}{n}))$, where w_G is the selection intensity of between-group competition and Λ describes the relative rate of within-group and between-group replication events. We note that, for any constant $s > 0$, the choice of $\pi(x) = s$ and $G(x) = x$ recovers the functions for within-group and between-group competition for the Luo-Mattingly model (Luo 2014; Luo and Mattingly 2017).

In the limit as the number of groups and group size tend to infinity ($m \to \infty, n \to \infty$), we can describe the composition of strategies in the group-structured population by $u(t, x)$, the probability density of groups composed x cooperators and $1 - x$ defectors at time t. Using either a heuristic derivation (Luo 2014; van Veelen et al. 2014; Cooney 2019b) or a weak convergence argument (Luo and Mattingly 2017), we can show that the large-population limit of the stochastic ball-and-urn process can be described by the following partial differential equation for the change of $u(t, x)$ in time

$$\frac{\partial u(t, x)}{\partial t} = \frac{\partial}{\partial x}\left(x(1 - x)\pi(x)u(t, x)\right) + \lambda\left(G(x) - \int_0^1 G(y)u(t, y)dy\right)u(t, x), \tag{1.2}$$

where $\lambda := \frac{\Lambda w_G}{w_I}$ describes the relative strength of within-group and between-group competition. The first term on the right-hand side of Eq. (1.2) describes the dynamics of within-group competition, in which defectors (respectively cooperators) increase

in frequency within groups when $\pi(x) > 0$ (respectively $\pi(x) < 0$). The second term in Eq. (1.2) describes the impact of between-group competition, and groups with composition x increase in frequency when their replication rate $G(x)$ exceeds the average group-replication rate in the population $\int_0^1 G(y)u(t, y)dy$.

Equation (1.2) is paired with an initial density given by

$$u(0, x) = u_0(x) \geq 0, \quad \int_0^1 u_0(x)dx = 1. \tag{1.3}$$

We can check that, if $u(x, t)$ is a solution to Eq. 1.2, it will be of unit mass for all t. Equation (1.2) is a hyperbolic PDE, whose characteristic curves are given by solutions of the following ODE

$$\frac{dx(t)}{dt} = -x(1 - x)\pi(x) \ , \ x(0) = x_0, \tag{1.4}$$

which we note is the well-known replicator equation for individual-level selection within a given group (Hofbauer and Sigmund 1998). The function $\pi(x)$, sometimes called the gain function (Bach et al. 2006; Kaznatcheev et al. 2017), describes the relative advantage of defectors over cooperators under within-group competition in an x-cooperator group.

We can also consider a measure-valued formulation corresponding to the multilevel dynamics described by Eq. (1.2). For an initial Borel probability measure $\mu_0(dx)$ and any $C^1([0, 1])$ test-function $v(x)$, the measure $\overline{\mu}_t(dx)$ changes in time according to

$$\frac{\partial}{\partial t} \int_0^1 v(x)\overline{\mu}_t(dx) = -\int_0^1 \frac{\partial v(x)}{\partial x}x(1 - x)\pi(x)\overline{\mu}_t(dx)$$

$$+ \lambda \int_0^1 v(x)\left[G(x) - \int_0^1 G(y)\overline{\mu}_t(dy)\right]\overline{\mu}_t(dx). \tag{1.5}$$

To study solutions for Eq. 1.5, we can introduce an auxiliary linear equation given by

$$\frac{\partial}{\partial t} \int_0^1 v(x)\mu_t(dx) = -\int_0^1 \frac{\partial v(x)}{\partial x}x(1 - x)\pi(x)\mu_t(dx) + \lambda \int_0^1 v(x)G(x)\mu_t(dx), \tag{1.6}$$

paired with initial the measure $\mu_0(dx)$. We can check that solutions $\overline{\mu}_t(dx)$ to the measure-valued multilevel dynamics of Eq. (1.5) can be related to solutions $\mu_t(dx)$ of Eq. (1.6) through the normalization given by

$$\overline{\mu}_t(dx) = \frac{\mu_t(dx)}{\int_0^1 \mu_t(dy)}. \tag{1.7}$$

The dynamics of Eq. (1.6) also have independent biological interest for studying multilevel selection in which x-cooperator reproduce with rate $G(x)$ and no groups

are removed from the population. Such models of an expanding group-structured population may be relevant in applications in which the group-level reproduction corresponds to cell division (Fontanari et al. 2006; Fontanari and Serva 2013, 2014a, b) or fission of social groups (Simon and Pilosov 2016; Gueron and Levin 1995).

We will also consider an extension of the multilevel dynamics in which groups belong to one of N possible subpopulations, with each subpopulation featuring its own reproduction rates $\pi_j(\cdot)$ and $G_j(\cdot)$. Within-group competition proceeds according to $\pi_j(\cdot)$, while between-group competition consists of group replicating with rate proportion to $G_j(\cdot)$ and replacing a randomly-chosen group from any of the N subpopulations. For example, each subpopulation could be defined by a different two-strategy game played within its groups, and then the corresponding multilevel dynamics describe the coevolution of cooperation and the fraction of groups playing each game.

Denoting the set of subpopulations by $\mathcal{N} = \{1, \cdots, N\}$, we describe the composition of x-cooperator groups in the subpopulation $j \in \mathcal{N}$ at time t by the density $u^j(t, x)$. This family of densities changes in time according to PDEs of the form

$$
\begin{aligned}
\frac{\partial u^j(t, x)}{\partial t} &= \frac{\partial}{\partial x} \left[x(1-x)\pi_j(x)u^j(t, x) \right] \\
&+ \lambda u^j(t, x) \left[G_j(x) - \sum_{k=1}^{N} G_k(y)u^k(t, y)dy \right],
\end{aligned}
\tag{1.8}
$$

for each $j \in \mathcal{N}$, and this system is paired with an initial density satisfying

$$
u^j(t, x) = u_0^j(x) \geq 0, \quad \sum_{k=1}^{N} \int_0^1 u_0^j(x)dx = 1.
\tag{1.9}
$$

We can also consider a measure-valued analogue of our multipopulation model by describing the strategic composition of groups in the jth subpopulation by the measure $\mu_t^j(dx)$. For a $C^1([0, 1])$ test-function $v_j(x)$, this measure changes according to the following equation

$$
\begin{aligned}
\frac{\partial}{\partial t} \int_0^1 v_j(x)\mu_t^j(dx) &= - \int_0^1 \frac{\partial v_j(x)}{\partial x} x(1-x)\pi_j(x)\mu_t^j(dx) \\
&+ \lambda \int_0^1 v_j(x) \left[G_j(x) - \sum_{k=1}^{n} \left(\int_0^1 G_k(y)\mu_t^k(dy) \right) \right] \mu_t^j(dx),
\end{aligned}
\tag{1.10}
$$

where the subpopulations are coupled through the nonlocal regulation term describing between-group competition. The system described by Eq. (1.10) is paired with initial measures given by the measures $\mu_0^j(dx)$ for $j \in \mathcal{N}$, which together satisfy the

normalization condition given by

$$\sum_{j=1}^{N} \mu_0^j([0,1]) = \sum_{j=1}^{N} \int_0^1 \mu_0^j(dx) = 1. \tag{1.11}$$

We can also associate with Eq. (1.10) a system of N decoupled linear equations of the form

$$\frac{\partial}{\partial t} \int_0^1 v_j(x)\mu_t^j(dx) = -\int_0^1 \frac{\partial v_j(x)}{\partial x} x(1-x)\pi_j(x)\mu_t^j(dx)$$

$$+ \lambda \int_0^1 v_j(x)G_j(x)\mu_t^j(dx). \tag{1.12}$$

Given solutions $\mu_t^1(dx), \cdots, \mu_t^N(dx)$ to the linear dynamics of Eq. (1.12), we can find a corresponding solution $\overline{\mu}_t^j(dx)$ to Eq. (1.10) for the jth subpopulation through the normalization given by

$$\int_0^1 v_j(x)\overline{\mu}_t^j(dx) = \frac{\int_0^1 v_j(x)\mu_t^j(dx)}{\sum_{k=1}^N \mu_t^k([0,1])} \tag{1.13}$$

1.1.1 Motivating example: two-strategy evolutionary games

To formulate assumptions about the behavior of the functions $\pi(x)$ and $G(x)$ characterizing within-group and between-group competition, we can consider the special case of the multilevel selection dynamics depend on payoffs from two-player, two-strategy social dilemmas (Cooney 2019b, 2020a). We consider games with symmetric payoff matrices of the form

$$\begin{array}{c} \quad\quad C \quad D \\ \begin{array}{c} C \\ D \end{array} \begin{pmatrix} R & S \\ T & P \end{pmatrix}, \end{array} \tag{1.14}$$

where the entries of the payoff matrix correspond to the reward for mutual cooperation (R), the sucker payoff from cooperating with a defector (S), the temptation to defect against a cooperator (T), and the punishment for mutual defection (P). In this paper, we will consider the multilevel dynamics corresponding to generalizations of four two-strategy social dilemmas: the Prisoners' Dilemma (PD), the Hawk-Dove game (HD), the Stag-Hunt (SH), and the Prisoners' Delight (PDel). These four games are characterized by the following rankings of payoffs

$$\text{PD} : T > R > P > S \tag{1.15a}$$

$$\text{HD} : T > R > S > P \tag{1.15b}$$

$$\text{SH} : R > T > P > S \tag{1.15c}$$

$$\text{PDel}: R > T > S > P. \tag{1.15d}$$

For a group composed of fractions x cooperators and $1 - x$ defectors, the average payoff for a cooperator and defector are given by

$$\pi_C(x) = Rx + S(1 - x) \tag{1.16a}$$
$$\pi_D(x) = Tx + P(1 - x). \tag{1.16b}$$

For this example of multilevel selection with replication rates arising from the payoffs of two-player, two-strategy games, we will assume that the group-level reproduction rate in an x-cooperator group depend on the average payoff of group members $G(x) = x\pi_C(x) + (1 - x)\pi_D(x)$ (Cooney 2019b, 2020a). Using the cooperator and defector payoffs from Eq. (1.16), we then have that the dynamics of Eq. (1.2) have the following dependence on the payoff matrix from Eq. (1.14)

$$\pi(x) = P - S - (R - S - T + P)x \tag{1.17a}$$
$$G(x) = P + (S + T - 2P)x + (R - S - T + P)x^2. \tag{1.17b}$$

From Eq. 1.17b, we can see that the group-reproduction function satisfies

$$G(1) = R > P = G(0) \tag{1.18}$$

for each of the social dilemmas described in Eq. (1.15). Under the game-theoretic model, the characteristic curves from Eq. (1.4) change with time according to

$$\frac{dx(t)}{dt} = -x(1 - x)\left[P - S - (R - S - T + P)x\right], \tag{1.19}$$

which has equilibria at $x = 0$, $x = 1$, and a possible interior equilibrium $x = x_{eq}$ given by

$$x_{eq} = \frac{P - S}{R - S - T + P}. \tag{1.20}$$

Furthermore, we note from Eq. (1.17a) that $\pi(0) = P - S$, $\pi(1) = T - R$, and that $\pi(x)$ interpolates linearly between $\pi(0)$ and $\pi(1)$ as x moves between 0 and 1. Using these facts and the payoff rankings from Eq. (1.15a), we can deduce that $\pi(x) > 0$ for $x \in [0, 1]$ for the PD game, and therefore defectors always have an individual-level advantage over cooperators in this game. Using Eq. (1.15d), we see by analogous reasoning that $\pi(x) < 0$ for $x \in [0, 1]$ for the PDel game, resulting in an individual-level advantage for cooperators over defectors in this game.

Using these properties, we formulate a generalization of the multilevel PD and PDel dynamics by considering $G(x), \pi(x) \in C^1([0, 1])$ satisfying $G(1) > G(0)$ and either $\pi(x) > 0$ (in the PD case) or $\pi(x) < 0$ (in the PDel case). Notably, this class of reproduction functions $G(x)$ and $\pi(x)$ allows us to consider group-reproduction rates $G(x)$ which are not given by the average of the individual-level replication rates

of cooperators $\pi_C(x)$ and defectors $\pi_D(x)$, and this broader class of models includes those used in the Luo-Mattingly model (Luo 2014), in models used to describe the evolution of protocells (Cooney et al. 2021), and within-group dynamics following the Fermi update rule for social learning (Traulsen et al. 2005). We will take this generalization of the PD and PDel dynamics as a generic picture of multilevel selection in populations without internal within-group equilibria, and will assume that the dynamics of multiple subpopulations described by Eq. (1.8) reflects a PD or PDel scenario within a given subpopulation. For games such as the HD and SH with internal within-group equilibria, we will apply the multiple subpopulation formulation using the fact that the dynamics above and below the within-group equilibria for such games reflect either a PD or PDel scenario.

For the more general class of models considered in this paper, we only impose assumptions on the value of the group-level replication rate $G(x)$ for levels of cooperation x that correspond to equilibria of the within-group dynamics. In particular, by only imposing the assumption that $G(1) > G(0)$ for our generalization of multilevel PD and PDel competitions, we can consider a variety of group-level replication rates for which $G(x)$ can exceed the all-cooperator replication rate $G(1)$ for a range of intermediate levels of cooperation $x \in (0, 1)$. In addition to the example considered above of PD games with $S + T > 2R$, the possibility of group-level competition most favoring a mix of cooperators and defectors arises in biological settings in which individuals pay a cost to provide a public good for their group. Experimental work in yeast has shown the possibility of populations achieving maximal growth rate by containing only a fraction of individuals with a gene required to produce proteins that break down sugars (MacLean et al. 2010), showing that a mix of producers and non-producers can best balance the collective benefit from the presence of the protein and the individual metabolic cost of producing this diffusible public good. This form of the generalized multilevel PD dynamics may arise in scenarios in which individual-level competition for reproduction is paired with a collective benefit for complementary genes or behaviors, with examples ranging from division-of-labor in animal groups (Boza and Számadó 2010) to the evolution of complementary genes in early cellular life (Maynard Smith and Száthmary 1993; Fontanari et al. 2006; Fontanari and Serva 2014b). Empirical studies for this kind of competition have been explored in several settings in the context of infectious disease, from the demonstration of heterotypic competition between complementary strains of influenza (Xue et al. 2016) to the evolution of genetic linkage and antiobiotic multiresistance that arises from competition between bacteriophages that confer resistance to alternate antibiotics (Sachs and Bull 2005).

1.2 Summary of main results

Now we present our main results for the dynamics of solutions to Eq. (1.2). In Sect. 1.2.1, we introduce the family of probability densities that are steady state solutions of Eq. (1.2) in the PD case, highlighting how the collective at steady state is limited by the group-level reproduction rate of the all-cooperator group. In Sect. 1.2.2, we present Theorem 1, showing convergence of the population to a steady state density

for the special of initial measures with a well-defined Hölder exponent near $x = 1$. In Sect. 1.2.3, focus on the characterization of long-time behavior of the multilevel PD dynamics for all possible initial measures, providing bounds on the long-time collective outcome in Theorem 2 and classifying the conditions for long-time extinction or persistence of cooperation in Theorem 3. In Sect. 1.2.4, we turn to the dynamics of Eq. (1.8) describing multilevel competition in the presence of multiple types of groups. We present Theorem 4, providing a sufficient condition for the population to concentrate upon the group type with maximal principal growth rate.

1.2.1 Steady state densities for multilevel PD dynamics

We first look to understand steady-state solutions to Eq. (1.2) in the Prisoners' Dilemma case, and the conditions under which such solutions can be achieved under the multilevel dynamics. In Sect. B, we show that, under generic conditions on λ, $G(\cdot)$, and $\pi(\cdot)$, the achievable steady states are delta-concentrations of full-defector groups $\delta(x)$ and full-cooperator groups $\delta(1 - x)$, as well as a family of density steady states which we characterize below.

In previous work on special cases of Eq. (1.2), the long-time behavior and convergence to steady state of the multilevel dynamics was studied for initial measures $\mu_0(dx)$ with given Hölder exponent near $x = 1$ (Luo and Mattingly 2017; Cooney 2019b, 2020a). This Hölder exponent and its associated Hölder constant quantify the extent to which the initial distribution concentrates or decays near the full-cooperator group, and is defined as follows.

Definition 1 The measure $\mu_t(dx)$ has Hölder exponent $\theta_t \geq 0$ near $x = 1$ with associated Hölder constant $C_\theta \in \mathbb{R}_{\geq 0} \cup \{\infty\}$ if it satisfies the following limiting behavior

$$\lim_{x \to 0} \frac{\mu_t([1-x, 1])}{x^\Theta} = \begin{cases} 0 & : \Theta < \theta \\ C_\theta & : \Theta = \theta \\ \infty & : \Theta > \theta \end{cases} . \tag{1.21}$$

Measures of the form $\mu(dx) = \theta(1 - x)^{\theta-1}dx$ for finite $\theta > 0$ have Hölder exponent θ near $x = 1$. Examples with Hölder exponent of 0 and ∞ are measures satisfying $\mu(\{1\}) > 0$ and $\mu([1 - \epsilon, 1]) = 0$ for some $\epsilon > 0$, respectively.

We show in Sect. 2 that steady state density solutions to Eq. (1.2) can be parametrized by λ and their Hölder exponent $\theta > 0$ near $x = 1$. We calculate that these steady states are given by the following densities

$$p_\theta^\lambda(x) = \frac{f_\theta^\lambda(x)}{\int_0^1 f_\theta^\lambda(x)dx},$$

$$f_\theta^\lambda(x) = x^{\nu_\theta - 1}(1 - x)^{\theta - 1}\frac{\pi(1)}{\pi(x)} \exp\left(-\lambda \int_x^1 \frac{\tilde{C}(s)ds}{\pi(s)}\right), \tag{1.22}$$

where the parameter ν_θ corresponds to

$$\nu_\theta := \left(\frac{1}{\pi(0)}\right)(\lambda[G(1) - G(0)] - \theta\pi(1)) \tag{1.23}$$

and the term $-\lambda\tilde{C}(x)$ is given by

$$-\lambda\tilde{C}(x) = \lambda\left(\frac{G(x) - G(0)}{x}\right) + \nu_\theta\left(\frac{\pi(x) - \pi(0)}{x}\right)$$
$$+ \lambda\left(\frac{G(x) - G(1)}{1 - x}\right) - \theta\left(\frac{\pi(x) - \pi(1)}{1 - x}\right). \tag{1.24}$$

The form of these steady state densities highlights the key contribution of the collective reproduction rates $G(1)$ and $G(0)$ and individual-level gain functions $\pi(1)$ and $\pi(0)$ for the all-cooperator and all-defector groups.

Furthermore, our assumption of $C^1([0, 1])$ replication rates $G(\cdot)$ and $\pi(\cdot)$ allows us to use Equation (1.24) to deduce the boundedness of $\lambda\tilde{C}(x)$ on $[0, 1]$. Because we are considering $\theta > 0$, we see from Eqs. (1.22) and (1.23) that the density $f_\theta^\lambda(x)$ is integrable provided that $\nu_\theta > 0$, which occurs when between-group competition is sufficiently strong so that

$$\lambda > \lambda^* := \frac{\theta\pi(1)}{G(1) - G(0)}. \tag{1.25}$$

This threshold condition is increasing in $\pi(1)$, the relative within-group advantage of a defector over a cooperator in a full-cooperator group, and is decreasing in $G(1) - G(0)$, the relative between-group advantage of a full-cooperator group over a full-defector group. From this we see that the ability to promote cooperation via multilevel selection can be understood as the collective incentive to cooperate winning out over the individual incentive to defect against cooperators. In addition, λ^* is an increasing function of θ, which means that larger cohorts of near full-cooperator groups are able to maintain a steady state density featuring cooperation over a large range in the strength of between-group competition.

We further show in Sect. 2 that the average level of group reproduction $G(x)$ achieved in the steady state distribution p_θ^λ is given by

$$\langle G(\cdot)\rangle_{p_\theta^\lambda} = \int_0^1 G(y)p_\theta^\lambda(y)dy = \left(\frac{\lambda^*}{\lambda}\right)G(0) + \left(1 - \frac{\lambda^*}{\lambda}\right)G(1). \tag{1.26}$$

The average group reproduction function (or collective payoff in the game-theoretic scenario) interpolates between the collective outcome of all-defector groups to that of all-cooperator groups as λ ranges from λ^* to ∞, and that the average at steady state $\langle G(\cdot)\rangle_{p_\theta^\lambda}$ is limited by the collective outcome for the all-cooperator group. In particular, this means that if $G(x)$ is maximized by an interior level of cooperation x^*, then the average collective reproduction at steady state does not achieve its optimal
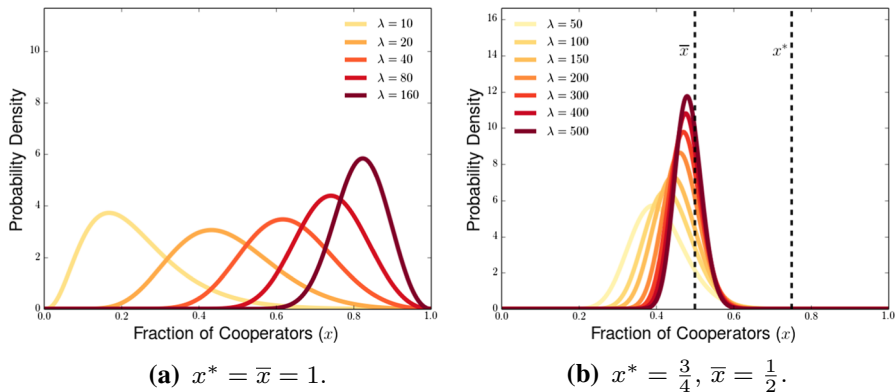
(a) $x^* = \overline{x} = 1$. **(b)** $x^* = \frac{3}{4}$, $\overline{x} = \frac{1}{2}$.

Fig. 1 Steady state densities from Eq. (1.22) for various values of between-group selection intensity λ for the game-theoretic case of Sect. 1.1.1 and scenarios in which $G(x)$ is maximized by full-cooperation (left) or by 75 percent cooperators (right). The dashed lines in the bottom panel correspond to the group type x^* maximizing $G(x)$ and the group type $\overline{x}$ for which $G(\overline{x}) = G(1)$. Densities have Hölder exponent $\theta = 3$ near $x = 1$, and payoff parameters are chosen from the family of games with $S = 1$, $P = 2$, $T = R + 2$ with either $R = 2.5$ (left) or $R = 3$ (right)

possible level even in the limit of infinitely strong between-group competition when $\lambda \to \infty$. This generalizes the so-called "shadow of lower-level selection" seen in previous work on multilevel selection in evolutionary games.

We may also visualize the impact of increasing the intensity of between-group competition by plotting the steady state densities from Eq. (1.22) for various values of λ. In Figure 1, we display these steady state densities for special cases of the game-theoretic examples from Sect. 1.1.1 in which the collective payoff of the group is either maximized by full-cooperation ($x^* = 1$, Figure 1a) or by a composition of 75 percent cooperators ($x^* = \frac{3}{4}$, Figure 1b). In the former case, we see arbitrarily high levels of cooperation are achieved by the group for sufficiently large λ, whereas in the latter case the densities do not come close to achieving the optimal composition of cooperators. In fact, in the latter case, the density appears to concentrate around $\overline{x} = \frac{1}{2}$, the unique interior level of cooperation satisfying $G(x) = G(1)$. In the appendix, we formalize this intuition from Figure 1 to show that the steady states of Eq. (1.2) in the PD case concentrate as $\lambda \to \infty$ upon measures supported only at points satisfying $G(x) = G(1)$.

1.2.2 Convergence to steady state densities

Next, we explore the conditions under which steady states of the form $p_\theta^\lambda(x)$ are achieved as the long-time behavior under the dynamics of Eq. (1.2). Considering initial populations that have a well-defined Hölder exponent, we show in Theorem 1 that solutions $\mu_t(dx)$ to Eq. (1.2) converge weakly to a steady state density when $\lambda > \lambda^*$. This result confirms and generalizes (Cooney 2020a, Conjecture 1), which addresses convergence to steady-state for multilevel selection in the case of replication rates arising from Prisoners' Dilemma games with the payoff matrix of Eq. (1.14).

Theorem 1 *Suppose that* $G(x), \pi(x) \in C^1([0, 1])$, $G(1) > G(0)$, *and* $\pi(x) > 0$ *for* $x \in [0, 1]$. *Consider an initial measure* $\mu_0(dx)$ *having a Hölder exponent* $\theta > 0$ *near* $x = 1$ *with corresponding positive, finite Hölder constant* C_θ. *If* $\lambda[G(1) - G(0)] > \theta\pi(1)$, *then* $\bar{\mu}_t$ *converges weakly to the probability measure defined by the density function* $p_\theta^\lambda(x)$ *defined in Eq.* (1.22):

$$\lim_{t\to\infty} \int_0^1 v(x)\overline{\mu}_t(dx) = \int_0^1 v(x)p_\theta^\lambda(x)dx \qquad (1.27)$$

where $v(x)$ *is an arbitrary continuous function on* [0, 1].

The above, and indeed, many of our main results to follow depend on the careful study of the characteristic curves solving Eq. (1.4) and the solutions along characteristics, which respectively describe the effects of within-group and between-group competition. Because $\pi(\cdot)$ is positive for the Prisoners' Dilemma case, $x = 0$ and $x = 1$ are the only steady states of this ODE and therefore the characteristic curves spend most of their time near $x = 0$ or $x = 1$. Thus, $G(0)$ and $G(1)$, the group replication rates near $x = 0$ and $x = 1$ respectively, and $\pi(1)$, the speed with which the ODE trajectory leaves $x = 1$, can be expected to control growth of the unnormalized solution $\mu_t(dx)$ to Eq. (1.6). This intuition is made precise by Lemma 7, where we use the continuous differentiability of $\pi(\cdot)$ and $G(\cdot)$ on [0, 1] to decompose $\mu_t(dx)$ into the product of a bounded, continuous function and an exponentially growing term whose growth rate is given by $\max\{\lambda G(0), \lambda G(1) - \theta\pi(1)\}$. This decomposition allow us to prove Theorem 1 using our knowledge of $G(\cdot)$ and $\pi(\cdot)$ at the endpoints 0 and 1, without requiring explicit expressions for the characteristic curves and solutions alongs characteristics only available in special cases (Luo and Mattingly 2017; Cooney 2019b). In addition to generalizing previous results, this approach allows us to glean further biological intuition into how the long-time behavior of Eq. (1.2) results from a tension between the collective incentive to achieve full-cooperation over full-defection $G(1) - G(0)$ and the individual-level incentive to defect $\pi(1)$ in a group with many cooperators.

1.2.3 Long-time behavior for more general initial measures

Not all initial measures have a well-defined Hölder exponent as defined in Definition 1, as the limit characterizing the Hölder exponent and constant in Eq. (1.21) does not necessarily exist. For such measures, the foregoing results do not apply, but we can still provide a characterization of the long-time behavior that holds for any initial measure and any relative strength of between-group selection. In Theorem 2, we derive long-time upper and lower bounds for the time-averaged collective payoff of the population, showing that the long-time collective outcome is limited by the replication rate of the all-cooperator group. In Proposition 3, we show that, for a given initial measure, there is a threshold strength of between-group competition required for cooperation to survive in the long-time population.

To supplement the Hölder exponent in characterizing initial measures by the behavior near $x = 1$, we introduce the following quantities that can be defined for any initial measure.

Definition 2 The infimum Hölder exponent $\overline{\theta}$ near $x = 1$ satisfies

$$\overline{\theta} := \sup \left\{ \Theta \geq 0 : \liminf_{x \to 0} \frac{\mu_t([1 - x, 1])}{x^{\Theta}} = 0 \right\} \tag{1.28}$$

Furthermore, the infimum Hölder constant $C_{\overline{\theta}}$ is given by

$$\liminf_{x \to 0} \frac{\mu_t([1 - x, 1])}{x^{\overline{\theta}}} = C_{\overline{\theta}}. \tag{1.29}$$

Definition 3 The supremum Hölder exponent $\underline{\theta}$ near $x = 1$ satisfies

$$\underline{\theta} := \sup \left\{ \Theta \geq 0 : \limsup_{x \to 0} \frac{\mu_t([1 - x, 1])}{x^{\Theta}} = 0 \right\} \tag{1.30}$$

Furthermore, the supremum Hölder constant $C_{\underline{\theta}}$ is given by

$$\limsup_{x \to 0} \frac{\mu_t([1 - x, 1])}{x^{\underline{\theta}}} = C_{\underline{\theta}}. \tag{1.31}$$

Remark 1 The infimum and supremum Hölder exponents satisfy the inequality $\underline{\theta} \leq \overline{\theta}$. This is true because, for any $\Theta \geq 0$,

$$\liminf_{x \to 0} \frac{\mu_t([1 - x, 1])}{x^{\Theta}} \leq \limsup_{x \to 0} \frac{\mu_t([1 - x, 1])}{x^{\Theta}}. \tag{1.32}$$

Therefore if the left-hand side is positive for a given Θ, then right-hand side is positive as well.

Furthermore, if there are θ and C_θ such that the infimum and supremum Hölder exponents satisfy $\overline{\theta} = \underline{\theta} = \theta$ and infimum and supremum Hölder constants satisfy $C_{\overline{\theta}} = C_{\underline{\theta}} = C_\theta$, then Definitions 2 and 3 imply that the limiting behavior of Eq. (1.21) is satisfied by the measure $\mu_t(dx)$. In other words, if the infimum and supremum Hölder data agree for a measure $\mu_t(dx)$ near $x = 1$, then $\mu_t(dx)$ has a well-defined Hölder exponent θ and Hölder constant C_θ near $x = 1$ in the sense of Definition 1.

Remark 2 It can be shown that, if our initial measure $\mu_0(dx)$ has infimum and supremum Hölder exponents $\overline{\theta}$ and $\underline{\theta}$ near $x = 1$ with constants $C_{\overline{\theta}}$ and $C_{\underline{\theta}}$, then the solution $\mu_t(dx)$ to Eq. (1.6) has the same infimum and supremum Hölder exponents $\overline{\theta}$ and $\underline{\theta}$ near $x = 1$ with constants $C_{\overline{\theta}} e^{[\lambda G(1) - \overline{\theta} \pi(1)]t}$ and $C_{\underline{\theta}} e^{[\lambda G(1) - \underline{\theta} \pi(1)]t}$. This allows us to see that the set of measures with well-defined Hölder exponent and Hölder constant near $x = 1$ is closed under our multilevel dynamics.

We can now use this characterization of the initial measure $\mu_0(dx)$ in terms of its infimum and supremum Hölder exponents $\overline{\theta}$ and $\underline{\theta}$ near $x = 1$ to study the long-time behavior of measure-valued solutions $\overline{\mu}_t(dx)$ to Eq. (1.2). To understand the collective outcome achieved by the population, we consider the average group-level replication rate across all the groups in the population, which we denote by

$$\langle G(\cdot)\rangle_{\overline{\mu}_t} := \int_0^1 G(x)\overline{\mu}_t(dx). \tag{1.33}$$

In Theorem 2, we show that the time-average of this collective reproduction rate $\langle G(\cdot)\rangle_{\overline{\mu}_s}$ eventually satisfies bounds in terms of the supremum and infimum Hölder exponents of the initial measure $\mu_0(dx)$.

Theorem 2 *Suppose that $G(x), \pi(x)$ satisfy the assumptions of Theorem 1 and that the initial distribution $\overline{\mu}_0(dx)$ has positive supremum and infimum Hölder exponents $\underline{\theta}$ and $\overline{\theta}$ near $x = 1$. Under these assumptions, the average group-level reproduction rate $\langle G(\cdot)\rangle_{\overline{\mu}_t}$ of solutions $\overline{\mu}_t(dx)$ to Eq. (1.2) satisfy*

$$\limsup_{t\to\infty} \frac{1}{t}\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s}\,ds = \max\left\{G(1) - \frac{\underline{\theta}\pi(1)}{\lambda}, G(0)\right\} \tag{1.34a}$$

$$\liminf_{t\to\infty} \frac{1}{t}\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s}\,ds = \max\left\{G(1) - \frac{\overline{\theta}\pi(1)}{\lambda}, G(0)\right\}. \tag{1.34b}$$

The bounds from Theorem 2 tell us that, in a time-averaged sense, the long-time collective outcome is limited by the group-reproduction rate $G(1)$ of the full-cooperator group. This extends our idea of the shadow of lower-level selection seen in Eq. (1.26), showing that this limitation on the collective reproduction rate holds for any initial measure and making an explicit connection between this limitation and the dynamics of Eq. (1.2).

The proof of Theorem 2 relies on the following implicit representation for the mass of the unnormalized solution $\mu_t(dx)$ of Eq. (1.6), which is derived in Sect. 3.1.1:

$$\mu_t([0,1]) = \int_0^1 \mu_t(dx) = \exp\left(\lambda\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s}\,ds\right). \tag{1.35}$$

This expression can be combined with our estimates for the principal growth rates for $\mu_t([0,1])$ derived in Lemmas 5, 7, and 8 to deduce the bounds of Eq. (1.34). This relationship between the principal growth rates and long-time average group-level replication rate for the population illustrates the importance of collective group-level in the structure of the multilevel replicator dynamics described by Eq. (1.2). Furthermore, the form of the bounds we obtain on the collective outcome highlights the key roles played by all-cooperator group in maintaining cooperation in the population, and how the tension between individual and group incentives hinges upon the interaction between reproduction rates $G(0)$, $G(1)$, and $\pi(1)$ and the infimum and supremum Hölder exponents $\overline{\theta}$ and $\underline{\theta}$ of the initial measure near $x = 1$.

Examining the bounds of Eq. (1.34), we see that $\limsup_{t\to\infty} \frac{1}{t}\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds = G(1) - \lambda^{-1}\underline{\theta}\pi(1) > G(0)$ when $\lambda[G(1) - G(0)] > \underline{\theta}\pi(1)$. This allows us to identify the following threshold value $\lambda^*(\underline{\theta})$

$$\lambda^*(\underline{\theta}) := \frac{\underline{\theta}\pi(1)}{G(1) - G(0)}, \tag{1.36}$$

such that the time-averaged value of $\langle G(\cdot)\rangle_{\overline{\mu}_t}$ exceeds the all-defector reproduction rate $G(0)$ infinitely often if $\lambda > \lambda^*(\underline{\theta})$. If $\lambda > \lambda^*(\underline{\theta})$ and $\underline{\theta} \neq \overline{\theta}$, we see from Eq. (1.34) that $\liminf_{t\to\infty} \frac{1}{t}\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds \neq \limsup_{t\to\infty} \frac{1}{t}\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds$, and therefore the time-average $\langle G(\cdot)\rangle_{\overline{\mu}_t}$ does not converge as $t \to \infty$. Because we assume that $G(x) \in C^1([0, 1])$, $G(x)$ is a valid test-function $v(x)$ for our measure-valued formulation of the multilevel dynamics as described by Eq. (1.5), and thus $\overline{\mu}_t(dx)$ cannot converge to a density steady state when there is disagreement between the supremum and infimum Hölder exponents $\overline{\theta}$ and $\underline{\theta}$ near $x = 1$ for the initial measure $\mu_0(dx)$.

Even though solutions $\overline{\mu}_t(dx)$ do not necessarily converge to any steady state in the long-time limit, we can use the threshold $\lambda^*(\underline{\theta})$ to characterize whether cooperation will survive in the long-run population given any initial measure $\mu_0(dx)$. In Theorem 3, we show that cooperation vanishes from the population when $\lambda < \lambda^*(\underline{\theta})$, while cooperation survives when $\lambda > \lambda^*(\underline{\theta})$. Mathematically, this consists of showing that the population converges to a delta-measure $\delta(x)$ concentrated upon the all-defector group when $\lambda < \lambda^*(\underline{\theta})$, while the fraction of cooperators in the positive exceeds a positive threshold infinitely often when $\lambda > \lambda^*(\underline{\theta})$. This sense in which cooperation survives is called weak persistence (Freedman and Moson 1990), and has often been used to characterize the survival and coexistence of strategies in evolutionary games and related ecological models under individual-level dynamics (Hofbauer 1981; Hofbauer and Sigmund 1998; Hofbauer and Schreiber 2004; Bratus et al. 2017; Freedman and Moson 1990). For the edge case in which $\lambda = \lambda^*(\underline{\theta})$, we know from Theorem 2 that the time-averaged collective outcome $\langle G(\cdot)\rangle_{\overline{\mu}_t}$ converges to that of the all-defector group $G(0)$, and show in Sect. 5.2 that the population will converge to $\delta(x)$ for a more restricted class of group-reproduction functions $G(x)$ and initial measures.

Theorem 3 *Suppose that $G(x), \pi(x)$ satisfy the assumptions of Theorem 2, and that the initial distribution $\overline{\mu}_0(dx)$ has positive supremum and infimum Hölder exponents $\underline{\theta}$ and $\overline{\theta}$ near $x = 1$. If $\lambda[G(1) - G(0)] < \underline{\theta}\pi(1)$, then $\overline{\mu}_t(dx) \rightharpoonup \delta(x)$ as $t \to \infty$. If $\lambda[G(1) - G(0)] > \underline{\theta}\pi(1)$, then the average fraction of cooperators satisfies*

$$\limsup_{t\to\infty} \int_0^1 x\overline{\mu}_t(dx) > 0. \tag{1.37}$$

Our result from Theorem 3 on extinction and weak persistence of cooperation holds for any initial measure $\mu_0(dx)$. Consequently, we can view weak persistence as serving as a more general criterion for identifying the survival of cooperation than convergence to a steady state density in the sense of Theorem 1. In this light, we can see Theorem 3 as providing a natural classification for the long-term behavior of

solutions to Eq. (1.2) for replication rates satisfying the assumptions of the multilevel PD dynamics.

To fully characterize the dynamics of our multilevel replicator equations for within-group replication rates that feature no internal equilibria, we now consider the long-time behavior for the case of the generalized Prisoners' Delight game (in which $G(1) > G(0)$ and $\pi(x) < 0$.) Because both within-groups and between-group competition push to increase cooperation under the PDel dynamics, we show in Proposition 1 that the population concentrates upon the full-cooperator group when there is any between-group competition and any cooperators in the initial population.

Proposition 1 *Suppose that $G(x), \pi(x) \in C^1([0, 1])$, $G(1) > G(0)$, and $\pi(x) < 0$ for $x \in [0, 1]$. If $\mu_0((0, 1]) > 0$ and $\lambda > 0$, then $\overline{\mu}_t(dx) \rightharpoonup \delta(x - 1)$ as $t \to \infty$.*

1.2.4 Results for multiple population dynamics

As we saw for the dynamics of a single interval, the growth of the non-normalized solution can be associated with a principal exponential growth rate that can be associated with the average payoff at steady state. For a given interval j with well-defined Hölder exponent θ^j near $x = 1$,

$$r_j = \begin{cases} \lambda G_j(1) & : \pi_j(x) < 0 \\ \lambda G_j(1) - \theta^j \pi_j(1) & : \pi_j(x) > 0, \ \lambda\left[G_j(1) - G_j(0)\right] - \theta^j \pi_j(1) > 0 \\ \lambda G_j(0) & : \pi_j(x) > 0, \ \lambda\left[G_j(1) - G_j(0)\right] - \theta^j \pi_j(1) < 0 \end{cases}.$$

(1.38)

However, in the case in which $\mu_0^j(dx)$ does not have a well-defined Hölder exponent near $x = 1$, we can possibly only bound the principal growth rate in terms of the infimum and supremum Hölder exponents $\overline{\theta}^j$ and $\underline{\theta}^j$ near $x = 1$. Recalling that $\underline{\theta}^j \le \overline{\theta}_j$, we can see that the principal growth rate $r_j \in [r_j^m, r_j^M]$, where we can our lower bound in terms of $\underline{\theta}^j$ as

$$r_j^m = \begin{cases} \lambda G_j(1) & : \pi_j(x) < 0 \\ \lambda G_j(1) - \overline{\theta}^j \pi_j(1) & : \pi_j(x) > 0, \ \lambda\left[G_j(1) - G_j(0)\right] - \overline{\theta}^j \pi_j(1) > 0 \\ \lambda G_j(0) & : \pi_j(x) > 0, \ \lambda\left[G_j(1) - G_j(0)\right] - \overline{\theta}^j \pi_j(1) < 0 \end{cases}$$

(1.39a)

and our upper bound in terms of $\underline{\theta}^j$ as

$$r_j^M = \begin{cases} \lambda G_j(1) & : \pi_j(x) < 0 \\ \lambda G_j(1) - \underline{\theta}^j \pi_j(1) & : \pi_j(x) > 0, \ \lambda\left[G_j(1) - G_j(0)\right] - \underline{\theta}^j \pi_j(1) > 0 \\ \lambda G_j(0) & : \pi_j(x) > 0, \ \lambda\left[G_j(1) - G_j(0)\right] - \underline{\theta}^j \pi_j(1) < 0 \end{cases}.$$

(1.39b)

In particular, this means that the possible growth rates for two intervals can potentially overlap. In the remainder of the paper, we will focus on the case in which there is an

interval j such that $r_j^m > r_k^M$ for all other intervals k. When such a condition holds, we show in Theorem 4 that the whole population will eventually concentrate upon the subpopulation with the dominant principal growth rate.

Theorem 4 *Suppose that each supopulation $j \in \mathcal{N}$ has reproduction functions $G_j(x), \pi_j(x) \in C^1([0, 1])$ satisfying $G_j(1) > G_j(0)$ and either $\pi_j(0) > 0$ for $x \in [0, 1]$ or $\pi_j(0) < 0$ for $x \in [0, 1]$ and that its initial measure $\mu_0^j(dx)$ have infimum and supremum Hölder exponents $\overline{\theta}^j > 0$ and $\underline{\theta}^j > 0$ near $x = 1$. Suppose there is a subpopulation k such that $r_k^m > r_j^M$ for $j \in \mathcal{N} - \{k\}$. Then, for all such j, we see that $\overline{\mu}_t^j([0, 1]) \to 0$ as $t \to \infty$. Furthermore, if $\mu_0^k(dx)$ has well-defined Hölder exponent θ^k and Hölder constant C_{θ^k} near $x = 1$ that are positive and finite, then $\overline{\mu}_t^k(dx) \to p_{\theta}^{\lambda,k}(x)dx$ as $t \to \infty$ if $\lambda[G_k(1) - G_k(0)] > \theta^k \pi_k(1)$, where $p_{\theta,k}^{\lambda}(x)$ is the steady state density given by Eq. (1.22) with $G(x) := G_k(x)$ and $\pi(x) := \pi_k(x)$.*

1.3 Outline of paper

The remainder of the main paper is structured as follows. In Sect. 2, we calculate the steady state densities for Eq. (1.2) in the PD case and characterize the average collective payoff at steady state. In Sect. 3, we demonstrate useful properties of time-dependent solutions to Eq. (1.2) and present the main estimates for the principal growth rates for solutions along characteristics. In Sect. 4, we use the estimates from Sect. 3 to prove Theorem 1, showing convergence of the population to steady states supporting cooperation when the initial measure has a well-defined Hölder exponent near full-cooperation and between-group competition is sufficiently strong. In Sect. 5, we derive long-time bounds on the time-average of the population's collective reproduction rate (Theorem 2), and we characterize how cooperation can either collapse or persist in the population (Theorem 3), depending on the supremum Hölder exponent of the initial population and the relative intensity of within-group and between-group competition. In Sect. 6, we characterize the long-time behavior of the multipopulation dynamics described by Eq. (1.10), showing that multilevel selection can promote concentration upon the group type favoring the highest long-time collective payoff (Theorem 2). In Sect. 7, we discuss our results and directions for future research.

We also present additional results in the appendix. In Sect. A, we demonstrate well-posedness for solutions to Eq. (1.2) in the measure-valued sense required for our results on long-time behavior. In Sect. 2, we discuss additional properties of the density steady state solutions for the PD multilevel dynamics, characterizing the regularity of the densities and showing the concentration of the steady states upon group compositions achieving the same payoff as an all-cooperator group in the limit of infinite between-group competition. Finally, in Sect. C, we formulate a generalized version of the multilevel HD and SH dynamics in terms of the multipopulation framework discussed in Sect. 6 and then show how we can apply Theorem 4 to understand the long-time behavior of solutions to Eq. (1.2) for the case of HD or SH games.

2 Steady state solutions of multilevel dynamics

In this section, we derive the steady state densities presented in Eq. (1.22) and (1.24) for the PD case of the multilevel dynamics. We show that this family of steady states can be parameterized by their Hölder exponent θ near the full-cooperator equilibrium at $x = 1$, and we use this parametrization to calculate the average group-level reproduction rate of the steady state population. We also characterize additional properties for these steady state densities in Sect. B.

From the results of Lemma 5 and Lemma 12 (provided in Sect. B), we know that the only possible long-time steady states of Eq. (1.2) in the PD case are delta-measures $\delta(x)$ and $\delta(x - 1)$, concentrated at the all-defector and all-cooperator equilibria, and a family densities $f(x)$ which satisfy the ordinary differential equation

$$0 = \frac{\partial}{\partial x} [x(1 - x)\pi(x)f(x)] + \lambda f(x) \left[G(x) - \int_0^1 G(y)f(y)dy \right], \quad (2.1)$$

where $f(x)$ is continuously differentiable and $f(x) > 0$ for $x \in (0, 1)$. For such strong solutions to the steady state problem, we can use separation of variables to see that $f(x)$ must satisfy

$$\frac{f'(x)}{f(x)} = -\frac{[x(1 - x)\pi(x)]'}{x(1 - x)\pi(x)} + \lambda \frac{\langle G(\cdot)\rangle_f - G(x)}{x(1 - x)\pi(x)}. \quad (2.2)$$

We can rewrite the last term of Eq. 2.2 as

$$\lambda \frac{\langle G(\cdot)\rangle_f - G(x)}{x(1 - x)\pi(x)} = \frac{\lambda}{\pi(0)} \left(\frac{\langle G(\cdot)\rangle_f - G(0)}{x} \right)$$
$$+ \frac{\lambda}{\pi(1)} \left(\frac{\langle G(\cdot)\rangle_f - G(1)}{1 - x} \right) + \frac{\lambda C(x)}{\pi(x)} \quad (2.3)$$

where, using the shorthand notation $\tilde{G}(x) := G(x) - \langle G(\cdot)\rangle_f$, we can write $C(x)$ as

$$C(x) = \left(\frac{1}{x(1 - x)} \right) \left[-\tilde{G}(x) + \frac{\tilde{G}(0)}{\pi(0)}(1 - x)\pi(x) + \frac{\tilde{G}(1)}{\pi(1)}x\pi(x) \right]$$
$$= \frac{\tilde{G}(0) - \tilde{G}(x)}{x} + \left(\frac{\tilde{G}(0)}{\pi(0)} \right) \left(\frac{\pi(x) - \pi(0)}{x} \right)$$
$$+ \frac{\tilde{G}(1) - \tilde{G}(x)}{1 - x} + \left(\frac{\tilde{G}(1)}{\pi(1)} \right) \left(\frac{\pi(x) - \pi(1)}{1 - x} \right). \quad (2.4)$$

Using Eqs. (2.3) and (2.3), we can see that steady state densities must satisfy the implicit expression

$$f(x) = \frac{1}{Z_f} x^{\frac{\lambda}{\pi(0)}[\langle G(\cdot)\rangle_f - G(0)] - 1} (1 - x)^{\frac{\lambda}{\pi(1)}[G(1) - \langle G(\cdot)\rangle_f] - 1} \pi(x)^{-1} \exp\left(-\lambda \int_x^1 \frac{C(s)}{\pi(s)} ds\right),$$
(2.5)

where Z_f is a normalizing constant $\int_0^1 f(x)dx = 1$, provided that such a constant exists. Because $G(x)$ and $\pi(x)$ are C^1 functions, we see from Eq. (2.4) that $C(x)$ remains bounded near 0 and 1. This means that the density $f(x)$ from Eq. (2.5) will be integrable on $(0, 1)$ if the average payoff $\langle G(\cdot)\rangle_f$ satisfies the following bounds

$$G(0) < \langle G(\cdot)\rangle_f < G(1).$$
(2.6)

In particular, this tells us that valid steady state densities cannot have a higher group-reproduction rate than the rate of a full-cooperator group $G(1)$, providing a signature of the shadow of lower-level selection. Furthermore, the implicit form of the density provided in Eq. (2.5) highlights the principal contributions of the group-level replication rates of all-cooperator and all-defector groups $G(1)$ and $G(0)$ in determining whether a given distribution of group compositions can be maintained at steady state.

From the implicit relation of Eq. (2.5), we see that there are infinitely many possible steady state densities $f(x)$ for a given relative selection strength λ, one for each value of $\langle G(\cdot)\rangle_f$ satisfying the bounds from Eq. (2.6). Because the Hölder exponent near $x = 1$ is preserved under the dynamics of Eq. (1.2), we will parametrize the measures corresponding to the densities from Eq. (2.5) by their Hölder exponents θ near $x = 1$ to obtain an explicit representation for our family of density steady states.

Noting that $C(x)$ is bounded on $[0, 1]$, we can compute that

$$\lim_{x \to 0} \frac{\int_{1-x}^1 f(y)dy}{x^\Theta} = \lim_{x \to 0} \frac{f(1 - x)}{\Theta x^{\Theta - 1}}$$

$$= \frac{1}{\Theta Z_f} \lim_{x \to 0} \left[x^{\frac{\lambda}{\pi(1)}[G(1) - \langle G(\cdot)\rangle_f] - \Theta} (1 - x)^{\frac{\lambda}{\pi(0)}[\langle G(\cdot)\rangle_f - G(0)] - 1} \pi(1 - x)^{-1} e^{-\lambda \int_{1-x}^1 \frac{C(s)}{\pi(s)} ds} \right]$$

$$= \begin{cases} 0 & : \Theta < \left(\frac{\lambda}{\pi(1)}\right)[G(1) - \langle G(\cdot)\rangle_f] \\ (\Theta \pi(1) Z_f)^{-1} & : \Theta = \left(\frac{\lambda}{\pi(1)}\right)[G(1) - \langle G(\cdot)\rangle_f] \\ \infty & : \Theta > \left(\frac{\lambda}{\pi(1)}\right)[G(1) - \langle G(\cdot)\rangle_f] \end{cases}.$$

Therefore we can deduce from Definition 1 that the Hölder exponent that the Hölder exponent near $x = 1$ for our steady state densities $f(x)$ is given by

$$\theta = \frac{\lambda}{\pi(1)} [G(1) - \langle G(\cdot)\rangle_f].$$
(2.7)

We can then use this expression to obtain the explicit family of steady states $f_\theta^\lambda(x)$ of Eq. (1.22). Furthermore, the average of the group-reproduction function $G(x)$ on

such steady states is given by

$$\langle G(\cdot) \rangle_f = G(1) - \frac{\pi(1)\theta}{\lambda}. \tag{2.8}$$

Using the expression from Eq. (1.25) for the threshold λ^* required for integrability of the density $f_\theta^\lambda(x)$, we can deduce that

$$\langle G(\cdot) \rangle_f = \left(\frac{\lambda^*}{\lambda} \right) G(0) + \left(1 - \frac{\lambda^*}{\lambda} \right) G(1). \tag{2.9}$$

This provides an improvement upon the bounds from Eq. (2.6), showing that $\langle G(\cdot) \rangle_f$ interpolates between $G(0)$ when $\lambda = \lambda^*$ and $G(1)$ as $\lambda \to \infty$.

3 Useful properties of multilevel dynamics

In this section, we provide some useful properties for the measure-valued solutions to Eq. 1.2 and the behavior of the dynamical properties of our model of multilevel selection. In Sect. 3.1, we use the method of characteristics to obtain a representation formula for the time-dependent solutions of the multilevel dynamics. In Sect. 3.2, we use this representation formula and assumptions about the supremum and infimum Hölder exponents of the initial distribution near the full-cooperator equilibrium to derive upper and lower bounds for the principal growth rates of solutions to the linear form of the multilevel dynamics given by Eq. (1.6).

3.1 Representing time-dependent solutions of multilevel dynamics

First, we characterize the impacts of within-group and between-group competition on solutions $\mu_t(dx)$ through properties of the characteristic curves and the solutions along characteristics. To do this, we first consider solutions to the linear problem of Eq. (1.6). We consider the following ordinary differential equation

$$\frac{dx}{dt} = -x(1-x)\pi(x), \quad \frac{dq}{dt} = \lambda G(x)q, \\ x(0) = y, \ q(0) = 1. \tag{3.1}$$

whose solution we denote by

$$x(t) = \phi_t(y), \ q(t) = \psi_t(y). \tag{3.2}$$

We can then represent the solution $\mu_t(dx)$ to the linear multilevel dynamics of Eq. (1.6) by pushing forward the initial measure along characteristic curves, allowing us to obtain

$$\int_0^1 v(x)\mu_t(dx) = \int_0^1 v(\phi_t(y))\psi_t(y)\mu_0(dy), \ w_t(x) = \psi_t(\phi_t^{-1}(x)) \tag{3.3}$$

for all continuous test functions $v(x)$. To further understand how the measure $\mu_t(dx)$ changes in time, we now study in Lemmas 1, 2, and 3 the behavior of the backwards characteristic curves ϕ_t^{-1} (describing within-group competition) and the solutions along characteristics $w_t(x)$ (describing between-group competition).

First we obtain an expression for the backward characteristic curves $\phi_t^{-1}(x)$.

Lemma 1 *Suppose that $G(x), \pi(x)$ satisfy the assumptions of Theorem 1 and let ϕ_t be as in Eq. (3.2). For $0 < x \leq 1$, we have that*

$$\exp(\pi(1)t)(1 - \phi_t^{-1}(x)) = (1 - x) \exp\left(\int_x^{\phi_t^{-1}(x)} \frac{Q(s)ds}{s\pi(s)}\right),$$

$$Q(s) = \frac{\pi(1) - \pi(s)}{1 - s} + \pi(s). \tag{3.4}$$

In particular,

$$\lim_{t\to\infty} \exp(\pi(1)t)(1 - \phi_t^{-1}(x)) = (1 - x) \exp\left(\int_x^1 \frac{Q(s)ds}{s\pi(s)}\right). \tag{3.5}$$

Proof Let $z(t) = 1 - \phi_t^{-1}(x)$. Given $q(t) = \phi_t(x_0)$ satisfies Eq. (3.1), $z(t)$ satisfies the following differential equation:

$$\frac{dz}{dt} = -z(1 - z)\pi(1 - z), \ z(0) = z_0 = 1 - x. \tag{3.6}$$

Separating variables, we find that

$$\int_{z_0}^z \frac{ds}{s(1 - s)\pi(1 - s)} = -t. \tag{3.7}$$

Using a partial fraction expansion allows us to rewrite the left-hand side as

$$\int_{z_0}^z \frac{ds}{s(1 - s)\pi(1 - s)} = \int_{z_0}^z \frac{ds}{\pi(1)s} + \int_{z_0}^z \frac{((\pi(1) - \pi(1 - s))/s + \pi(1 - s))ds}{\pi(1)(1 - s)\pi(1 - s)}, \tag{3.8}$$

which isolates the key contribution of the individual-level advantage $\pi(1)$ for defectors in an all-cooperator group. We can then deduce that

$$\ln\left(\frac{z}{z_0}\right) - \int_{1-z_0}^{1-z} \frac{Q(s)ds}{s\pi(s)} = -\pi(1)t. \tag{3.9}$$

Exponentiating both sides, we obtain Eq. (3.4). We obtain Eq. (3.5) by noting that $\phi_t^{-1}(x) \to 1$ as $t \to \infty$ and that $Q(s)$ is a bounded function for $0 < x \leq s \leq 1$ since $\pi(s)$ is a C^1 function by assumption. □

We next study $w_t(x)$ to describe the effect of between-group competition. Lemma 2 allows us to extract a principal contribution $\lambda G(1)$ to the exponential growth rate for solutions along characteristics to the multilevel dynamics, showing the key importance of the all-cooperator group in promoting long-time cooperation.

Lemma 2 *Suppose that $G(x)$, $\pi(x)$ satisfy the assumptions of Theorem 1 and consider $w_t(x)$ given in Eq. (3.3). For $0 < x \le 1$, we have that*

$$\exp(-\lambda G(1)t)w_t(x) = \exp\left(-\lambda \int_x^{\phi_t^{-1}(x)} \frac{R(s)}{s\pi(s)} ds\right)$$

$$R(s) = \frac{G(1) - G(s)}{1 - s}. \tag{3.10}$$

In particular, we have

$$\lim_{t \to \infty} \exp(-\lambda G(1)t)w_t(x) = \exp\left(-\lambda \int_x^1 \frac{R(s)}{s\pi(s)} ds\right). \tag{3.11}$$

Proof Let $q(t) = w_t(x)$. It is readily seen from (3.1) and (3.2) that $q(t)$ satisfies the differential equation:

$$\frac{dp}{dt} = p(1-p)\pi(p), \ \frac{dq}{dt} = \lambda G(p)q,$$
$$p(0) = p_0 = x, \ q(0) = 1. \tag{3.12}$$

Note here that $p(t) = \phi_t^{-1}(x)$. Let $r(t) = q(t)\exp(-\lambda G(1)t)$. The function $r(t)$ satisfies the equation

$$\frac{dr}{dt} = \lambda(G(p) - G(1))r. \tag{3.13}$$

Thus, using Eq. (3.12), we have

$$\frac{dr}{dp} = \frac{\lambda(G(p) - G(1))}{p(1-p)\pi(p)} r. \tag{3.14}$$

Using $r(0) = q(0) = 1$, we have

$$r(t) = \exp\left(-\lambda \int_{p_0}^{p(t)} \frac{R(s)}{s\pi(s)} ds\right). \tag{3.15}$$

We obtain Eq. (3.11) by noting that $\phi_t^{-1}(x) \to 1$ as $t \to \infty$ and that $G(s)$ is C^1 and thus the integral in Eq. (3.11) is bounded as the upper bound of the integral tends to 1. $\qquad\square$

Next, we obtain an expression for $\partial\phi_t^{-1}(x)/\partial x$.

Lemma 3 *Suppose that $G(x)$, $\pi(x)$ satisfy the assumptions of Theorem 1 and consider $\phi_t^{-1}(x)$ as defined in Eq. (3.2). We have that*

$$\frac{\partial}{\partial x}\phi_t^{-1}(x) = \frac{(\phi_t^{-1}(x)(1 - \phi_t^{-1}(x))\pi(\phi_t^{-1}(x)))}{x(1-x)\pi(x)} \tag{3.16}$$

Proof Let $J(t) = \partial\phi_t^{-1}(x)/\partial x$. Then, using (3.1), we see that J must satisfy the following differential equation:

$$\frac{dp}{dt} = p(1-p)\pi(p), \quad \frac{dJ}{dt} = \big((1-p)\pi(p) - p\pi(p) + p(1-p)\pi'(p)\big)J,$$
$$p(0) = p_0 = x, \quad J(0) = 1, \tag{3.17}$$

where π' is the derivative of π, which exists given our assumption that π is C^1. From the above, we see that:

$$\frac{dJ}{dp} = \left(\frac{1}{p} - \frac{1}{1-p} + \frac{\pi'(p)}{\pi(p)}\right)J. \tag{3.18}$$

Integrating the above differential equation, we obtain the desired formula:

$$J(t) = \frac{p(t)(1-p(t))\pi(p(t))}{p_0(1-p_0)\pi(p_0)}. \tag{3.19}$$

Having characterized solutions $\mu_t(dx)$ to the associated linear problem of Eq. (1.6), in Sect. 3.1.1 we show another way to interpret measure-valued solutions $\overline{\mu}_t(dx)$ to Eq. (1.2) in terms of the linear dynamics. $\qquad\square$

3.1.1 Expressing solutions to multilevel dynamics via exponential normalization relation

In this section, we will show solutions $\overline{\mu}_t(dx)$ to (1.2) can be expressed in terms of the mass of the solutions $\mu_t([0, 1])$ to Eq. (1.6) and on the average group-level reproduction rate across the groups in the population. This average collective outcome across a measure $\mu(dx)$ is defined as

$$\langle G(\cdot)\rangle_\mu := \frac{\int_0^1 G(x)\mu(dx)}{\int_0^1 \mu(dx)}. \tag{3.20}$$

Using the normalization relation of Eq. (1.7) and the fact that $\overline{\mu}_t(dx)$ is a probability measure, we can compute that

$$\langle G(\cdot)\rangle_{\mu_t} = \frac{\int_0^1 G(x)\mu(dx)}{\int_0^1 \mu(dx)} = \int_0^1 G(x)\left(\frac{\mu_t(dx)}{\int_0^1 \mu_t(dy)}\right) = \int_0^1 G(x)\overline{\mu}_t(dx) = \langle G(\cdot)\rangle_{\overline{\mu}_t}, \tag{3.21}$$

so the unnormalized and normalized solutions $\mu_t(dx)$ and $\overline{\mu}_t(dx)$ feature the same average group-level reproduction rates.

By considering the test-function $v(x) = 1$ and using the assumption that $\mu_0(dx)$ is a probability measure, we can calculate the mass of $\mu_t(dx)$ solving Eq. (1.6) satisfies the following ordinary differential equation

$$\frac{d}{dt}\mu_t([0,1]) = \frac{d}{dt}\int_0^1 \mu_t(dx) = \lambda \int_0^1 G(x)\mu_t(dx) \qquad (3.22)$$

Dividing both sides by $\mu_t([0,1])$, we can see from Eq. (3.20) that

$$\frac{1}{\mu_t([0,1])}\frac{d}{dt}\mu_t([0,1]) = \frac{1}{\int_0^1 \mu_t(dx)}\frac{d}{dt}\int_0^1 \mu_t(dx) = \lambda \frac{\int_0^1 G(x)\mu_t(dx)}{\int_0^1 \mu_t(dx)} = \lambda \langle G(\cdot)\rangle_{\mu_t}.$$
$$(3.23)$$

Integrating this differential equation, we see that the mass of the solution $\mu_t(dx)$ to the linear multilevel dynamics is given by

$$\mu_t([0,1]) = \int_0^1 \mu_t(dx) = \exp\left(\lambda \int_0^t \langle G(\cdot)\rangle_{\mu_s} ds\right). \qquad (3.24)$$

Then, applying this to the normalization relation from Eq. (1.7), we can express solutions $\overline{\mu}_t$ to the full multilevel dynamics in terms of $\mu_t(dx)$ by

$$\int_0^1 v(x)\overline{\mu}_t(dx) = \frac{\int_0^1 v(x)\mu_t(dx)}{\mu_t([0,1])} = \exp\left(-\lambda \int_0^t \langle G(\cdot)\rangle_{\mu_s} ds\right)\int_0^1 v(x)\mu_t(dx). \qquad (3.25)$$

This representation of the solution $\overline{\mu}_t(dx)$ for the full nonlinear multilevel dynamics of Eq. (1.2) in terms of the solutions $\mu_t(dx)$ of the linear dynamics is particularly useful for understanding the long-time dynamics of the collective outcome $\int_0^t \langle G(\cdot)\rangle_{\mu_s} ds$. In particular, we will use this and the fact that solutions $\overline{\mu}_t(dx)$ are normalized to obtain the long-time bounds on time-averaged collective fitness presented in Theorem 2. Noting that $\overline{\mu}_t([0,1]) = 1$ and that $\langle G(\cdot)\rangle_{\mu_s} = \langle G(\cdot)\rangle_{\overline{\mu}_s}$ from Eq. (3.21), we can apply the test-function $v(x) = 1$ in Eq. (3.25) to see that

$$1 = \overline{\mu}_t([0,1]) = \exp\left(-\lambda \int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds\right)\int_0^1 \mu_t(dx). \qquad (3.26)$$

Rearranging this equation allows us to deduce Eq. (1.35), which highlights the connection between principal growth rates for $\mu_t(dx)$ and the average group-level reproduction rate $\langle G(\cdot)\rangle_{\overline{\mu}_t}$.

3.2 Upper and lower bounds for principal growth rates of solutions

Now we can use our push-forward representation to estimate the growth rate for the tails of solutions $\mu_t(dx)$ to the linear multilevel dynamics of Eq. (1.6). In Lemma 4, we use the formula from $w_t(x)$ from Lemma 2 to find upper and lower bounds for the growth rates of solutions along characteristics described by $\psi_t(x) = w_t(\phi_t(x))$. Next, in Lemma 5, we apply Lemma 4 to obtain bounds on the growth rate of the mass $\mu_t([a, 1])$ for any $a > 0$. These bounds are expressed in terms of the group-reproduction rate of the full-cooperator group $G(1)$, the individual-level advantage of defectors in an otherwise full-cooperator group $\pi(1)$, and the infimum and supremum Hölder exponents $\overline{\theta}$ and $\underline{\theta}$ of the initial measure $\mu_0(dx)$ near $x = 1$. The form of these rates highlights the conflict between the individual-level incentive to defect and the collective incentive to achieve full-cooperation.

Lemma 4 *Suppose that $G(x)$ and $\pi(x)$ satisfy the assumptions of Theorem 1. Then there exist positive constants $\underline{M}, \overline{M} < \infty$ such that, for all $x \in (0, 1)$,*

$$\underline{M}e^{\lambda G(1)t}\phi_t(x)^{\lambda\pi(0)^{-1}[G(1)-G(0)]} \leq \psi_t(x) \leq \overline{M}e^{\lambda G(1)t}. \tag{3.27}$$

Proof For our upper bound, we can use Eq. (3.10) to estimate that

$$\psi_t(x) = w_t(\phi_t(x)) = e^{\lambda G(1)t} \exp\left(\lambda \int_{\phi_t(x)}^{x} \left[\frac{G(q) - G(1)}{q(1 - q)\pi(q)}\right] dq\right)$$

$$\leq e^{\lambda G(1)t} \underbrace{\exp\left(\lambda \int_0^1 \frac{[G(q) - G(1)]_+}{q(1 - q)\pi(q)} dq\right)}_{:=\overline{M}}, \tag{3.28}$$

where we have ignored non-positive contributions to the integral using the notation

$$[G(q) - G_a(1)]_+ = \begin{cases} G(q) - G(1) : G(q) > G(1) \\ 0 \qquad\quad : G(q) \leq G(1) \end{cases}. \tag{3.29}$$

Noting that the integrand in the last term of Eq. (3.28) is bounded as $q \to 1$ because $G(q)$ is C^1 and bounded as $q \to 0$ because $G(0) < G(1)$, we can deduce that $\overline{M} < \infty$.

To find a corresponding lower bound on $\psi_t(x)$, we can first use Eq. (3.10) to rewrite the quantity $-R(q)/[q\pi(q)]$ from as

$$-\frac{R(q)}{q\pi(q)} = \frac{G(0) - G(1)}{\pi(0)q} + \frac{1}{(1 - q)\pi(q)}$$
$$\times \left[\frac{G(q) - G(0)}{q} + \frac{G(0) - G(1)}{\pi(0)}\left(\pi(q) - \frac{\pi(q) - \pi(0)}{q}\right)\right], \tag{3.30}$$

where we can check that the second term on the righthand side is bounded below on $[0, 1]$ because $G(x)$ and $\pi(x)$ are in $C^1([0, 1])$. This means that there exists $N > -\infty$

such that

$$-\int_{\phi_t(x)}^x \frac{R(q)}{q\pi(q)} dq \geq N + \left[\frac{G(1)-G(0)}{\pi(0)}\right] \log(\phi_t(x)), \tag{3.31}$$

and then, letting $\underline{M} := e^{\lambda N}$, we can use Eq. (3.10) to estimate that

$$\psi_t(x) = e^{\lambda G(1)t} \exp\left(\lambda \int_{-\phi_t(x)}^x \frac{R(q)}{q\pi(q)} dq\right) \geq \underline{M} e^{\lambda G(1)t} \phi_t(x)^{\lambda\pi(0)^{-1}[G(1)-G(0)]}. \tag{3.32}$$

Lemma 5 *Suppose that $G(x)$ and $\pi(x)$ satisfy the assumptions of Theorem 1. Consider $a \in (0,1)$ and $\mu_0(dx)$ with supremum and infimum Hölder exponents near $x = 1$ satisfying $0 < \underline{\theta} \leq \overline{\theta}$. For any $\Theta < \underline{\theta}$, there exists $B(\Theta) < \infty$ such that*

$$\mu_t([a,1]) \leq B(\Theta)e^{[\lambda G(1)-\Theta\pi(1)]t}, \tag{3.33a}$$

while for any $\Theta > \underline{\theta}$, there exists $b(\Theta) > 0$ and a sequence of times $\{t_n\}_{n\in\mathbb{N}}$ tending to infinity such that

$$\mu_{t_n}([a,1]) \geq b(\Theta)e^{[\lambda G(1)-\Theta\pi(1)]t_n}. \tag{3.33b}$$

Similarly, for $\Theta > \overline{\theta}$, there exists $d(\Theta) > 0$ such that

$$\mu_t([a,1]) \geq d(\Theta)e^{[\lambda G(1)-\Theta\pi(1)]t}, \tag{3.33c}$$

and for $\Theta < \overline{\theta}$, there exists $D(\Theta) < \infty$ and a sequence of times $\{s_n\}_{n\in\mathbb{N}}$ tending to infinity such that

$$\mu_{s_n}([a,1]) \leq D(\Theta)e^{[\lambda G(1)-\Theta\pi(1)]s_n}. \tag{3.33d}$$

Furthermore, when the Hölder constants $C_{\underline{\theta}}$ or $C_{\overline{\theta}}$ of $\mu_0(dx)$ near $x = 1$ are positive and finite, we can obtain versions of each of the bounds from Eq. (3.33) with $\Theta = \underline{\theta}$ or $\Theta = \overline{\Theta}$, respectively.

Proof Using the upper bound on $\psi_t(x)$ from Lemma 4, we know that there exists an $\overline{M} < \infty$ such that

$$\int_a^1 \mu_t(dx) = \int_{\phi_t^{-1}(a)}^1 \psi_t(x)\mu_0(dx) \leq \overline{M} e^{\lambda G(1)t} \int_{\phi_t^{-1}(a)}^1 \mu_0(dx) = \overline{M} e^{\lambda G(1)t} \mu_0\left(\left[\phi_t^{-1}(a), 1\right]\right). \tag{3.34}$$

Using the lower bound from Lemma 4, we know that there is an $\underline{M} > 0$ such that

$$\int_a^1 \mu_t(dx) = \int_{\phi_t^{-1}(a)}^1 \psi_t(x)\mu_0(dx) \geq \underline{M} e^{\lambda G(1)t} \int_{\phi_t^{-1}(a)}^1 \phi_t(x)^{\lambda\pi(0)^{-1}[G(1)-G(0)]}\mu_0(dx). \tag{3.35}$$

Using our assumptions that $G(0) < G(1)$ and $\pi(0) > 0$, as well as the fact that $\phi_t(x) \in [a, 1]$ for $x \in [\phi_t^{-1}(a), 1]$, we can further estimate that

$$\int_a^1 \mu_t(dx) \geq \underline{M} a^{\lambda \pi(0)^{-1}[G(1)-G(0)]} e^{\lambda G(1)t} \mu_0\left(\left[\phi_t^{-1}(a), 1\right]\right). \qquad (3.36)$$

Our next step is to obtain bounds on $\mu_0\left(\left[\phi_t^{-1}(a), 1\right]\right)$. We can use the formula from Lemma 1 for $1 - \phi_t^{-1}(x)$ to see that

$$\begin{aligned}
\mu_0\left(\left[\phi_t^{-1}(a), 1\right]\right) &= \mu_0\left(\left[1 - \left(1 - \phi_t^{-1}(a)\right), 1\right]\right) \\
&= \mu_0\left(\left[1 - e^{-\pi(1)t}(1-a)\exp\left(\int_a^{\phi_t^{-1}(a)} \frac{Q(q)}{q\pi(q)}dq\right), 1\right]\right)
\end{aligned} \qquad (3.37)$$

From our assumption that $\mu_0(dx)$ has supremum Hölder exponent $\underline{\theta} > 0$, we know that $\limsup_{x \to 0} x^{-\Theta} \mu_0([1-x, 1]) = 0$ for $\Theta < \underline{\theta}$. For such Θ, there is therefore a constant C such that

$$\mu_0\left(\left[\phi_t^{-1}(a), 1\right]\right) \leq C e^{-\Theta \pi(1)t}(1-a)^\Theta \exp\left(\Theta \int_a^{\phi_t^{-1}(a)} \frac{Q(q)}{q\pi(q)}dq\right). \qquad (3.38)$$

Combining this with our estimate from Eq. (3.34) and the expression for $Q(q)$ from Eq. (3.4), we can see that

$$\mu_t([a, 1]) \leq \underbrace{C\overline{M}(1-a)^\Theta \exp\left(\Theta \int_a^1 \frac{[\pi(1) - q\pi(q)]_+}{q\pi(q)}dq\right)}_{:=B(\Theta)} e^{[\lambda G(1) - \Theta \pi(1)]t}. \qquad (3.39)$$

Similarly, we note that $\limsup_{x \to 0} x^{-\Theta} \mu_0([1-x, 1]) > 0$ for $\Theta > \underline{\theta}$. Combining this with the expression from Eqs. (3.36) and (3.37), we see that there is $C > 0$ and a sequence of times $\{t_n\}$ tending to infinity such that

$$\mu_{t_n}([a, 1]) \geq \underbrace{C\underline{M}(1-a)^\Theta \exp\left(\Theta \int_a^1 \frac{[\pi(1) - q\pi(q)]_-}{q\pi(q)}dq\right)}_{:=b(\Theta)} e^{[\lambda G(1) - \Theta \pi(1)]t}. \qquad (3.40)$$

Further, we know that $b(\Theta) > 0$ and $B(\Theta) < \infty$ because $\pi(x)$ is C^1. We can obtain the desired bounds depending on the infimum Hölder exponent $\overline{\theta}$ of $\mu_0(dx)$ near $x = 1$ with an analogous approach. $\qquad \square$

We can also derive approximate lower bounds for the growth rate of $\mu_t(dx)$ in terms of the growth rate near the equilibrium of the within-group dynamics.

Lemma 6 *Suppose that $G(x), \pi(x)$ satisfy the assumptions of Theorem 1 and that $\mu_0(dx)$ has positive Hölder exponent $\underline{\theta}$ near $x = 1$ and consider the quantity $\check{G}_w := \min_{x \in [0,a]} G(x)$. For w sufficiently close to 0, there exists a constant $A_w > 0$ such that*

$$\mu_t([0,1]) \geq A_w e^{\lambda \check{G}_w t}. \tag{3.41}$$

Proof Due to our assumption that $\mu_0(dx)$ has supremum Hölder exponent $\underline{\theta} > 0$, there is $w' < 1$ such that $\mu_0([0, w']) > 0$. Because $\phi_t(x)$ decreases in time and satisfies $\phi_t(x) \to 0$ for $x \in (0,1)$ as $t \to \infty$, we have that, for any $w < w'$, there is a T_w such that $\phi_{T_w}(w') = w$ and $\phi_t(w') \leq w$ for $t > T_w$. Now we can estimate the integral of the group-reproduction function along characteristic curves by

$$\int_0^t G(\phi_s(x))ds = \int_{T_w}^t G(\phi_s(x))ds + \int_0^{T_w} G(\phi_s(x))ds \geq \check{G}_w t + \left(\check{G}_{w'} - \check{G}_w\right)T_w. \tag{3.42}$$

We can then apply this estimate and the fact that $\mu_0(dx)$ is a probability measure to deduce that

$$
\begin{aligned}
\int_0^1 \mu_t(dx) &= \int_0^1 \exp\left(\lambda \int_0^t G(\phi_s(x))ds\right)\mu_0(dx) \\
&\geq \int_0^{w'} \exp\left(\lambda \int_0^t G(\phi_s(x))ds\right)\mu_0(dx) \\
&\geq e^{\lambda\left(\check{G}_{w'} - \check{G}_w\right)T_w} e^{\lambda\check{G}_w t} \int_0^{w'} \mu_0(dx) \geq e^{\lambda\left(\check{G}_{w'} - \check{G}_w\right)T_w} e^{\lambda\check{G}_w t}.
\end{aligned} \tag{3.43}
$$

$\square$

4 Convergence to steady state population for initial measures with well-defined Hölder exponents

In this section, we consider the long-time behavior of solutions $\overline{\mu}_t(dx)$ to Eq. (1.2) for initial conditions with well-defined Hölder exponents and Hölder constants. In Sect. 4.1, we prove Theorem 1, demonstration weak convergence of the population to steady state densities for sufficiently strong between-group competition and initial measures with well-defined nonzero, finite Hölder exponents and constants. In Sect. 4.2, we state and prove Proposition 2, which tells us that a population with an initial partial delta-peak at full-cooperation will fix full-cooperation in the group-structure population in the long-time limit.

4.1 Weak convergence for measure-valued initial population

Before presenting the proof of Theorem 1, we will first rewrite our expression for the steady state densities from Eq. (1.22) into a form that is most compatible with the expressions related to the method of characteristics derived in Lemmas 1, 2, and 3 . We start by considering the following expression for steady state density solutions to Eq. (1.2)

$$f_\theta^\lambda(x) = (1-x)^{\theta-1} \frac{\pi(1)}{x\pi(x)} \exp\left(\int_x^1 \frac{M_\theta(s)ds}{s\pi(s)}\right), \quad M_\theta(s) = -\lambda R(s) + \theta Q(s). \tag{4.1}$$

Combining the steady state expressions from Eqs. (1.22) and (4.1), we can write $M_\theta(s)$ using the following decomposition

$$M_\theta(s) = \frac{\nu_\theta}{s} + \frac{C(s)}{\pi(s)} \tag{4.2}$$

where ν and $C(s)$ are given by Eqs. (1.23) and (1.24), respectively.

Now we will study the convergence of $\mu_t(dx)$ to these density steady states. We will start by the integrating the righthand side of Eq. (3.3) by parts to see that

$$\int_0^1 v(\phi_t(y))\psi_t(y)\mu_0(dy) = v(1)\psi_t(1)F(1^-) - v(0)\psi_t(0)F(0^+)$$
$$- \int_0^1 \frac{\partial}{\partial y}(v(\phi_t(y))\psi_t(y))F(y)dy, \quad F(y) = \mu_0([0, y]). \tag{4.3}$$

In the above, we used the fact that $\phi_t(0) = 0$ and $\phi_t(1) = 1$. For initial measures $\mu_0(dx)$ with positive supremum Hölder exponent $\underline{\theta}$, $F(1^-) = 1$ and therefore

$$\int_0^1 v(\phi_t(y))\psi_t(y)\mu_0(dy) = v(0)\psi_t(0)(1 - F(0^+))$$
$$+ \int_0^1 \frac{\partial}{\partial y}(v(\phi_t(y))\psi_t(y))(1 - F(y))dy. \tag{4.4}$$

In Lemma 7, we estimate the growth rate of both terms in Eq. (4.4), and show that the time-dependent family of integrands in the second term is bounded by an integrable function that is independent of time. Using Lemma 7, we can then apply the Dominated Convergence Theorem to help prove convergence to steady state densities in Theorem 1.

Lemma 7 *Suppose that $G(x), \pi(x)$ satisfy the assumptions of Theorem 1 and that $\mu_0(dx)$ has supremum Hölder exponent $\underline{\theta}$ near $x = 1$ that is nonzero and finite. If $\nu_\theta = \lambda[G(1) - G(0)] - \underline{\theta}\pi(1) > 0$, then, for any $\Theta < \underline{\theta}$, there exists a constant*

$C^v(\Theta) < \infty$ such that

$$\left| \frac{\partial}{\partial y} \left(v(\phi_t(y))\psi_t(y) \right) (1 - F(y)) \right| \le C^v(\Theta) e^{[\lambda G(1) - \Theta \pi(1)]t} d(x)$$

$$d(x) := \begin{cases} (1-x)^{\Theta-1} x^{v-1} & \text{if } v < 1, \\ (1-x)^{\Theta-1} & \text{if } v \ge 1. \end{cases} \tag{4.5}$$

Furthermore, we can use Eq. (4.4) to see that there are positive constants C_1^v and $C_2^v(\Theta)$ such that

$$\int_0^1 v(\phi_t(y))\psi_t(y)\mu_0(dx) \le C_1^v e^{\lambda G(0)t} + C_2^v(\Theta) e^{[\lambda G(1) - \Theta \pi(1)]t}. \tag{4.6}$$

If, in addition, the supremum Hölder constant of $\mu_0(dx)$ is finite and nonzero, then we can obtains bounds analogous to those of Eqs. (4.5) and (4.6) for $\Theta = \theta$.

Proof We start with the integral in Eq. (4.4). Changing variables to $x = \phi_t(y)$, we have:

$$\int_0^1 \frac{\partial}{\partial y} \left(v(\phi_t(y))\psi_t(y) \right) (1 - F(y)) dy = \int_0^1 \frac{\partial}{\partial x} \left((v(x)w_t(x)) (1 - F(\phi_t^{-1}(x))) \right) dx \tag{4.7}$$

Our goal is to estimate the following quantity

$$\int_0^1 I_1 I_2 dx := e^{-(\lambda G(1) - \theta \pi(1))t} \int_0^1 \frac{\partial}{\partial x} \left((v(x)w_t(x)) (1 - F(\phi_t^{-1}(x))) \right) dx,$$

$$I_1 = e^{-\lambda G(1)t} \frac{\partial}{\partial x} \left((v(x)w_t(x)) \right), \quad I_2 = e^{\theta \pi(1)t} (1 - F(\phi_t^{-1}(x))). \tag{4.8}$$

Let us first consider I_1. Using Eq. (3.10), we have that

$$I_1 = \left(\frac{\partial v}{\partial x} - \lambda K(x)v(x) \right) \exp\left(-\lambda \int_x^{\phi_t^{-1}(x)} \frac{R(s)}{s\pi(s)} ds \right),$$

$$K(x) = \frac{R(\phi_t^{-1}(x))}{\phi_t^{-1}(x)\pi(\phi_t^{-1}(x))} \frac{\partial}{\partial x} \phi_t^{-1}(x) - \frac{R(x)}{x\pi(x)}. \tag{4.9}$$

Using Eq. (3.16) and the definition of R in Eq. (3.10), we have:

$$K(x) = \frac{G(x) - G(\phi_t^{-1}(x))}{x(1-x)\pi(x)}. \tag{4.10}$$

Using the fact that G is C^1 and that $\phi_t^{-1}(x) \geq x$, we have that

$$\left| G(x) - G(\phi_t^{-1}(x)) \right| = \left| \int_x^{\phi_t^{-1}(x)} G'(z)dz \right| \leq \left\| G' \right\|_\infty (\phi_t^{-1}(x) - x) \quad (4.11)$$

Denoting $\pi_{\min} := \min_{0 \leq s \leq 1} \pi(s) > 0$, we may thus estimate $K(x)$ by

$$|K(x)| \leq \frac{\left\| G' \right\|_\infty (\phi_t^{-1}(x) - x)}{x(1-x)\pi(x)} \leq \frac{\left\| G' \right\|_\infty \phi_t^{-1}(x)}{\pi_{\min} x(1-x)}, \quad (4.12)$$

and therefore deduce that

$$\left| \frac{\partial v}{\partial x} - \lambda K(x)v(x) \right| \leq \left\| \frac{\partial v}{\partial x} \right\|_\infty + \frac{\lambda \left\| G' \right\|_\infty \left\| v \right\|_\infty \phi_t^{-1}(x)}{\pi_{\min} x(1-x)}$$
$$\leq \underbrace{\left(\left\| \frac{\partial v}{\partial x} \right\|_\infty + \frac{\lambda \left\| G' \right\|_\infty \left\| v \right\|_\infty}{\pi_{\min}} \right)}_{:=C_1} \frac{\phi_t^{-1}(x)}{x(1-x)}. \quad (4.13)$$

Returning to Eq. (4.9), we see that I_1 can be estimated as

$$|I_1| \leq C_1 \frac{\phi_t^{-1}(x)}{x(1-x)} \exp\left(-\lambda \int_x^{\phi_t^{-1}(x)} \frac{R(s)}{s\pi(s)}ds \right). \quad (4.14)$$

We now estimate I_2 in Eq. (4.8). Because $\mu_0(dx)$ has supremum Hölder exponent $\underline{\theta}$ near $x = 1$ and $1 - F(\phi_t^{-1}(x)) = \mu_0\left(\left(\phi_t^{-1}(x), 1 \right] \right)$, we can use Eq. (3.37) to see that, for any $\Theta < \underline{\theta}$, there is a constant $C_2(\Theta) < \infty$ such that

$$|I_2| \leq \exp(\Theta\pi(1)t)(1 - \phi_t^{-1}(x))^\Theta \leq C_2(\Theta)(1-x)^\Theta \exp\left(\Theta \int_x^{\phi_t^{-1}(x)} \frac{Q(s)}{s\pi(s)}ds \right). \quad (4.15)$$

Combining Eqs. (4.14) and (4.15), we have that

$$|I_1 I_2| \leq C_1 C_2(\Theta)(1-x)^{\Theta-1} \frac{\phi_t^{-1}(x)}{x} I_3, \ \ I_3 = \exp\left(\int_x^{\phi_t^{-1}(x)} \frac{M_\Theta(s)}{s\pi(s)}ds \right), \quad (4.16)$$

where $M_\Theta(s)$ is defined as in Eq. (4.1). To estimate I_3, we see from Eq. (4.2) that

$$\int_x^{\phi_t^{-1}(x)} \frac{M_\Theta(s)}{s\pi(s)}ds = -\nu_\Theta \ln\left(\frac{\phi_t^{-1}(x)}{x} \right) + \int_x^{\phi_t^{-1}(x)} \frac{C(s)}{\pi(s)}ds, \quad (4.17)$$

and we can use the fact that $C(s)$ is bounded on $[0, 1]$ to see that

$$|I_3| \leq C_3 \left(\frac{x}{\phi_t^{-1}(x)} \right)^{\nu}, \quad C_3 = \exp \left(\int_0^1 \frac{C(s)}{\pi(s)} ds \right) < \infty. \quad (4.18)$$

Combining this with Eq. (4.16), we see that

$$|I_1 I_2| \leq C^{\nu}(\Theta)(1 - x)^{\Theta - 1} \left(\frac{x}{\phi_t^{-1}(x)} \right)^{\nu_\Theta - 1}, \quad C^{\nu}(\Theta) = C_1 C_2(\Theta) C_3. \quad (4.19)$$

Since $\phi_t^{-1}(x) \geq x$, we can use the definitions of $d(x)$, I_1, and I_2 from Eqs. (4.5) and (4.8) to see that $|I_1 I_2| \leq d(x) e^{[\lambda G(1) - \Theta \pi(1)]t}$, where we note that $d(x)$ is integrable because $\Theta > 0$ and $\nu_\Theta > \nu_\theta > 0$ for any $\Theta < \theta$. We can further use this estimate of $|I_1 I_2|$,., the fact that $\psi_t(0) = e^{\lambda G(0)t}$, and the choice of constants $C_1^{\nu} = |v(0)|(1 - F(0^+))$ and $C_2^{\nu}(\Theta) = C^{\nu}(\Theta) \int_0^1 d(x) dx$ to obtain the estimate of Eq. (4.6). $\qquad \square$

Proof of Theorem 1 We first prove Eq. (1.27) when $v(x)$ is a C^1 function. We will evaluate the following quantity as $t \to \infty$

$$e^{-(\lambda G(1) - \theta \pi(1))t} \int_0^1 v(x) \mu_t(dx) = e^{-(\lambda G(1) - \theta \pi(1))t} \int_0^1 v(\phi_t(y)) \psi_t(y) \mu_0(dx), \quad (4.20)$$

which we can express through Eq. (4.4). From the boundary term of Eq. (4.4), we can use the fact that $\psi_t(0) = e^{\lambda G(0)t}$ and our assumption that $\lambda [G(1) - G(0)] > \theta \pi(1)$ to see that

$$\lim_{t \to \infty} e^{-(\lambda G(1) - \theta \pi(1))t} v(0) \psi_t(0)(1 - F(0^+))$$
$$= \lim_{t \to \infty} e^{-(\lambda(G(1) - G(0)) - \theta \pi(1))t} v(0)(1 - F(0^+)) = 0, \quad (4.21)$$

where the last equality follows from the fact that $\mu_0(dx)$ has Hölder exponent θ near $x = 1$. Recalling that $\theta > 0$ and $\nu_\theta > 0$ by assumption, we can apply Lemma 7 to see that the integrand on the righthand side of Eq. (4.4) is bounded by an integrable function independent of t. We may thus use the Dominated Convergence Theorem to pass to the limit as $t \to \infty$ in the integral in Eq. (4.8).

Using Eqs. (3.11) and (4.10), we find that

$$\lim_{t \to \infty} K(x) = \lim_{t \to \infty} \frac{G(x) - G(\phi_t^{-1}(x))}{x(1 - x)\pi(x)} = -\frac{1}{x\pi(x)} \left(\frac{G(1) - G(x)}{1 - x} \right) = -\frac{R(x)}{x\pi(x)}.$$

Combining this with Eq. (4.9), we can compute that

$$\lim_{t \to \infty} I_1 = \left(\frac{\partial v}{\partial x} + \frac{\lambda R(x) v(x)}{x\pi(x)} \right) \exp \left(-\lambda \int_x^1 \frac{R(s)}{s\pi(s)} ds \right)$$

$$= \frac{\partial}{\partial x} \left(v(x) \exp \left(-\lambda \int_x^1 \frac{R(s)}{s\pi(s)} ds \right) \right). \tag{4.22}$$

Furthermore, can use Eq. (3.5) and the fact that $\mu_0(dx)$ has Hölder exponent θ with constant C_θ near $x = 1$ to see that

$$\lim_{t \to \infty} I_2 = \lim_{t \to \infty} \frac{1 - F(\phi_t^{-1}(x))}{(1 - \phi_t^{-1}(x))^\theta} \left(\exp(\pi(1)t)(1 - \phi_t^{-1}(x)) \right)^\theta$$

$$= C_\theta (1 - x)^\theta \exp \left(\theta \int_x^1 \frac{Q(s)ds}{s\pi(s)} \right). \tag{4.23}$$

Therefore we deduce that

$$\lim_{t \to \infty} \int_0^1 I_1 I_2 dx = \int_0^1 \left[\frac{\partial}{\partial x} \left(v(x) \exp \left(-\lambda \int_x^1 \frac{R(s)}{s\pi(s)} ds \right) \right) \right.$$

$$\left. C_\theta (1 - x)^\theta \exp \left(\theta \int_x^1 \frac{Q(s)ds}{s\pi(s)} \right) \right] dx$$

$$= \left(v(x) C_\theta (1 - x)^\theta \exp \left(\int_x^1 \frac{M_\theta(s)ds}{s\pi(s)} \right) \right) \Big|_0^1$$

$$- \int_0^1 \left[\left(v(x) \exp \left(-\lambda \int_x^1 \frac{R(s)}{s\pi(s)} ds \right) \right) \right.$$

$$\left. \frac{\partial}{\partial x} \left(C_\theta (1 - x)^\theta \exp \left(\theta \int_x^1 \frac{Q(s)ds}{s\pi(s)} \right) \right) \right] dx, \tag{4.24}$$

where we integrated by parts in the second equality. Using Eq. (4.2) and the fact that $v_\theta > 0$, we see that the boundary term vanishes. After some simplifications, we see that

$$\lim_{t \to \infty} \int_0^1 I_1 I_2 dx = C_\theta \theta \int_0^1 v(x) f_\theta^\lambda(x) dx. \tag{4.25}$$

Combining this with Eq. (4.21), we see that:

$$\lim_{t \to \infty} e^{-[\lambda G(1) - \theta \pi(1)]t} \int_0^1 v(x) \mu_t(dx) = C_\theta \theta \int_0^1 v(x) f_\theta^\lambda(x) dx. \tag{4.26}$$

Using the test-function $v(x) = 1$, we further have that

$$\lim_{t \to \infty} e^{-[\lambda G(1) - \theta \pi(1)]t} \int_0^1 \mu_t(dx) = C_\theta \theta \int_0^1 f_\theta^\lambda(x) dx. \tag{4.27}$$

Using the normalization relation from Eq. (1.7), we may now compute the limit in Eq. (1.27).

$$\lim_{t\to\infty}\int_0^1 v(x)\overline{\mu}_t(dx) = \lim_{t\to\infty}\frac{e^{-[\lambda G(1)-\theta\pi(1)]t}\int_0^1 v(x)\mu_t(dx)}{e^{-[\lambda G(1)-\theta\pi(1)]t}\int_0^1 \mu_t(dx)} = \frac{\int_0^1 v(x)f_\theta^\lambda(x)dx}{\int_0^1 f_\theta^\lambda(x)dx}$$

$$= \int_0^1 v(x)p_\theta^\lambda(x)dx, \tag{4.28}$$

where we used Eqs. (4.26) and (4.27) in the second equality. We have thus established (1.27) when $v(x)$ is C^1. To see that Eq. (1.27) is valid for merely continuous $v(x)$, we may use a standard approximation argument. Note that we may approximate $v(x)$ arbitrarily closely by a C^1 function $w(x)$ in the sup norm:

$$\text{for any } \epsilon > 0, \text{ there exists } w(x) \in C^1 \text{ such that } \|v - w\|_\infty \le \epsilon. \tag{4.29}$$

Using this and the fact that $\overline{\mu}_t(dx)$ and $p_\theta^\lambda(x)dx$ are probability measures, we see that

$$\left|\int_0^1 v\overline{\mu}_t(dx) - \int_0^1 vp_\theta^\lambda(x)dx\right| \le \int_0^1 |v - w|\,\overline{\mu}_t(dx) + \left|\int_0^1 w\overline{\mu}_t(dx) - \int_0^1 wp_\theta^\lambda(x)dx\right|$$

$$+ \int_0^1 |v - w|\,p_\theta^\lambda(x)dx \le 2\epsilon + \left|\int_0^1 w\overline{\mu}_t(dx) - \int_0^1 wp_\theta^\lambda(x)dx\right|, \tag{4.30}$$

As $t \to \infty$, the last expression tends to 0 since $w(x)$ is C^1. Since ϵ is arbitrary, we obtain the desired conclusion. □

4.2 Convergence to delta-function at full-cooperation

Next, we consider the case in which a positive fraction of groups in the initial population are concentrated at the all-cooperator composition. We show in Proposition 2 that if the initial population contains a positive probability of full-cooperator groups, then the whole population will concentrate upon full-cooperation in the long-time limit.

Proposition 2 *Suppose that $G(x), \pi(x)$ satisfy the assumptions of Theorem 1 and that the initial population can be written as*

$$\mu_0(dx) = a_1\delta(1 - x) + (1 - a_1)\rho_0(dx) \tag{4.31}$$

for $a_1 > 0$ and $\rho_0(dx)$ a probability measure with a supremum Hölder exponent of $\underline{\theta} > 0$ near $x = 1$. Then the solution to Eq. (1.2) $\overline{\mu}_t(dx)$ converges weakly to the delta-function $\delta(1 - x)$ concentrated at full-cooperation as $t \to \infty$.

Proof Using the push-forward representation of $\mu_t(dx)$, we find that

$$\int_0^1 v(x)\overline{\mu}_t(dx) = \frac{e^{-\lambda G(1)t}\int_0^1 v(x)\mu_t(dx)}{e^{-\lambda G(1)t}\int_0^1 \mu_t(dx)}$$

$$= \frac{a_1 v(1) + (1 - a_1) e^{-\lambda G(1)t} \int_0^1 v(\psi_t(x))\psi_t(x)\rho_0(dx)}{a_1 + (1 - a_1) e^{-\lambda G(1)t} \int_0^1 \psi_t(x)\rho_0(dx)}. \quad (4.32)$$

Using Lemma 7, we know that there exist positive constants $C_1^v, C_2^v < \infty$ such that for any $\tilde{\theta} \in (0, \underline{\theta})$

$$(1 - a_1) e^{-\lambda G(1)t} \left| \int_0^1 v(\psi_t(x))\psi_t(x)\rho_0(dx) \right|$$
$$\leq (1 - a_1) \left[C_1^v e^{\lambda[G(0)-G(1)]t} + C_2^v e^{-\tilde{\theta}\pi(1)t} \right]$$

Because $G(0) < G(1)$ and $\tilde{\theta} > 0$, we can further see that

$$(1 - a_1) e^{-\lambda G(1)t} \left| \int_0^1 v(\psi_t(x))\psi_t(x)\rho_0(dx) \right| \to 0 \text{ as } t \to \infty.$$

Applying this limit to the numerator and denominator in Eq. 4.32, we can then conclude that

$$\lim_{t \to \infty} \int_0^1 v(x)\overline{\mu}_t(dx) = v(1), \quad (4.33)$$

and we have shown that $\overline{\mu}_t(dx) \rightharpoonup \delta(1 - x)$. $\qquad \square$

Remark 3 Proposition 2 generalizes results for the Luo-Mattingly model and models from evolutionary games in which full-cooperation maximizes collective payoff (Luo and Mattingly 2017; Cooney 2020b). The proofs in those cases had relied on the fact that $G(1) \geq \langle G(x) \rangle_{\mu_t(dx)}$ at all times t, but did not require the additional assumption that the portion of the initial measure not concentrated at full-cooperation have a positive supremum Hölder exponent near $x = 1$. In the present case, the population concentrates upon full-cooperation even when group reproduction function is maximized by an interior level of cooperation.

5 Long-time behavior of multilevel PD dynamics for general initial measures

In this section, we consider the long-time behavior of solutions of Eq. (1.2) for initial measures $\mu_0(dx)$ with a given supremum Hölder exponent $\underline{\theta}$ near $x = 1$. For the Prisoner's Dilemma case, we present the proof of Theorem 2 characterizing long-time bounds on the time-averaged collective group-level reproduction rate in Sect. 5.1 and the proof of Theorem 3 for extinction or weak persistence of cooperation in Sect. 5.2. In Sect. 5.3, we prove Proposition 1, showing that $\overline{\mu}_t(dx)$ concentrates upon full-cooperation in the Prisoners' Delight case.

5.1 Bounds on average group reproduction function

Before presenting the proof of Theorem 2, we first provide a result analogous to Lemma 7 showing the existence of a sequence tending on which solutions to Eq. (1.6) can be bounded in terms of the infimum Hölder exponent near $x = 1$.

Lemma 8 *Suppose that $G(x), \pi(x)$ satisfy the assumptions of Theorem 1 and that $\mu_0(dx)$ has infimum Hölder exponent $\overline{\theta}$ near $x = 1$ that is nonzero and finite. Then, for $\Theta < \overline{\theta}$, there are positive constants E_1^v and $E_2^v(\Theta)$ and a sequence $\{t_n\}_{n \in \mathbb{N}}$ satisfying $t_n \to \infty$ such that*

$$\int_0^1 v(\phi_{t_n}(y))\psi_{t_n}(y)\mu_0(dy) \le E_1^v e^{\lambda G(0)t_n} + E_2^v(\Theta)e^{[\lambda G(1) - \Theta\pi(1)]t_n}. \tag{5.1}$$

Proof of Theorem 2 We will first consider the case in which $\overline{\theta} \ge \underline{\theta} > 0$, and then mention how to generalize our argument to the case in which $\underline{\theta} = 0$. Using the exponential normalization of Eq. (3.26) and the assumption that $\mu_0(dx)$ has supremum Hölder exponent $\underline{\theta}$ near $x = 1$, we can apply Lemma 7 for the test-function $v(x) = 1$ to see that, for $\Theta < \underline{\theta}$, there are constants C_1^1, $C_2^1(\Theta)$, and $\tilde{C}$ such that

$$\overline{\mu}_t([0, 1]) \le \left(C_1^1 e^{\lambda G(0)t} + C_2^1(\Theta)e^{[\lambda G(1) - \Theta\pi(1)]t} \right) \exp\left(-\lambda \int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds \right)$$

$$\le \tilde{C} \exp\left([\max\{\lambda G(1) - \Theta\pi(1), \lambda G(0)\}] t - \lambda \int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds \right). \tag{5.2}$$

Noting that $\overline{\mu}_t([0, 1]) = 1$, we must eventually have that $\frac{\lambda}{t}\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds \le \max\{\lambda G(1) - \Theta\pi(1), \lambda G(0)\}$, as otherwise the righthand side of Eq. (5.2) will take on a value less than 1 for sufficiently large t. Because this is true for all $\Theta < \underline{\theta}$, we can deduce that $\limsup_{t \to \infty} \frac{\lambda}{t}\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds \le \max\{\lambda G(1) - \underline{\theta}\pi(1), \lambda G(0)\}$.

Next, we look to confirm the reverse inequality. First, we may use Lemma 5 and Eq. (3.26) to see that, for any $\Theta > \underline{\theta}$, there is a sequence of times $\{t_n\}_{n \in \mathbb{N}}$ satisfying $t_n \to \infty$ and a constant $b(\Theta)$ such that

$$\overline{\mu}_{t_n}([0, 1]) \ge b(\Theta) \exp\left([\lambda G(1) - \Theta\pi(1)] t_n - \lambda \int_0^{t_n} \langle G(\cdot)\rangle_{\overline{\mu}_s} ds \right). \tag{5.3}$$

Using this estimate, we may deduce that $\limsup_{t \to \infty} \frac{\lambda}{t}\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds \ge \lambda G(1) - \Theta\pi(1)$ for $\Theta > \underline{\theta}$. Second, we denote $\check{G}_w := \min_{x \in [0, w]} G(x)$ and apply Lemma 6 to see that, for w sufficiently close to 0, there exists $A_w > 0$ such that

$$\overline{\mu}_t([0, 1]) \ge A_w \exp\left(\lambda \left[\check{G}_w t - \int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds \right] \right). \tag{5.4}$$

This allows us to deduce that $\limsup_{t \to \infty} \frac{\lambda}{t}\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds \ge \lambda \check{G}_w$ for w sufficiently close to 0. Combining our two lower bounds, we see that

$\limsup_{t\to\infty} \frac{\lambda}{t}\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds \geq \max\{\lambda G(1) - \Theta\pi(1), \lambda\check{G}_w\}$ for any $\Theta > \underline{\theta}$ and w close enough to 0. Because this bound holds for all such Θ and w, we can use the fact that $\check{G}_w \to G(0)$ to deduce that $\limsup_{t\to\infty} \frac{\lambda}{t}\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds \geq \max \lambda G(1) - \underline{\theta}\pi(1), \lambda G(0)$. Combining this with the reverse inequality, we can conclude that $\limsup_{t\to\infty} \frac{\lambda}{t}\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds = \max\{\lambda G(1) - \underline{\theta}\pi(1), \lambda G(0)\}$.

To study the corresponding limit infimum, we apply Lemma 8 for the test-function $v(x) = 1$ and the exponential normalization of Eq. (3.26) to see that there are constants E_1^1, E_2^1, and $\tilde{E}$ and a sequence $\{\tau_n\}_{n\in\mathbb{N}}$ satisfying $\tau_n \to \infty$ such that, for any $\Theta < \overline{\theta}$,

$$\overline{\mu}_{\tau_n}([0,1]) \leq \left(E_1^1 e^{\lambda G(0)\tau_n} + E_2^1(\Theta)e^{[\lambda G(1)-\Theta\pi(1)]\tau_n}\right)\exp\left(-\lambda\int_0^{\tau_n}\langle G(\cdot)\rangle_{\overline{\mu}_s} ds\right)$$
$$\leq \tilde{E}\exp\left(\max\{\lambda G(1) - \Theta\pi(1), \lambda G(0)\}\tau_n - \lambda\int_0^{\tau_n}\langle G(\cdot)\rangle_{\overline{\mu}_s} ds\right).$$
(5.5)

Because $\overline{\mu}_{\tau_n}([0,1])$ remains normalized for all time, Eq. (5.5) tells us that $\liminf_{t\to\infty} \frac{\lambda}{t}\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds \leq \max\{\lambda G(1) - \Theta\pi(1), \lambda G(0)\}$ for any $\Theta < \overline{\theta}$, and therefore we can deduce that $\liminf_{t\to\infty} \frac{\lambda}{t}\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds \leq \max\{\lambda G(1) - \overline{\theta}\pi(1), \lambda G(0)\}$.

To study the analogous lower bound on the time-average collective group-reproduction rate, we first we apply Lemma 5 and Eq. (3.26) to see that, for $\Theta \geq \overline{\theta}$, there exists $d(\Theta) > 0$ such that

$$\overline{\mu}_t([0,1]) \geq d(\Theta)\exp\left([\lambda G(1) - \Theta\pi(1)]t - \lambda\int_0^t\langle G(\cdot)\rangle_{\overline{\mu}_s} ds\right). \quad (5.6)$$

Combining the bounds of Eqs. (5.4) and (5.6) allows us to see that $\liminf_{t\to\infty} \frac{\lambda}{t}\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds \geq \max\{\lambda G(1) - \Theta\pi(1), \lambda\check{G}_w\}$ for any $\Theta > \overline{\theta}$ and w sufficiently close to 0. Because this holds for all such Θ and w, we can conclude that $\liminf_{t\to\infty} \frac{\lambda}{t}\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds \geq \max\{\lambda G(1) - \overline{\theta}\pi(1), \lambda G(0)\}$. Because we have also confirmed the reverse inequality, we can then conclude that $\liminf_{t\to\infty} \frac{\lambda}{t}\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds = \max\{\lambda G(1) - \overline{\theta}\pi(1), \lambda G(0)\}$.

Finally, dividing both sides by λ in our expressions for $\liminf_{t\to\infty} \frac{\lambda}{t}\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds$ and $\limsup_{t\to\infty} \frac{\lambda}{t}\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds$ then provides us with the long-time bounds presented in Eq. (1.34) on the time-averaged collective outcome $\frac{1}{t}\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds$ for the population.

For the cases in which $\underline{\theta} = 0$ or $\overline{\theta} = 0$, we can still obtain the upper bounds from Eqs. (5.3) and (5.6) using our existing proof. To obtain corresponding lower bounds on $\overline{\mu}_t([0,1])$, we can use the bound of solutions along characteristics from Lemma 4 and the fact that $\mu_0(dx)$ is a probability measure. Then applying these bounds allows us to deduce the limiting time-averaged behavior of Eq. (1.34) in these cases as well. □

5.2 Long-time extinction or persistence of cooperation

Proof of Theorem 3 For the case in which $\lambda\,[G(1) - G(0)] > \underline{\theta}\pi(1)$, the long-time persistence of cooperation follows from the bounds from Eq. (1.34) for the time-averaged collective outcome. Because $\limsup_{t\to\infty} \frac{1}{t}\int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s} ds = G(1) - \lambda^{-1}\underline{\theta}\pi(1) > G(0)$, the time-average of $\langle G(\cdot)\rangle_{\overline{\mu}_t} = \int_0^1 G(x)\overline{\mu}_t(dx)$ will achieve a value bounded away from $G(0)$ at an infinite sequence of times, and therefore we can show that the fraction of cooperators $\int_0^1 x\overline{\mu}_t(dx)$ must be bounded away from 0 on an infinite sequence as well.

Next we turn to proving the extinction of cooperation when $\lambda\,[G(1) - G(0)] < \underline{\theta}\pi(1)$. To show weak convergence of $\overline{\mu}_t(dx)$ to $\delta(x)$, we consider any continuous test function $v(x)$ and look to show that $\int_0^1 v(x)\overline{\mu}_t(dx) \to v(0)$ as $t \to \infty$. We can use the continuity of $v(x)$ and the fact that $\overline{\mu}_t(dx)$ is a probability distribution to show that for any $\epsilon > 0$, there is a $\delta > 0$ such that

$$\left| \int_0^1 v(x)\overline{\mu}_t(dx) - v(0) \right| \le \epsilon + 2||v||_\infty \int_\delta^1 \overline{\mu}_t(dx) \qquad (5.7)$$

We note that $\mu_0(dx)$ must have a supremum Hölder exponent $\underline{\theta} > 0$ near $x = 1$ in order to satisfy our assumption on λ. Therefore we can apply Lemma 6 and the normalization relation from Eq. (1.7) to say that, for w sufficiently close to 0, there exists $A_w > 0$ such that

$$\int_\delta^1 \overline{\mu}_t(dx) = \frac{\int_\delta^1 \mu_t(dx)}{\int_0^1 \mu_t(dx)} \le A_w^{-1} e^{-\lambda \check{G}_w t} \int_\delta^1 \mu_t(dx), \qquad (5.8)$$

where $\check{G}_w = \min_{x\in[0,w]} G(x)$. We can apply Lemma 5 to the integral on the righthand side of Eq. (5.8) to say that, for $\Theta < \underline{\Theta}$, there exists $B(\Theta) < \infty$ such that

$$\int_\delta^1 \overline{\mu}_t(dx) \le B(\Theta) A_w^{-1} \exp\left(\left\{ \lambda\left[G(1) - \check{G}_w \right] - \Theta\pi(1) \right\} t \right) \qquad (5.9)$$

Because $G(\cdot) \in C^1([0, 1])$, we know that we can make $\check{G}_w$ arbitrarily close to $G(0)$ by choosing w sufficiently close to 0. Because our assumption on λ is the strict inequality $\lambda\,[G(1) - G(0)] < \underline{\theta}\pi(1)$, we know, for any given λ, that we can choose Θ sufficiently close to $\underline{\theta}$ and w sufficiently close to 0 such that

$$\lambda\left[G(1) - \check{G}_w \right] < \Theta\pi(1) \qquad (5.10)$$

as well. This condition then guarantees that

$$2||v||_\infty \int_\delta^1 \overline{\mu}_t(dx) \le 2B(\Theta) A_w^{-1} ||v||_\infty$$

$$\times \exp\left(\left\{\lambda\left[G(1) - \check{G}_w\right] - \Theta\pi(1)\right\} t\right) \to 0 \text{ as } t \to \infty,$$

which, in combination with Eq. (5.7), allows us to conclude that $\mu_t(dx) \rightharpoonup \delta(x)$ as $t \to \infty$ when $\lambda\left[G(1) - G(0)\right] < \underline{\theta}\pi(1)$. $\qquad \square$

Because we show in Theorem 3 that cooperation dies out for $\lambda < \lambda^*$ and weakly persists when $\lambda > \lambda^*$, it is natural to ask what happens in the edge case when $\lambda = \lambda^*$. For the special case in which the group reproduction achieves a unique minimum at $x = 0$, we can show that the population still concentrates at $\delta(x)$ when $\lambda\left[G(1) - G(0)\right] = \underline{\theta}\pi(1)$. We rely on the following lemma, which was previously used to study special cases of the model under consideration (Luo and Mattingly 2017; Cooney 2020b).

Lemma 9 *Consider $\pi(x), G(x) \in C^1([0, 1])$ and suppose that $G(x) > G(0)$ for $x \in (0, 1]$. If $\int_0^\infty \left[\langle G(\cdot)\rangle_{\mu_s} - G(0)\right] ds < \infty$, then $\langle G(\cdot)\rangle_{\mu_t} \to G(0)$ and $\bar{\mu}_t(dx) \rightharpoonup \delta(x)$ as $t \to \infty$.*

Proposition 3 *Suppose the initial distribution $\mu_0(dx)$ has supremum Hölder exponent $\underline{\theta}$ near 1 with corresponding Hölder constant $C_{\underline{\theta}} < \infty$. If $G(x) > G(0)$ for $x \in (0, 1]$ and $\lambda\left[G(1) - G(0)\right] = \underline{\theta}\pi(1)$, the $\overline{\mu}_t(dx) \rightharpoonup \delta(x)$ as $t \to \infty$.*

Proof of Proposition 3 We can use the continuity of the test function $v(x)$ and the fact that $\bar{\mu}_t(dx)$ is a probability distribution to see that, for any $\epsilon > 0$, there is a $\delta > 0$ such that

$$\left|\int_0^1 v(x)\overline{\mu}_t(dx) - v(0)\right| \le \epsilon + 2||v||_\infty \int_\delta^1 \overline{\mu}_t(dx) \tag{5.11}$$

From our assumptions that $\lambda\left[G(1) - G(0)\right] = \underline{\theta}\pi(1)$ and $\mu_0(dx)$ has Hölder exponent $\underline{\theta} > 0$ with positive, finite Hölder constant, can use Lemma 5 and the exponential normalization from Eq. (3.25) to see that there is a constant $B(\underline{\theta}) < \infty$ such that

$$\int_\delta^1 \overline{\mu}_t(dx) \le B(\underline{\theta})e^{[\lambda G(1) - \underline{\theta}\pi(1)]t} \exp\left(-\lambda \int_0^t \langle G(\cdot)\rangle_{\mu_s} ds\right)$$

$$= B(\underline{\theta}) \exp\left(-\lambda \int_0^t \left[\langle G(\cdot)\rangle_{\mu_s} - G(0)\right] ds\right). \tag{5.12}$$

Because $G(x) \ge G(0)$ for $x \in [0, 1]$, either $\int_0^\infty \left[\langle G(\cdot)\rangle_{\mu_s} - G(0)\right] ds < \infty$ or $\int_0^\infty \left[\langle G(\cdot)\rangle_{\mu_s} - G(0)\right] ds \to \infty$ as $t \to \infty$. In the former case, Lemma 9 tells us that $\overline{\mu}_t(dx) \rightharpoonup \delta(x)$ as $t \to \infty$. In the alternate case, $\int_0^\infty \left[\langle G(\cdot)\rangle_{\mu_s} - G(0)\right] ds \to \infty$ corresponds to $e^{-\int_0^t \left[\langle G(\cdot)\rangle_{\mu_s} - G(0)\right] ds} \to 0$, and we can use Eq. (5.11) allows us to conclude that

$$\left|\int_0^1 v(x)\overline{\mu}_t(dx) - v(0)\right| \le \epsilon \text{ as } t \to \infty,$$

and therefore we see that $\overline{\mu}_t(dx) \rightharpoonup \delta(x)$ when $\lambda\left[G(1) - G(0)\right] = \underline{\theta}\pi(1)$. $\qquad \square$

Remark 4 A key aspect of the proof of convergence to $\delta(x)$ in Theorem 3 was that the strict inequality condition that $\lambda < \lambda^*$ provided us the freedom to consider a slightly smaller exponent $\tilde{\theta} < \underline{\theta}$ and still obtain exponential decay to $\delta(x)$. In the equality case with $\lambda = \lambda^*$ considered in Proposition 3, we had to assume a finite supremum Hölder constant $C_{\underline{\theta}}$ to get a bound involving the exact supremum Hölder exponent $\underline{\theta}$.

5.3 Convergence to full-cooperation in the prisoners' delight

Before proceeding to the proof of Proposition 1, we introduce several lemmas that allow us to estimate the measure $\mu_t(dx)$ solving Eq. (1.2) under the Prisoners' Delight scenario. These lemmas serve as analogues to Lemma 5 and 6, and can be proved with a similar approach.

Lemma 10 *Suppose that $G(x), \pi(x)$ satisfy the assumptions of Proposition 1. For any $a \in (0, 1)$, there exists $M_a < \infty$ such that*

$$\rho_t([0, a]) := \mu_t((0, a]) \leq M_a e^{\lambda G(0)t}. \tag{5.13}$$

Lemma 11 *Suppose that $G(x), \pi(x)$ satisfy the assumptions of Proposition 1. Considering the quantity $\hat{G}_a := \min_{x \in [a,1]} G(x)$, we see that for z sufficiently close to 1, there exists $A_z > 0$ such that*

$$\mu_t([0, 1]) \geq A_z e^{\lambda \hat{G}_z t}. \tag{5.14}$$

Proof of Proposition 1 To show that $\overline{\mu}_t(dx)$ converges to a delta-function at $x = 1$, we consider a continuous test function $v(x)$ and use the fact that $\overline{\mu}_t(dx)$ is a probability measure to see that, for any $\epsilon > 0$, there is a $\delta > 0$ such that

$$\left| \int_0^1 v(x)\overline{\mu}_t(dx) - v(1) \right| \leq \int_0^{1-\delta} |(v(x) - v(1))| \overline{\mu}_t(dx)$$
$$+ \int_{1-\delta}^1 |(v(x) - v(1))| \overline{\mu}_t(dx) \leq \epsilon + 2||v||_\infty \int_0^{1-\delta} \overline{\mu}_t(dx) \tag{5.15}$$

We can use the assumption that $a_0 := \mu_0(\{0\}) < 1$ to write our initial measure by a decomposition of the form

$$\mu_0(dx) = a_0 \delta(x) + (1 - a_0) \rho_0(dx) \tag{5.16}$$

where $\rho_0(dx)$ is a probability measure satisfying $\rho_0((0, 1]) = 1$. Because we have assumed that $\mu_0(dx) \neq \delta(x)$, we can apply Lemma 11 to see that, for z sufficiently close 1, there is $A_z > 0$ such that $\mu_t([0, 1]) \geq A_z e^{\lambda \hat{G}_z t}$. Combining this with the normalization relation of Eq. (1.7) and the decomposition of our initial measure from

Eq. ([5.16](#)), we can estimate that

$$
\int_0^{1-\delta} \overline{\mu}_t(dx) = \frac{\int_0^{1-\delta} \mu_t(dx)}{\int_0^1 \mu_t(dx)} \leq A_z^{-1} e^{-\lambda \hat{G}_z t} \left[a_0 e^{\lambda G(0)t} + (1-a_0)\rho_t\left([0,1-\delta]\right) \right]
$$
(5.17)

From Lemma [10](#), we know that there is $M_{1-\delta} < \infty$ such that $\rho_t\left([0,1-\delta]\right) \leq M_{1-\delta} e^{\lambda G(0)t}$. Combining this with our estimates from Eqs. ([5.15](#)) and ([5.17](#)) allows us to see that there exists $\tilde{M} < \infty$ such that

$$
\left| \int_0^1 v(x)\overline{\mu}_t(dx) - v(1) \right| \leq \epsilon + 2\tilde{M}||v||_\infty \exp\left(\lambda \left[G(0) - \hat{G}_z \right] t \right).
$$
(5.18)

Because $G(1) > G(0)$ for the PDel game, we know that we can pick z sufficiently close to 1 such that $\hat{G}_z > G(0)$. For such choices of z, we can deduce that

$$
2\tilde{M}||v||_\infty \exp\left(\lambda \left[G(0) - \hat{G}_z \right] t \right) \to 0 \text{ as } t \to \infty
$$

as long as $\lambda > 0$, and we can conclude that $\overline{\mu}_t(dx) \rightharpoonup \delta(1-x)$ as $t \to \infty$ for $\lambda > 0$ and $\mu_0(dx) \neq \delta(x)$. $\qquad \square$

6 Generalization to multilevel competition with N populations

In this section, we discuss results for the N-population multilevel selection model whose dynamics are described by Eq. ([1.10](#)). We provide the proof of Theorem [4](#), demonstrating a sufficient condition for the long-time behavior of the population to feature concentration upon the subpopulation with the maximal principal growth rate. In Sect. [C](#), we apply these results to study the multilevel dynamics for the generalizations of the Hawk-Dove and Stag-Hunt games in a single population by reformulating these results as a two-population problem and characterize the long-time behavior for Eq. ([1.2](#)) for these games.

Proof of Theorem [4] Using the normalization from Eq. ([1.13](#)), we can estimate the $\overline{\mu}_t^j\left([0,1]\right)$ of having groups in subpopulation $j \in \mathcal{N} - \{k\}$ by

$$
\overline{\mu}_t^j\left([0,1]\right) = \frac{\mu_t^j\left([0,1]\right)}{\sum_{i=1}^N \mu_t^i\left([0,1]\right)} \leq \frac{\mu_t^j\left([0,1]\right)}{\mu_t^j\left([0,1]\right) + \mu_t^k\left([0,1]\right)}.
$$
(6.1)

If faced with the relevant case for the principal growth r_k^m, we can use the quantities $\check{G}_w = \min_{x \in [0,w]} G(x)$ for w sufficiently close to 0, $\hat{G}_z = \min_{x \in [z,1]} G(x)$ for z

sufficiently close to 1, or $\overline{\Theta}^k$ sufficiently close to but greater than $\overline{\theta}^k$ to introduce a modified principal growth rate $\tilde{r}_k^m$ that satisfies

$$
\tilde{r}_k^m = \left\{
\begin{array}{ll}
\lambda \hat{G}_k^z & : \pi_k(x) < 0 \\
\lambda G_k(1) - \overline{\Theta}^k \pi_k(1) & : \pi_k(x) > 0 \,, \; \lambda \left[G_k(1) - G_k(0) \right] - \overline{\theta}^k \pi_k(1) > 0 \\
\lambda \tilde{G}_k^w & : \pi_k(x) > 0 \,, \; \lambda \left[G_k(1) - G_k(0) \right] - \overline{\theta}^k \pi_k(1) < 0
\end{array}
\right\} > r_j^M .
\tag{6.2}
$$

Across the three cases, we can apply Lemmas 5, 6, and 11 to see that there exists $L > 0$ such that $\mu_t^k([0,1]) \geq L e^{\tilde{r}_k^m t}$. We can then combine this with Eq. (6.1) to estimate that

$$
\overline{\mu}_t^j([0,1]) = \frac{e^{-\tilde{r}_k^m} \mu_t^j([0,1])}{e^{-\tilde{r}_k^m} \mu_t^j([0,1]) + e^{-\tilde{r}_k^m} \mu_t^k([0,1])} \leq \frac{e^{-\tilde{r}_k^m t} \mu_t^j([0,1])}{e^{-\tilde{r}_k^m t} \mu_t^j([0,1]) + L}
$$
$$
\leq L^{-1} e^{-\tilde{r}_k^m t} \mu_t^j([0,1]) .
\tag{6.3}
$$

Our next step is to estimate $\mu_t^j([0,1])$. If $\pi_j(x) < 0$ for $x \in [0,1]$, we may estimate that

$$
\int_0^t (G(\phi_s(x)) - G(1)) \, ds = \int_x^{\phi_t(x)} \frac{G(q) - G(1)}{q(1-q)|\pi(q)|} dq \leq \int_0^1 \frac{[G(q) - G(1)]_+}{q(1-q)|\pi(q)|} dq < \infty,
\tag{6.4}
$$

and therefore there exists an $A_1 < \infty$ such that

$$
\mu_t^j([0,1]) = e^{\lambda G(1)t} \int_0^1 e^{\lambda \int_0^t [G(\phi_s(x)) - G(1)] ds} \mu_0(dx) \leq A_1 e^{\lambda G(1)t} = A_1 e^{r_j^M t}.
\tag{6.5}
$$

If instead $\pi_j(x) > 0$ for $x \in [0,1]$, we can use the fact that $\mu_0^j(dx)$ has supremum Hölder exponent $\underline{\theta} > 0$ to apply Lemma 7. Therefore we can introduce a modified supremum Hölder exponent $\underline{\Theta} < \underline{\theta}$ sufficiently close to $\underline{\theta}$ and a modified principal growth rate $\tilde{r}_j^M \in (r_j^M, \tilde{r}_k^m)$, which allows us to see that there exist constants $C_1^1, C_2^1, A_2 < \infty$ such that

$$
\mu_t^j([0,1]) \leq C_1^1 e^{\lambda G_j(0)t} + C_2^1 e^{[\lambda G_j(1) - \underline{\Theta}^j \pi_j(1)]t} \leq A_2 e^{\max\{\lambda G_j(0), \lambda G_j(1) - \underline{\Theta}^j \pi_j(1)\}t} \leq A_2 e^{\tilde{r}_j^M t}.
\tag{6.6}
$$

Between the two cases for the sign of $\pi_j(x)$, we can see that there exists $A < \infty$ and $\tilde{r}_j^M < \tilde{r}_m^k$ such that $\mu_t^j([0,1]) \leq A e^{\tilde{r}_j^M t}$. Combining this estimate with Eq. (6.3), we can use the fact that $\tilde{r}_k^m > \tilde{r}_j^M$ to deduce that

$$
\overline{\mu}_t^j([0,1]) \leq AL^{-1} \exp\left(\left[\tilde{r}_j^M - \tilde{r}_k^m \right] t \right) \to 0 \text{ as } t \to \infty,
\tag{6.7}
$$

and we can conclude that the population will concentrate upon subpopulation k in the long-time limit.

Now we turn to describing the long-time behavior of $\overline{\mu}_t^k(dx)$ in the case for which $\mu_0^k(dx)$ has a well-defined Hölder exponent and constant near $x = 1$. We can write describe the distribution of group compositions in subpopulation k as

$$
\int_0^1 v_k(x)\overline{\mu}_t^k(dx) = \frac{\int_0^1 v_k(x)\mu_t^k(dx)}{\displaystyle\sum_{j=1}^N \mu_t^j([0,1])}
$$

$$
= \frac{e^{[\theta^k \pi_k(1) - \lambda G_k(1)]t} \int_0^1 v_k(x)\mu_t^k(dx)}{e^{[\theta^k \pi_k(1) - \lambda G_k(1)]t} \mu_t^k([0,1]) + \displaystyle\sum_{\substack{j=1 \\ j \neq k}}^N \left\{ e^{[\theta^k \pi_k(1) - \lambda G_k(1)]t} \mu_t^j([0,1]) \right\}}. \tag{6.8}
$$

Using the same approach as in the proof of Theorem 1, we can see that

$$
\left\{
\begin{array}{l}
e^{[\theta^k \pi_k(1) - \lambda G_k(1)]t} \int_0^1 v_k(x)\mu_t^k(dx) \;\to\; \int_0^1 v_k(x) f_{\theta^k}^\lambda(x)dx \\[2mm]
e^{[\theta^k \pi_k(1) - \lambda G_k(1)]t} \mu_t^k([0,1]) \;\to\; \int_0^1 f_{\theta^k}^\lambda(x)dx
\end{array}
\right\} \quad \text{as } t \to \infty. \tag{6.9}
$$

Because $r_k^m = \lambda G_k(1) - \theta^k\pi(1) > r_j^M$ in this case we know from above that $e^{[\theta^k \pi_k(1) - \lambda G_k(1)]t}\mu_t^j([0,1]) \to 0$ for $j \neq k$. Applying this to Eq. (6.8) allows us to conclude that

$$
\int_0^1 v_k(x)\overline{\mu}_t^k(dx) \to \frac{\int_0^1 v_k(x) f_{\theta^k}^\lambda(dx)}{\int_0^1 f_{\theta^k}^\lambda(dx)} = \int_0^1 v_k(x) p_{\theta_k}^\lambda(dx) \quad \text{as } t \to \infty. \tag{6.10}
$$

7 Discussion

In this paper, we have analyzed the long-time behavior in a PDE model of multilevel selection, in which a tension exists between the individual-level incentive to defect and group-level competition favoring groups that cooperate. We show that defectors take over the group-structured population when within-group competition is stronger than between-group competition, and that cooperation can weakly persist in the population for all time when the relative strength of between-group competition exceeds a threshold value. We also provide sufficient conditions for the population to converge to a long-time steady state density featuring coexistence of cooperators and defectors, and further characterize the average level of cooperation and group-reproduction rate at steady state in the limit of strong between-group competition. These results generalize and extend previous work on PDE models of multilevel selection with within-group and between-group dynamics arising from frequency-independent competition (Luo

2014; van Veelen et al. 2014; Luo and Mattingly 2017) or from the payoffs of evolutionary games (Cooney 2019b, 2020a, 2019a).

By considering arbitrary C^1 functions $\pi(x)$ and $G(x)$ to describe the within-group and between-group competition, we have a general analysis to study how long-term cooperation depends on the tug-of-war of between individual-level and group-level incentives. In this more general setting, we are able to understand the key role the full-cooperator group plays in determining the level of cooperation and collective average payoff supported by the long-time behavior of multilevel dynamics. In particular, we see that increasing the relative advantage of the full-cooperator group or increasing the initial cohort of many-cooperator groups (corresponding to lower Hölder exponent θ) helps to promote the evolution of cooperative behavior via multilevel selection, consistent with analytical and simulation results from finite population models of multilevel selection (Markvoort et al. 2014; Traulsen and Nowak 2006; Traulsen et al. 2008).

Considering a broader class of initial conditions for which the infimum and supremum Hölder exponents or constants disagree also reveals important properties of our multilevel dynamics. In particular, convergence to a steady state solution of Eq. (1.2) in not guaranteed for generic initial probability measures of group compositions, and instead a more natural notion for quantifying the ability for cooperation to survive via multilevel selection is the weak persistence of cooperation for sufficiently strong between-group competition. This distinction between weak persistence and convergence to steady state may also be relevant for exploring multilevel selection for more complex strategy spaces, as it may be more difficult to identify or quantify the possible steady state behaviors beyond a one-dimensional state space for group compositions. As a question for future work, one could look for analogous weak persistence thresholds for PDE models of multilevel selection including additional evolutionary forces like genetic drift or migration (Ogura and Shimakura 1987a, b; Fontanari and Serva 2013, 2014a, b), or models in which the assumption of fixed group size is relaxed and group-level events such as fission and fusion can help to drive the evolution of cooperation (Simon 2010; Simon and Nielsen 2012; Simon et al. 2013; Simon and Pilosov 2016).

As in the special cases previously studied for evolutionary games (Cooney 2019b, 2020a), we establish that the collective payoff of the steady state population is limited by the payoff of a full-cooperator group. This means that the so-called "shadow of lower-level selection" is present for all group reproduction functions which are maximized by an intermediate level of cooperation: no level of between-group competition produces a steady-state population that achieves the maximum possible group-reproduction rate. Even in the limit of infinitely strong between-group competition, the population still concentrates as a delta-function at a level of cooperation that produces the same collective-reproduction rate as that of a (sub-optimal) full-cooperator group. In addition, we have now established a dynamical analogue to the shadow of lower-level selection in terms of the bounds on the time-average of the group-reproduction rate in the population, highlighting that this limitation of the collective outcome to that of a full-cooperator group can be seen through the dynamics of our model of multilevel selection. Given the bounds on the time-averaged collective outcome, a natural question for future research is whether there is a sense

in which the potentially oscillatory long-time solutions of the multilevel dynamics concentrate upon group compositions with the group-reproduction rate $G(1)$ for sufficiently strong between-group competition, extending the concentration behavior seen for steady-state densities.

Our analysis of the threshold selection strength and average payoff at steady state also provides a window to understanding how mechanisms that alter within-group and between-group competition may facilitate cooperation. In particular, we see that decreasing the incentive to defect in a many-cooperator group, $\pi(1)$, can help decrease the between-group competition strength needed to allow long-time survival of cooperation, but it cannot increase the maximum possible achievable group-average payoff in the limit of strong between-group competition. Because altering within-group interactions – through the mechanisms of assortment, other-regarding preference, indirect reciprocity, and network reciprocity – can increase $\pi(0)$ to a positive value and provide cooperators an individual-level advantage in a many-defector group (Grafen 1979; Maynard Smith 1982; Taylor and Nowak 2007; Cooney 2019a), we can also see that such within-group mechanisms can change the dynamics of multilevel selection from following the generalized Prisoners' Dilemma assumptions to following the generalized Hawk-Dove or Stag-Hunt assumptions. Through our results on the long-time behavior of the multilevel dynamics under each of each of these generalized games, we can further explore the synergistic effects of within-group population structure and between-group competition on the evolution of cooperation (Cooney 2019a).

This paper constitutes an in-depth analysis of a class of hyperbolic PDEs that generalize the equation studied by Luo and Mattingly (Luo 2014) to include variety of replication functions corresponding to different within-group and between-group competition scenarios. In turn, Eq. (1.2) corresponds to the two-level replicator equation that arises from a generalization of the two-level Moran process introduced by Luo (2014) under one possible scaling in the limit of infinite group size and of an infinite number of groups. Because our PDE is derived from an individual-based model with finite populations, an important consideration is the extent to which the behavior of solutions to the PDE correspond to properties of the underlying finite population dynamics. In particular, an interesting question for future research is the extent to which behaviors like persistence of cooperation or convergence to steady state densities in the PDE corresponds to any analogous support for cooperation in the original stochastic two-level selection model with finite populations or in alternate deterministic scaling limits that retain the diffusive effects of finite group size (Luo 2014; Luo and Mattingly 2017; Cooney 2019b).

Our results in Sect. 6 on multiple population dynamics also provide motivation for future work on the coevolution of cooperative strategies and the games played by members of groups. Because the principle growth rate of a subpopulation often corresponds to the average collective outcome at steady state, we can see the long-time behavior of the multipopulation competition as picking out the pair of gain and group-reproduction functions $\pi_j(x)$ and $G_j(x)$ most capable of producing a beneficial level of cooperation under the two-level dynamics for a given set of initial distributions. Future modeling work can examine how competition between different group types can impact the evolution of group properties such as within-group population structures or social norms (Cooney 2019a; Santos et al. 2007). In addition, the approach of

reformulating multilevel selection models multipopulation models can be extended to evolutionary games with more complicated within-group dynamics, allowing for the analysis of long-time behavior for nonlinear public goods games (Archetti and Scheuring 2011; Pacheco et al. 2009) or for models of efficient extraction of common-pool resources (Tavoni et al. 2012; Tilman et al. 2017). Going forward, the study of multilevel selection models with multiple group types can serve as a potential next step for mathematically understanding the role which multilevel selection can play on the cusp of major evolutionary transitions.

Data Availability Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

A Well-posedness of multilevel dynamics

In this section, we address the well-posedness of measure-valued solutions to Eq. (1.2). Our approach will be to first establish well-posedness of solutions $\mu_t(dx)$ to the linear multilevel dynamics of Eq. (1.6), and then to use the exponential normalization from Eq. (3.25) to demonstrate that Eq. (1.2) is well-posed as well.

For the auxiliary linear problem of Eq. (1.6), we can use Eqs. (3.1) and (3.3) to obtain the following representation formula for its solution $\mu_t(dx)$ starting from an initial measure $\mu_0(dx)$:

$$\int_0^1 v(x)\mu_t(dx) = \int_0^1 v(\phi_t(x)) \exp\left(\int_0^t G(\phi_s(x))ds\right)\mu_0(dx). \qquad \text{(A.1)}$$

The righthand Eq. (A.1) is a positive linear functional, so we can deduce that $\mu_t(dx)$ is a Borel measure due to the Riesz-Markov-Kakutani theorem. We can also check from this representation formula to see that $\mu_t(dx)$ solves Eq. (1.6) for any test-function $v(x) \in C^1([0, 1])$. Existence for any $C([0, 1])$ test-function can then be established by density of $C^1([0, 1])$ functions in $C([0, 1])$.

To further explore the well-posedness of solutions of the full multilevel dynamics of Eq. (1.2), it is helpful to establish a bijective correspondence between the measures $\mu_t(dx)$ and $\overline{\mu}_t(dx)$ solving Eqs. (1.6) and (1.2). Taking inspiration from the exponential normalization relation from Eq. (3.25), we use the following expressions to show that, for any continuous test-function $v(x)$, $\mu_t(dx)$ and $\overline{\mu}_t(dx)$ can be expressed a as functions of the other: expressions

$$\int_0^1 v(x)\overline{\mu}_t(dx) = \exp\left(-\lambda \int_0^t \langle G(\cdot)\rangle_{\mu_s}ds\right)\int_0^1 v(x)\mu_t(dx). \qquad \text{(A.2a)}$$

$$\int_0^1 v(x)\mu_t(dx) = \exp\left(\lambda \int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s}ds\right)\int_0^1 v(x)\overline{\mu}_t(dx). \qquad \text{(A.2b)}$$

This correspondence can be established for $C^1([0, 1])$ test-functions by applying the push-forward representation of Eq. (3.3) to the righthand side, and then differentiating both sides with respect to time. Eq. (A.2) can then be confirmed for $C([0, 1])$ test-functions by the density of continuously differentiable functions within the space of continuous functions.

From the normalization relation provided by Eq. (1.7), we can use the existence of the solution $\mu_t(dx)$ to Eq. (1.6) to deduce the existence of a Borel measure $\overline{\mu}_t(dx)$ solving Eq. (1.2) in the weak sense. For uniqueness, we will make use of the mapping from Eq. (A.2b) to help deduce uniqueness for Eq. (1.2) from uniqueness of the linear problem.

It is also important to consider the space in which our solutions $\overline{\mu}_t(dx)$ live. From the weak form of the multilevel dynamics presented in Eq. (1.5), it natural to consider $C^1([0, 1])$ test-functions due to the advection term and to consider a family measures $\mu_t(dx)$ with a differentiable dependence in time due to the time-derivative in the PDE. From the push-forward representation formula of Eq. (3.3) and the fact that our results on long-time behavior only require test-functions $v(x) \in C^1([0, 1])$, it also makes sense to consider uniformly continuous test functions and measures $\mu_t(dx)$ that vary continuously in time. As a result, in Proposition 4, we establish well-posedness of solutions $\overline{\mu}_t(dx)$ of Eq. (1.2) both taking values in the space of Borel measures with uniformly continuous time-dependence and taking values in the dual space of $C^1([0, 1])$ functions with continuously differentiable time-dependence.

Proposition 4 *Denote by $\mathcal{M}([0, 1])$ and $C^1([0, 1])^*$ the set of Borel measures on $[0, 1]$ and the dual space of $C^1([0, 1])$, respectively. Given the initial measure $\mu_0(dx) \in \mathcal{M}([0, 1])$ and any $T \geq 0$, there exists a unique $\overline{\mu}_t(dx) \in C([0, T]; \mathcal{M}([0, 1])) \cap C^1([0, T]; C^1([0, 1])^*)$ such that $\overline{\mu}_t(dx)$ is a weak solution to Eq. (1.2).*

Proof We start by establishing the well-posedness of solutions $\mu_t(dx)$ to the linear model from Eq. (1.6). Knowing from the representation formula of Eq. (3.3) that $\mu_t(dx)$ exists in the appropriate space, we now suppose that Eq. (1.6) has two solutions $\mu_t^1(dx)$ and $\mu_t^2(dx)$ for given initial measure $\mu_0(dx)$. Considering the difference $\tilde{\mu}_t(dx) := \mu_t^1(dx) - \mu_2^1(dx)$ with $\tilde{\mu}_0(dx) = 0dx$. Considering a test-function $v(t, x) \in C^{1,1}([0, T] \times [0, 1])$ we can use the linearity of Eq. (1.6) to see that

$$\frac{\partial}{\partial t} \int_0^1 v(t, x) \tilde{\mu}_t(dx)$$
$$= \int_0^1 \left[\frac{\partial v(t, x)}{\partial t} - x(1 - x)\pi(x)\frac{\partial v(t, x)}{\partial x} + \lambda G(x)v(t, x) \right] \tilde{\mu}_t(dx). \quad \text{(A.3)}$$

We can simplify Eq. (A.3) by choosing $v(t, x)$ solve the dual problem to the linear dynamics of Eq. (1.6), given by

$$\frac{\partial v(t, x)}{\partial t} - x(1 - x)\pi(x)\frac{\partial v(t, x)}{\partial x} = -\lambda G(x)v(t, x) \;,\; v(T, x) = \omega(x) \in C^1([0, 1]).$$
$$\text{(A.4)}$$

Noting that such a solution $v(t, x)$ exists by the method of characteristics, we can see with this choice that the righthand side of Eq. (A.3) will vanish. This allows us to integrate Eq. (A.3) in time to see that

$$\int_0^1 v(T, x)\tilde{\mu}_T(dx) = \int_0^1 v(0, x)\tilde{\mu}_0(dx) \implies \int_0^1 \omega(x)\tilde{\mu}_T(dx) = 0. \quad (A.5)$$

Because this identity holds for each $\omega(x) \in C^1([0, 1])$, we can deduce that the solution $\mu_t(dx)$ to Eq. (1.6) is unique in $C^1([0, 1])^*$, and further that it is unique in $\mathcal{M}([0, 1])$ by the density of continuous test-functions in $C^1([0, 1])$.

Existence of a weak solution $\overline{\mu}_t(dx)$ to Eq. (1.2) in $C([0, T]; \mathcal{M}([0, 1])) \cap C^1([0, T]; C^1([0, 1])^*)$ follows from the existence of a solution $\mu_t(dx)$ to Eq. (1.6) and the mapping of Eq. (3.25). To establish uniqueness, we consider solutions $\overline{\mu}_t^1(dx)$ and $\overline{\mu}_t^2(dx)$ to Eq. (1.2) for initial measure $\mu_0(dx)$, and apply the mapping of Eq. (A.2b) and the uniqueness of the solution $\mu_t(dx)$ to Eq. (1.6) to see that, for any C^1 test-function $v(x)$,

$$\int_0^1 v(x)\mu_t(dx) = \int_0^1 v(x) \exp\left(\lambda \int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s^1} ds\right) \overline{\mu}_t^1(dx)$$

$$= \int_0^1 v(x) \exp\left(\lambda \int_0^t \langle G(\cdot)\rangle_{\overline{\mu}_s^2} ds\right) \overline{\mu}_t^2(dx). \quad (A.6)$$

We can then use this to deduce that there is a function of time $c(t)$ such that

$$\int_0^1 v(x)\overline{\mu}_t^1(dx) = c(t) \int_0^1 v(x)\overline{\mu}_t^2(dx). \quad (A.7)$$

Because $\overline{\mu}_t^1(dx)$ and $\overline{\mu}_t^2(dx)$ are probability measures for all $t \geq 0$, we can further deduce that $c(t) \equiv 1$ and conclude that $\overline{\mu}_t^1(dx) = \overline{\mu}_t^2(dx)$ in $C^1([0, 1])^*$ and in $\mathcal{M}([0, 1])$ by density of continuous test-functions in $C^1([0, 1])$. $\qquad \square$

B Properties of density steady state solutions

In this section, we discuss additional properties of steady state solutions to Eq. (1.2) for the PD case. In Sect. B.1, we characterize the properties of weak steady-state solutions and show that such distributions can only consist of combinations of delta-concentrations at equilibria of the within-group dynamics and densities supported on non-equilibrium levels of cooperation that are strong solutions to the steady-state ODE. In Sect. B.2, we strengthen this observation by showing that the only steady states that can be achieved as the long-time behavior for solutions of the generalized PD case of Eq. (1.2) are the delta functions $\delta(x)$ and $\delta(x - 1)$ concentrated at the all-defector and all-cooperator equilibria, as well as the family of steady state densities $p_\theta^\lambda(x)$ described in Sect. 2.

In Sect. B.3, we prove Propositions 6 and 7, showing that, in the limit of strong between-group competition, the steady state densities $p_\theta^\lambda(x)$ concentrate upon points at which the group-level replication rate is equal to $G(1)$, the collective replication rate of the all-cooperator group. For the special case of within-group and between-group replication rates arising from the payoff matrix of a PD game in which group payoff is maximized by an intermediate level of cooperation, this means that the population will always feature less cooperation than in collectively optimal. This concentration upon a delta-measure at a suboptimal level of cooperation strengthens observations on the shadow of lower-level selection derived in prior work, in which it was shown that the modal level of cooperation at steady state featured a suboptimal level of cooperation in the limit of infinite between-group selection strength (Cooney 2019b, 2020a; Cooney et al. 2021).

B.1 Possible steady state solutions of multilevel dynamics

One aspect of the long-term behavior we would like to characterize is the conditions under which solutions to the measure-valued dynamics $\overline{\mu}_t$ converge to weak solutions $\mu(dx)$ of the steady-state ODE given by

$$\int_0^1 \frac{dv(x)}{dx} x(1-x)\pi(x)\mu(dx) = \int_0^1 v(x)\left[G(x) - \int_0^1 G(y)\mu(dy)\right]\mu(dx),$$
(B.1)

and, furthermore, when such time-independent solutions to Eq. B.1 have corresponding densities $f(x)$ that are strong solutions to the steady-state ODE given by

$$-\frac{\partial}{\partial x}[x(1-x)\pi(x)f(x)] = \lambda f(x)\left[G(x) - \int_0^1 G(y)f(y)dy\right].$$
(B.2)

In Lemma 12, we show that if solutions $\overline{\mu}_t(dx)$ converges to a limit $\mu_\infty(dx)$ as $t \to \infty$, then the limit must be a strong solution to the steady state ODE at all non-equilibrium points of the within-group dynamics. Under the PD dynamics, this means that probability mass can accumulate at the within-group equilibrium at the endpoints 0 or 1, and that either the steady state $\mu_\infty(dx)$ vanishes identically on $(0, 1)$ or is everywhere nonzero on $(0, 1)$.

Lemma 12 *If a weak solution $\overline{\mu}_t(dx)$ to Eq. (1.2) converges weakly to a limit function $\mu_\infty(dx)$ as $t \to \infty$, then the limit $\mu_\infty(dx)$ is a weak solution to the steady state ODE in Eq. (B.1). Furthermore, denoting the set of points $E = \{x \in (0, 1) : \pi(x) \neq 0\}$, we see that there $\mu_\infty(dx)$ has corresponding density $f_\infty(x)$ for $x \in E$, which solves the strong form of the steady state ODE given in Eq. (B.2). For any interval $\mathcal{I} \subset E$, either $f_\infty(x) > 0$ for all $x \in \mathcal{I}$ or $f_\infty \equiv 0$ for all $x \in \mathcal{I}$.*

Proof of Lemma 12 Using the measure-valued formulation of the multilevel dynamics, we integrate Eq. (1.5) in time from T to $T + 1$ and obtain

$$\int_0^1 v(x)\overline{\mu}_{T+1}(dx) - \int_0^1 v(x)\overline{\mu}_T(dx) = -\int_T^{T+1}\int_0^1 \frac{\partial v(x)}{\partial x}x(1-x)\pi(x)\overline{\mu}_t(dx)dt$$

$$+ \lambda \int_T^{T+1}\int_0^1 v(x)\left[G(x) - \int_0^1 G(y)\overline{\mu}_t(dy)\right]\overline{\mu}_t(dx)dt. \tag{B.3}$$

Because $\overline{\mu}_t(dx)$ converges weakly to $\mu_\infty(dx)$ by assumption, we know that, for any C^1 test function $V(x)$, $\int_0^1 V(x)\overline{\mu}_t dx \to \int_0^1 V(x)\mu_\infty(dx)$ as $t \to \infty$. In particular, this tells us that the lefthand side of Eq. (B.3) vanishes in the limit as $T \to \infty$. Furthermore, because $G(\cdot)$, $\pi(\cdot)$, and $v(\cdot)$ are C^1 functions, we see that the following limits hold as $t \to \infty$

$$\int_0^1 \frac{\partial v(x)}{\partial x}x(1-x)\pi(x)\overline{\mu}_t(dx) \to \int_0^1 \frac{\partial v(x)}{\partial x}x(1-x)\pi(x)\mu_\infty(dx)$$

$$\int_0^1 v(x)\left[G(x) - \int_0^1 G(y)\overline{\mu}_t(dx)\right]\overline{\mu}_t(dx) \to$$

$$\int_0^1 v(x)\left[G(x) - \int_0^1 G(y)\mu_\infty(dy)\right]\mu_\infty(dx).$$

Therefore we can take the limit as $T \to \infty$ on both sides of Eq. (B.3), which allows us to see that the limiting measure $\mu_\infty(dx)$ is a weak solution to the steady state ODE from Eq. (B.1).

Using the shorthand $Q(dx) = x(1-x)\pi(x)\mu_\infty(dx)$, we can rewrite Eq. (B.1) to see that

$$\int_0^1 \frac{\partial v(x)}{\partial x}Q(dx) = \lambda \int_0^1 v(x)\left[G(x) - \int_0^1 G(y)\mu_\infty(dy)\right]\mu_\infty(dx). \tag{B.4}$$

Because Eq. (B.4) tells us that the distributional derivative of $Q(dx)$ is a finite (signed) measure, we can deduce that there is a function $q(x)$ such that $Q(dx) = q(x)dx$ that has bounded variation, and furthermore that $q(x) \in L^1$ ([0, 1]) (Evans and Gariepy 2015). Then, for any closed interval $\mathcal{J} \subset E$ such that $\min_{x \in \mathcal{J}} |\pi(x)| > 0$, we can deduce that there is a density $f_\infty(dx)$ such that

$$f_\infty(x) = \frac{q(x)}{x(1-x)\pi(x)} \in L^1(\mathcal{J}).$$

Then, restricting ourselves to test-functions with support contained in $\mathcal{J}$, we can rewrite Eq. (B.4) as

$$\int_{\mathcal{J}} \frac{\partial v(x)}{\partial x}q(x)dx = \int_{\mathcal{J}} v(x)\left[G(x) - \int_0^1 G(y)f_\infty(y)dy\right]f_\infty(x)dx. \tag{B.5}$$

Because $f_\infty(x)$ is integrable on $\mathcal{J}$, we see from the righthand side of Eq. (B.5) that $q(x)$ has an integrable weak derivative, and therefore $q(x)$ and $f_\infty(x)$ are absolutely continuous. Applying this again to Eq. (B.5) tells us that $q(x)$ has a continuous weak derivative, and therefore $q(x), f_\infty(x) \in C^1(\mathcal{J})$. Furthermore, this means that $f_\infty(x)$ is actually a strong solution to the steady state ODE of Eq. (B.2) for $x \in \mathcal{J}$. We can then extend our definition of the interval $\mathcal{J}$ as needed to show that this also holds for any $x \in E$.

To discuss positivity, we can rewrite the steady-state relation Eq. (B.2) for $f_\infty(x)$ as

$$\frac{\partial f_\infty(x)}{\partial x} = -\left(\frac{1}{x(1-x)\pi(x)}\right)$$
$$\times \left(\frac{\partial}{\partial x}[x(1-x)\pi(x)] + \lambda\left[G(x) - \int_0^1 G(y)f_\infty(y)dy\right]\right) f_\infty(x). \tag{B.6}$$

This ODE has a unique solution on any interval $\mathcal{J}$ on which $\min_{x \in \mathcal{J}} |\pi(x)| > 0$, as the righthand side is Lipschitz in $f_\infty(x)$ and is continuous in x away from the equilibria of the within-group dynamics. Consequently, if $f_\infty(x) = 0$ for an $x \in \mathcal{J}$, then we can see from Eq. (B.6) that $\frac{\partial f_\infty(x)}{\partial x} = 0$, and therefore $f_\infty(x)$ is identically 0 on $\mathcal{J}$. This allows us to conclude that $f_\infty(x)$ is either strictly positive or identically 0 on any connected component of the set E of points in $(0, 1)$ that are not equilibria of the within-group dynamics. $\qquad \square$

B.2 Achievable long-time steady states for PD dynamics

Lemma 12 tells us that the only possible steady states of Eq. (1.2) are convex combinations delta-functions supported at equilibria of the within-group equilibria and densities that are strong solutions to the steady state ODE on intervals between within-group equilbria. As a first step to exploring which steady states can actually be achieved through the long-time behavior of Eq. (1.2) for the PD case, we can consider initial measures of the form

$$\mu_0(dx) = a_0\delta(x) + a_1\delta(1-x) + (1 - a_0 - a_1)\, p_\theta^\lambda(x)dx \tag{B.7}$$

for $a_0, a_1 \geq 0$ satisfying $a_0 + a_1 \leq 1$ and $p_\theta^\lambda(x)$ given by Eq. (1.22). In Proposition 5, we characterize the long-time behavior of solutions $\overline{\mu}_t(dx)$ to Eq. (1.2) for such initial measures, and we see that the only possible long-time steady states are $\delta(x)$, $\delta(1-x)$, and $p_\theta^\lambda(x)$ if there is any between-group competition (i.e. when $\lambda > 0$).

Proposition 5 *Suppose that $\lambda[G(1) - G(0)] > \pi(1)\theta$ and that the population has initial measure $\mu_0(dx)$ given by Eq. (B.7). If $a_1 > 0$, then the solution $\mu_t(dx)$ to Eq. (1.2) satisfies $\overline{\mu}_t(dx) \rightharpoonup \delta(1-x)$ as $t \to \infty$. If $a_1 = 0$ and $a_0 < 1$, then we have instead that $\overline{\mu}_t(dx) \rightharpoonup p_\theta^\lambda(x)$.*

Proof Noting from Proposition 4 that solutions $\mu_t(dx)$ to Eq. (1.6) are unique, we can check the solution $\mu_t(dx)$ corresponding to the initial measure of Eq. (B.7) is given by

$$\mu_t(dx) = a_0 e^{\lambda G(0)t}\delta(x) + a_1 e^{\lambda G(1)t}\delta(1-x) + (1 - a_0 - a_1)\, e^{[\lambda G(1) - \theta\pi(1)]t}\, p_\theta^\lambda(x)dx.$$
(B.8)

We can use Eq. (B.8) and the normalization relation from Eq. (1.7) to further see that solutions $\overline{\mu}_t(dx)$ to the full multilevel dynamics satisfy

$$
\begin{aligned}
\overline{\mu}_t(dx) &= \frac{e^{-\lambda G(1)t}\mu_t(dx)}{e^{-\lambda G(1)t}\int_0^1 \mu_t(dy)} \\
&= \frac{a_0 e^{\lambda[G(0) - G(1)]t}\delta(x) + a_1\delta(1-x) + (1 - a_0 - a_1)\, e^{-\theta\pi(1)t}\, p_\theta^\lambda(x)dx}{a_0 e^{\lambda[G(0) - G(1)]t} + a_1 + (1 - a_0 - a_1)\, e^{-\theta\pi(1)t}}.
\end{aligned}
$$
(B.9)

Then, using the fact that that $G(1) > G(0)$ and $\theta\pi(1) > 0$, we can further see that $\overline{\mu}_t(dx) \rightharpoonup \delta(1-x)$ as long as $a_1 > 0$. If, instead, $a_1 = 0$, we can see that

$$
\begin{aligned}
\overline{\mu}_t(dx) &= \frac{e^{[\theta\pi(1) - \lambda G(1)]t}\mu_t(dx)}{e^{[\theta\pi(1) - \lambda G(1)]t}\int_0^1 \mu_t(dy)} \\
&= \frac{a_0 e^{(\theta\pi(1) - \lambda[G(1) - G(0)])t}\delta(x) + (1 - a_0)\, p_\theta^\lambda(x)dx}{a_0 e^{(\theta\pi(1) - \lambda[G(1) - G(0)])t} + (1 - a_0)}.
\end{aligned}
$$
(B.10)

Because $\lambda[G(1) - G(0)] > \theta\pi(1)$ by assumption, we can use the expression in Eq. (B.10) to conclude that $\overline{\mu}_t(dx) \rightharpoonup p_\theta^\lambda(x)dx$ as $t \to \infty$. □

B.3 Concentration of steady-state densities in the limit of infinite between-group competition

In this section, we study the behavior of the steady state densities $p_\theta^\lambda(x)$ in the limit of infinite strength of between-group competition. We characterize the concentration of the steady state densities to measures supported upon values x at which $G(x) = G(1)$, and therefore concentration at levels of cooperation that are not necessarily optimal for group-level replication.

To highlight the dependence of our steady-state density on λ, we can rewrite the expression of our steady state $p_\theta^\lambda(x)$ from Eq. (1.22) in the form

$$p_\theta^\lambda(x) = \frac{b(x)\exp(\lambda h(x))}{\int_0^1 b(y)\exp(\lambda h(y))\, dy},$$
(B.11)

where $b(x)$ and $h(x)$ are given by the formulas

$$b(x) := (1 - x)^{\theta - 1} \frac{\pi(1)}{x\pi(x)} \exp\left(\int_x^1 \frac{\theta\,[\pi(1) - s\pi(s)]}{s\pi(s)} ds\right) \quad \text{(B.12a)}$$

$$h(x) := \int_x^1 \left[\frac{G(s) - G(1)}{s(1 - s)\pi(s)}\right] ds. \quad \text{(B.12b)}$$

From the form of Eq. (B.11), we expect $p_\theta^\lambda(x)$ to concentrate around the global maximizer of $h(x)$ as $\lambda \to \infty$. The critical points x_c of $h(x)$ satisfy $G(x_c) = G(1)$, and we further compute that

$$\left. h''(x) \right|_{x=x_c} = \left. \frac{-x(1 - x)\pi(x)G'(x) - (x(1 - x)\pi(x))'\,[G(1) - G(x)]}{x^2(1 - x)^2\pi(x)^2} \right|_{x=x_c}$$

$$= \frac{-G'(x_c)}{x_c(1 - x_c)\pi(x_c)}.$$

Because $\pi(x) > 0$ for $x \in [0, 1]$ under the PD assumptions, we see that the local maxima of $h(x)$ are upcrossings of $G(1)$, and the local minima of $h(x)$ are downcrossings of $G(1)$.

Example 1 As an example of the possible collective optima and upcrossings of $G(1)$ that can occur in our multilevel dynamics, we can turn to the example of the game-theoretic model of Sect. (1.1.1). We recall from Eq. (1.17) that, for dynamics arising from games with the payoff matrix of Eq. (1.14), that the group-level replication rate $G(x)$ is given by the quadratic function

$$G(x) = P + (S + T - 2P)x + (R - S - T + P)x^2.$$

In this case, $G(x)$ is maximized by the following level of cooperation

$$x^* = \begin{cases} \dfrac{S + T - 2P}{-2(R - S - T + P)} & : 2R < S + T,\ R - S - T + P < 0 \\ 1 & : \text{otherwise} \end{cases}, \quad \text{(B.13)}$$

so an intermediate level of cooperation can optimize group-level reproduction when the total payoff $S+T$ generated by the interaction of a cooperator and defector exceeds the total payoff $2R$ generated by two cooperators. In addition, we see from the fact that $G(x)$ is a quadratic function of x in the game-theoretic setting that $G'(x)$ changes sign at most once in $[0, 1]$, and that $G(x)$ experiences a single upcrossing of $G(1)$ in $[0, 1]$. This upcrossing occurs at the level of cooperation $\bar{x}$ given by

$$\bar{x} = \begin{cases} \dfrac{R - P}{-(R - S - T + P)} & : 2R < S + T,\ R - S - T + P < 0 \\ 1 & : \text{otherwise} \end{cases}, \quad \text{(B.14)}$$

so the upcrossing occurs in the interior for the same conditions in which the collective optimum x^* is achieved by an intermediate level of cooperation. By comparing Eqs, (B.13) and (B.14), we see that $\overline{x} < x^*$ whenever the collective optimum features a mix of cooperators and defectors.

For the generalization of the PD and HD dynamics studied in this paper with continuously differentiable replication rates, $G(x)$ can either be maximized by interior levels of cooperation or by full-cooperation and can feature arbitrarily many upcrossings of $G(1)$. As a result, in the broader class of models, it is possible for $h(x)$ to be maximized at $x = 1$ even when $G(x)$ is maximized at an interior level of cooperation.

In Proposition 6, we consider the case in which $h(x)$ is maximized by an interior level of cooperation $\overline{x} \in (0, 1)$, and use the Laplace integration method (Bender and Orszag 1999) to show that $p_\theta^\lambda(x)$ concentrates upon $\overline{x}$ as $\lambda \to \infty$.

Proposition 6 *Consider the steady state densities given by Eq. (B.12) under the assumptions that $G(x), \pi(x) \in C^1([0, 1])$, $G(1) > G(0)$ and $\pi(x) > 0$ for $x \in [0, 1]$. Suppose that $G(x)$ has an interior maximizer x^* satisfying $G(x^*) > G(1)$, and that $h(x)$ has a unique maximizer $\overline{x} < 1$. Then the family of steady states $p_\theta^\lambda(x) \rightharpoonup \delta(x - \overline{x})$ as $\lambda \to \infty$.*

Furthermore, if $\overline{x}$ is the only upcrossing of $G(1)$ in $[0, 1]$, then $\overline{x} < x^*$ and the level of cooperation achieved as $\lambda \to \infty$ is less than the level that achieves the maximal group reproduction rate $G(x^*)$. In particular, we note from Example 1 that this is true for all PD and HD games with corresponding quadratic $G(x)$ given by Eq. (1.17).

When $h(x)$ is maximized by $\overline{x} = 1$, the Laplace method breaks down because of the behavior of $b(x)$ at $x = 1$, so we employ a different approach to show that $p_\theta^\lambda(x)$ concentrates upon full-cooperation as $\lambda \to \infty$ under the additional assumption that $\theta > 1$.

Proposition 7 *Consider the steady state densities given by Eq. (B.12) under the assumptions on $G(x)$ and $\pi(x)$ from Proposition 6. Suppose $h(x)$ achieves a unique maximum at $x = 1$ and consider $\theta > 1$. Then, for our family of steady-state solutions $p_\theta^\lambda(x) \rightharpoonup \delta(x - 1)$ as $\lambda \to \infty$.*

In particular, $h(x)$ has a global maximum at $x = 1$ when $G(x)$ is non-decreasing on $[0, 1]$, and therefore the population concentrates upon the optimal level of cooperation as $\lambda \to \infty$ when full-cooperation is the best possible group composition. However, there exist group reproduction functions $G(x)$ for which $h(x)$ is still maximized by full-cooperation even though intermediate levels of cooperation are collectively-optimal. In such cases, Proposition 7 tells us the population can concentrate at a level of cooperation greater than is optimal for group-level reproduction, so the shadow of lower-level selection can also manifest itself by promoting too much cooperation.

Proof of Proposition 6 We start by integrating both sides of Eq. B.11 against a test-function $v(x)$, obtaining

$$\int_0^1 v(x) p_\theta^\lambda(x) dx = \int_0^1 v(x) \frac{b(x) \exp(\lambda h(x))}{\int_0^1 b(y) \exp(\lambda h(y)) \, dy} dx = \frac{\int_0^1 v(x) b(x) \exp(\lambda h(x)) \, dx}{\int_0^1 b(x) \exp(\lambda h(x)) \, dx}.$$

We can further rearrange this expression to obtain

$$\int_0^1 v(x)p_\theta^\lambda(x)dx = v(\overline{x}) \left(\frac{\sqrt{\frac{2\pi}{\lambda|h''(\overline{x})|}}b(\overline{x})\exp(\lambda h(\overline{x}))}{\int_0^1 b(y)\exp(\lambda h(y))\,dy} \right) \left(\frac{\int_0^1 v(y)b(y)\exp(\lambda h(y))\,dy}{v(\overline{x})\sqrt{\frac{2\pi}{\lambda|h''(\overline{x})|}}b(\overline{x})\exp(\lambda h(\overline{x}))} \right)$$

(B.15)

Because $b(x)$ is continuous and nonzero at $\overline{x} < 1$, we can use the interior critical-point case of the Laplace method (Bender and Orszag 1999) to obtain the following asymptotic formulas for the two terms in parenthesis

$$\lim_{\lambda\to\infty} \left[\frac{\sqrt{\frac{2\pi}{\lambda|h''(\overline{x})|}}b(\overline{x})\exp(\lambda h(\overline{x}))}{\int_0^1 b(y)\exp(\lambda h(y))\,dy} \right] = 1 \text{ and}$$

$$\lim_{\lambda\to\infty} \left[\frac{\sqrt{\frac{2\pi}{\lambda|h''(\overline{x})|}}v(\overline{x})b(\overline{x})\exp(\lambda h(\overline{x}))}{\int_0^1 v(y)b(y)\exp(\lambda h(y))\,dy} \right] = 1.$$

(B.16)

Taking the limit as $\lambda \to \infty$ in Eq. (B.15), we can apply the asymptotic formulas from Eq. (B.16) to see that

$$\lim_{\lambda\to\infty} \int_0^1 v(x)p_\theta^\lambda(x)dx = v(\overline{x}),$$

and we conclude that $p_\theta^\lambda(x) \rightharpoonup \delta(x - \overline{x})$ as $\lambda \to \infty$. $\qquad\square$

Proof of Proposition 7 We assume, for contradiction, that there is a test-function $v(x) \in C^1$ such that $\int_0^1 v(x)p_\theta^\lambda(x)dx \not\to v(1)$. In that case, there exists an $\epsilon > 0$ such that for any $\Lambda > 0$, there is a $\lambda > \Lambda$ for which

$$\epsilon < \left| \int_0^1 \psi(x)p_\theta^\lambda(x)dx - v(1) \right| \le \int_0^1 |v(x) - v(1)|p_\theta^\lambda(x)dx, \qquad \text{(B.17)}$$

where the second inequality follows because $p_\theta^\lambda(x)$ is a probability density. Because $v(\cdot)$ is continuous, there is $\delta > 0$ such that $|v(x) - v(1)| < \frac{\epsilon}{2}$ for $x \in [1 - \delta, 1]$, so we can further estimate that

$$\left| \int_0^1 v(x)p_\theta^\lambda(x)dx - v(1) \right| \le \frac{\epsilon}{2} + \int_0^{1-\delta} |v(x) - v(1)|p_\theta^\lambda(x)dx$$

$$< \frac{\epsilon}{2} + 2||v||_\infty \int_0^{1-\delta} p_\theta^\lambda(x)dx. \qquad \text{(B.18)}$$

Combining the results of Eqs. B.17 and B.18 , we see that there are $\epsilon, \delta > 0$ such that for any $\Lambda > 0$, there exists $\lambda > \Lambda$ for which

$$\int_0^{1-\delta} p_\theta^\lambda(x)dx > \frac{\epsilon}{4||v||_\infty} > 0. \tag{B.19}$$

For such λ, we can then consider the steady state probability $\int_0^{1-\delta} p_\theta^\lambda(x)dx$ found on the interval $[0, 1-\delta]$. Using Eq. (B.11) and the fact that $b(x) \geq 0$, we can estimate that, for any $\Delta > 0$,

$$\int_0^{1-\delta} p_\theta^\lambda(x) = \frac{\int_0^{1-\delta} b(x)e^{\lambda h(x)}dx}{\int_0^1 b(x)e^{\lambda h(x)}dx} \leq \frac{\int_0^{1-\delta} b(x)e^{\lambda h(x)}dx}{\int_{1-\Delta}^1 b(x)e^{\lambda h(x)}dx}. \tag{B.20}$$

Because we have assumed that $h(x)$ has a global maximum at $x = 1$ and that $G(x)$ is a C^1 function, we know that $h(x)$ is locally non-decreasing as $x \to 1^-$. Therefore, choosing sufficiently small $\Delta > 0$, we see that there is an $A > 0$ such that

$$\int_{1-\Delta}^1 b(x)e^{\lambda h(x)}dx \geq e^{\lambda h(1-\Delta)} \int_{1-\Delta}^1 b(x)dx \geq Ae^{\lambda h(1-\Delta)} \tag{B.21}$$

Turning to the numerator of Eq. B.20, we can use the fact that $(1-x)^{\theta-1} \leq 1$ for $\theta > 1$ to see that the integrand in the numerator satisfies

$$b(x)e^{\lambda h(x)} \leq \frac{1}{x\pi(x)} \exp\left(\lambda h(x) + \theta \int_x^1 \frac{[\pi(1) - s\pi(s)]}{s\pi(s)}ds\right)$$
$$= \frac{1}{x\pi(x)} \exp\left(\int_x^1 \left\{\frac{1}{s\pi(s)}\left[\frac{\lambda[G(s) - G(1)] + \theta[\pi(1) - s\pi(s)]}{1-s}\right]\right\}ds\right). \tag{B.22}$$

In particular, we note that when $s = 0$, the term in square brackets takes the value $\lambda[G(0) - G(1)] + \theta\pi(1)$, which is negative when steady state densities of the form $p_\theta^\lambda(x)$ exist. Because $G(x), \pi(x) \in C^1[0, 1]$, the integrand on the last line of Eq. (B.22) is negative for x close enough to 0. In addition, we can deduce that $b(x)e^{\lambda h(x)} \to 0$ as $x \to 0$ for sufficiently large λ due to Eqs. (4.1), (4.2), (B.11), and (B.12). Combining these observations, we know there is d sufficiently close to 1 and $M_1, M_2 < \infty$ such that

$$b(x)e^{\lambda h(x)} \leq M_1 \exp\left(\lambda h(\max(x, 1-d)) + \theta \int_{\max(x,1-d)}^1 \frac{[\pi(1) - s\pi(s)]}{s\pi(s)}ds\right)$$
$$\leq M_2 \exp\left(\lambda \max_{x \in [1-d, 1-\delta]} h(x)\right). \tag{B.23}$$

Further denoting $\hat{h}_{d,\delta} = \max_{x\in[1-d,1-\delta]} h(x)$, we can combine the estimates from Eqs. (B.20), (B.21), and (B.23) to see that

$$\int_0^{1-\delta} p_\theta^\lambda(x)dx \le MA^{-1}(1-\delta)\exp\left(\lambda\left[\hat{h}_{d,\delta} - h(1-\Delta)\right]\right). \quad \text{(B.24)}$$

Because $h(x)$ has a unique global maximum at $x = 1$, there is a $\Delta^* > 0$ such that, for any $\Delta < \Delta^*$, $h(1 - \Delta) > h(x)$ for $x \in [0, 1 - \Delta^*)$. Namely, choosing $\Delta < \min(\Delta^*, \delta)$ allows us to additionally deduce that $h(1 - \Delta) > \hat{h}_{d,\delta}$ and conclude that

$$\int_0^{1-\delta} p_\theta^\lambda(x)dx \to 0 \ \text{ as } \lambda \to \infty, \quad \text{(B.25)}$$

contradicting the existence of a subsequence of λ-values tending to infinity on which the inequality of Eq. (B.19) holds. Therefore we can conclude that $p_\theta^\lambda(x) \rightharpoonup \delta(x - 1)$ as $\lambda \to \infty$. □

C Application of multipopulation framework to analyze generalization of Hawk-Dove and stag-hunt dynamics

In this section, we consider the long-time behavior of solutions to Eq. (1.2) when the individual-level and group-level replication rates resemble the dynamics of Hawk-Dove (HD) or Stag-Hunt (SH) games. We do this by decomposing the distribution of groups playing these games into two subpopulations featuring levels of cooperation above and below the interior within-group equilibrium x_{eq}, and then study the evolution of these two conditional distributions using a two-population version of the multipopulation dynamics studied in Sect. 6. In Sect. C.1, we formulate the individual and group replication rates for the two subpopulations considered for the HD and SH games, and, in Sect. C.2, we present the results for the long-time behavior of the multilevel HD and SH dynamics in light of Theorem 4.

C.1 Assumptions HD and SH dynamics and formulation as a 2-population scenario

To generalize the multilevel Hawk-Dove and Stag-Hunt dynamics, we adapt our assumptions on our replication rates $\pi(x)$ and $G(x)$ to reflect the properties of the payoff rankings for these two games. In particular, we see from Eqs. 1.19 and 1.20 that $\pi(x) > 0$ for $x \in (x_{eq}, 1]$ and $\pi(x) < 0$ for $x \in [0, x_{eq})$ for HD games, while the opposite signs hold for the SH game.

For the group payoff function, we can use the rankings of payoffs for the HD and SH games and Eq. 1.20 to see that

$$G(1) - G(x_{eq}) = \frac{(R - S)(R - T)}{R - S - T + P} > 0$$

$$G(0) - G(x_{eq}) = \frac{(S - P)(T - P)}{R - S - T + P} < 0,$$

$$(\text{C.1})$$

and therefore we the group reproduction function satisfies $G(0) < G(x_{eq}) < G(1)$ for both games.

To reformulate the dynamics of the HD and SH games in terms of the multipopulation dynamics studied in Sect. 6, it is helpful to try to understand ODE of Eq. (1.4) in comparison to a logistic ODE supported on one of the intervals $[0, x_{eq})$ or $(x_{eq}, 1]$. We do this by rewriting Eq. (1.4) in the form

$$\frac{dx(t)}{dt} = -x(1 - x)(x - x_{eq})\left(\frac{\pi(x)}{x - x_{eq}}\right) , \quad x(0) = x_0. \tag{C.2}$$

Next, we can map the intervals $[0, x_{eq})$ or $(x_{eq}, 1]$ into $[0, 1]$ using the rescaled variables

$$X_1 = \frac{x}{x_{eq}} \quad . \quad X_2 = \frac{x - x_{eq}}{1 - x_{eq}}. \tag{C.3}$$

Then we can describe the respective within-group dynamics below and above the equilibrium x_{eq} using the modified gain functions

$$\Pi_1(X_1) := \begin{cases} \dfrac{(1 - x_{eq}X_1)\pi(x_{eq}X_1)}{1 - X_1} : X_1 \in [0, 1) \\ -x_{eq}(1 - x_{eq})\pi'(x_{eq}) : X_1 = 1 \end{cases}$$

$$\Pi_2(X_2) := \begin{cases} \dfrac{(x_{eq} + (1 - x_{eq})X_2)\pi(x_{eq} + (1 - x_{eq})X_2)}{X_2} : X_2 \in (0, 1] \\ x_{eq}(1 - x_{eq})\pi'(x_{eq}) : X_2 = 0 \end{cases}.$$

$$(\text{C.4})$$

Further introducing the measures $\overline{\mu}_t^1(dX_1) = 1_{X_1 \in [0,1]}\overline{\mu}_t(dX_1)$ and $\overline{\mu}_t^2(dX_2) = 1_{X_2 \in [0,1]}\overline{\mu}_t(dX_2)$ as well as the modified group-reproduction functions

$$G_1(X_1) = G(x_{eq}X_1) \tag{C.5a}$$

$$G_2(X_2) = G\left(x_{eq} + (1 - x_{eq})X_2\right), \tag{C.5b}$$

we can see that these measures $\overline{\mu}_t^1(dX_1)$ and $\overline{\mu}_t^2(dX_2)$ change according to Eq. (1.10) for our choices of modified gain and group-reproduction functions. To guarantee that modified gain functions $\Pi_1(X_1), \Pi_2(X_2) \in C^1([0, 1])$ in line with the assumption of Theorem 4, we see from Eq. (C.4) that it suffices to assume that the original gain function satisfies $\pi(x) \in C^2([0, 1])$.

C.2 Summary of long-time behavior for HD and SH dynamics

We start with the case of the Stag-Hunt game. Because $\pi(x) > 0$ for $x \in (0, x_{eq})$ and $\pi(x) < 0$ for $x \in (x_{eq}, 1)$ for the SH game, we see that $r_2^m = \lambda G_2(1) = \lambda G(1)$, while $r_1^M \leq \lambda G_1(1) = G(x_{eq})$. Therefore the two-population representation of the SH game will always satisfy the hypothesis of Theorem 4 regarding a dominant subpopulation, and so the support groups below x_{eq} vanishes in the long-time limit. Because the dynamics above x_{eq} resemble a PDel game, we can use an approach inspired by the proof of Proposition 1 to show fixation upon full-cooperation.

Proposition 8 *Suppose that $G(x) \in C^1([0, 1])$, $\pi(x) \in C^2([0, 1])$, $G(0) < G(x_{eq}) < G(1)$, $\pi(x)$ has a single root $x_{eq} \in (0, 1)$, and that $\pi'(x_{eq}) > 0$. We further assume that the initial distribution contains groups with levels of cooperation exceeding the within-group equilibrium, i.e. $\mu_0\left(\left(x_{eq}, 1\right]\right) > 0$. If $\lambda > 0$, $\overline{\mu}_t(dx) \rightharpoonup \delta(1 - x)$ as $t \to \infty$.*

This generalizes a previous result for SH games with the payoff matrices of Eq. (1.14) (Cooney 2020a), whose proof relied on the fact that the quadratic $G(x)$ of Eq. (1.17b) is increasing for $x > x_{eq}$ under the SH payoff rankings. By contrast, Proposition 8 only requires the ranking of the values group-reproduction function at the equilibria of the within-group dynamics.

We now turn to the Hawk-Dove game and consider initial measures with infimum and supremum Hölder exponents near $x = 1$ satisfying $\infty > \overline{\theta} \geq \underline{\theta} > 0$. Because $\pi(x) < 0$ for $x \in (0, x_{eq})$ and $\pi(x) > 0$ for $x \in (x_{eq}, 1)$, we can find that the principal growth rates on our two intervals are given by $r_1^M = \lambda G(x_{eq})t$ and $r_2^m = \max\{\lambda G(x_{eq}), \lambda G(1) - \overline{\theta}\pi(1)\}$. We can show for either possible value of r_2^m that the probability $\overline{\mu}_t\left([0, x_{eq})\right)$ vanishes as $t \to \infty$, with the case of $r_2^m = \lambda G(1) - \overline{\theta}\pi(1)$ following from Theorem 4 and the case of $r_2^m = \lambda G(x_{eq})$ requiring an argument analogous to the proof of Proposition 1. We can further characterize the long-time behavior of the multilevel HD dynamics by analyzing the distribution of groups above x_{eq} in the same way we studied the PD dynamics in previous sections.

We show in Theorem 5 that the population concentrates upon a delta-function $\delta(x - x_{eq})$ at the within-group equilibrium when $\lambda\left[G(1) - G(x_{eq})\right] < \underline{\theta}\pi(1)$, while we show in Theorem 6 that the fraction of cooperators in the population exceeds the equilibrium level x_{eq} infinitely often when $\lambda\left[G(1) - G(x_{eq})\right] > \overline{\theta}\pi(1) > \underline{\theta}\pi(1)$. Under the stronger assumption that the initial measure has a well-defined positive, finite Hölder exponent and constant near $x = 1$, we show in Theorem 7 that the population converges to a steady state density whose support consists of groups with fractions of cooperation between x_{eq} and 1.

Theorem 5 *Suppose that $G(x) \in C^1([0, 1])$, $\pi(x) \in C^2([0, 1])$, $G(0) < G(x_{eq}) < G(1)$. We further assume that the initial distribution $\mu_0(dx)$ has supremum Hölder exponent θ near $x = 1$, that $\pi(x)$ has a single root $x_{eq} \in (0, 1)$, and that $\pi'(x_{eq}) > 0$. If $\lambda\left(G(1) - G(x_{eq})\right) < \theta\pi(1)$, then $\mu_t(dx) \rightharpoonup \delta(x - x_{eq})$.*

Theorem 6 *Suppose that $G(x)$ and $\pi(x)$ satisfy the assumptions of Theorem 5 and that the initial distribution $\overline{\mu}_0(dx)$ has nonzero infimum and supremum Hölder exponents $\overline{\theta}$ and $\underline{\theta}$ near $x = 1$ with corresponding Hölder constants $C_{\overline{\theta}}$ and $C_{\underline{\theta}}$ that are finite and nonzero. If $\lambda\left[G(1) - G(x_{eq})\right] > \overline{\theta}\pi(1) > \underline{\theta}\pi(1)$, then*

$$\limsup_{t \to \infty} \int_{x_{eq}}^1 x \overline{\mu}_t(dx) > x_{eq}. \tag{C.6}$$

Now we consider the case of convergence to steady state for the multilevel HD dynamics. When the initial measure has Hölder exponent $\theta > 0$ near $x = 1$ and the strength of between-group satisfies $\lambda \left[G(1) - G(x_{eq}) \right] > \theta \pi(1)$, the principal growth rates r_1^M and r_2^m associated with the measures $\overline{\mu}_t^1(dX_1)$ and $\overline{\mu}_t^2(dX_2)$ satisfying $r_2^m = \lambda G(1) - \theta \pi(1) > \lambda G(x_{eq}) = r_1^M$. As a consequence, we can apply Theorem 4 to conclude that the probably of groups below x_{eq} satisfies $\overline{\mu}_t^1([0, 1]) \to 0$ as $t \to \infty$. For the distribution of groups above x_{eq}, Theorem 4 tells us that $\overline{\mu}_t^2(dX_2) \rightharpoonup p_\theta^{\lambda,2}(X_2)$, where $p_\theta^{\lambda,2}(X_2)$ is a steady state density given by

$$p_\theta^{\lambda,2}(X_2) = \frac{f_\theta^{\lambda,2}(X_2)}{\int_0^1 f_\theta^{\lambda,2}(y)dy} \tag{C.7a}$$

$$f_\theta^{\lambda,2}(X_2) = X_2^{v_2-1}(1 - X_2)^{\theta-1} \left(\frac{\Pi_2(1)}{\Pi_2(X_2)} \right) \exp\left(\int_{X_2}^1 \frac{-\lambda \tilde{C}_2(s)}{\Pi_2(s)} ds \right) \tag{C.7b}$$

$$v_2 = \frac{1}{\Pi_2(0)} \left(\lambda \left[G_2(1) - G_2(0) \right] - \theta \Pi_2(1) \right) \tag{C.7c}$$

$$-\lambda \tilde{C}_2(s) = \lambda \left(\frac{G_2(s) - G_2(0)}{s} \right) + v_2 \left(\frac{\Pi_2(s) - \Pi_2(0)}{s} \right)$$
$$+ \lambda \left(\frac{G_2(s) - G_2(1)}{1 - s} \right) - \theta \left(\frac{\Pi_2(s) - \Pi_2(1)}{1 - s} \right). \tag{C.7d}$$

We can use this representation of the steady-state conditional measure $\overline{\mu}_t(dX_2)$ for group compositions featuring more than x_{eq} cooperators, paired with fact that the probability of groups below x_{eq} vanishes in the long-time limit, to find the long-time steady-state achieved by the multilevel dynamics in the Hawk-Dove case. Applying to Eq. (C.7) the change of variables $X_2 = \frac{x - x_{eq}}{1 - x_{eq}}$ and the definitions of the modified replication rates $\Pi_2(X_2)$ and $G_2(X_2)$ from Eqs. (C.4) and (C.5), we find that the steady states of the multilevel HD dynamics take the following form:

$$q_\theta^\lambda(x) = \begin{cases} 0 & : x < x_{eq} \\ \dfrac{g_\theta^\lambda(x)}{\int_{x_{eq}}^1 g_\theta^\lambda(y)dy} & : x \geq x_{eq}, \end{cases} \tag{C.8}$$

with $g_\theta^\lambda(x)$ given by

$$g_\theta^\lambda(x) = \left(x - x_{eq} \right)^{v_H-1} (1 - x)^{\theta-1} \left(\frac{\pi(1)}{x\pi(x)} \right)$$
$$\exp\left(\int_{\frac{x-x_{eq}}{1-x_{eq}}}^1 \frac{(-\lambda/(1-x_{eq}))s\tilde{C}_H(s)}{(x_{eq} + (1-x_{eq})s)\pi(x_{eq} + (1-x_{eq})s)} ds \right) \tag{C.9}$$

and where ν_H and $-\lambda C_H(s)$ are given by

$$\nu_H = \frac{\lambda\left[G(1) - G(x_{eq})\right] - \theta\pi(1)}{x_{eq}(1 - x_{eq})\pi'(x_{eq})} \tag{C.10a}$$

$$
-\lambda\tilde{C}_H(s) = \lambda\left(\frac{G\left(x_{eq} + (1 - x_{eq})s\right) - G(x_{eq})}{s} + \frac{G\left(x_{eq} + (1 - x_{eq})s\right) - G(1)}{1 - s}\right)
$$
$$
+ \theta\left(\frac{s^{-1}\left(x_{eq} + (1 - x_{eq})s\right)\pi\left(x_{eq} + (1 - x_{eq})s\right) - \pi(1)}{1 - s}\right)
$$
$$
+ \nu_H\left(\frac{\left[x_{eq} + (1 - x_{eq})s\right]\pi\left(x_{eq} + (1 - x_{eq})s\right) - sx_{eq}(1 - x_{eq})\pi'(x_{eq})}{s^2}\right)
\tag{C.10b}
$$

In Theorem 7, we characterize the long-time weak convergence of solutions of the multilevel HD dynamics to steady states in the form described by Eqs. (C.8), (C.9), (C.10) for the case in which initial conditions have well-define Hölder data near $x = 1$ and in which between-group selection is sufficiently strong. This result confirms and generalizes (Cooney 2020a, Conjecture 2), which addresses convergence to steady-state in the case for which the replication rates arise from Hawk-Dove games with the payoff matrix of Eq. (1.14).

Theorem 7 *Suppose that $G(x)$ and $\pi(x)$ satisfy the assumptions of Theorem 5 and that the initial distribution $\mu_0(dx)$ has Hölder exponent θ near $x = 1$ with corresponding positive, finite Hölder constant C_θ. If $\lambda\left[G(1) - G(x_{eq})\right] > \theta\pi(1)$, then, for any continuous test-function $v(x)$, the solution $\overline{\mu}_t(dx)$ to Eq. (1.2) satisfies*

$$\lim_{t\to\infty}\int_0^1 v(x)\overline{\mu}_t(dx) = \int_0^1 v(x)q_\theta^\lambda(x)dx. \tag{C.11}$$

Using the same approach as in Sect. 2, we can characterize the threshold level of between-group competition required for an integrable steady state

$$\lambda_H^*(\theta) := \frac{\theta\pi(1)}{G(1) - G(x_{eq})}. \tag{C.12}$$

Then we can use this expression to see that the average payoff at steady state is given by

$$\langle G(\cdot)\rangle_f = \left(\frac{\lambda_H^*(\theta)}{\lambda}\right)G(x_{eq}) + \left(1 - \frac{\lambda_H^*(\theta)}{\lambda}\right)G(1), \tag{C.13}$$

so average payoff interpolates between $G(x_{eq})$ at $\lambda = \lambda_H^*$ and $G(1)$ as $\lambda \to \infty$. Because $G(1) > G(x_{eq})$, we see that the average group payoff increases with λ and that group payoff is limited by the average payoff of a full-cooperator group, even if group payoff $G(x)$ is maximized by an interior fraction of cooperators.

References

Akdeniz A, van Veelen M (2020) The cancellation effect at the group level. Evolution 74(7):1246–1254

Archetti M, Scheuring I (2011) Coexistence of cooperation and defection in public goods games. Evolution 65(4):1140–1148

Bach LA, Helvik T, Christiansen FB (2006) The evolution of n-player cooperation-threshold games and ess bifurcations. J Theor Biol 238(2):426–434

Bender CM, Orszag SA (1999) Advanced mathematical methods for scientists and engineers I: asymptotic methods and perturbation theory, vol 1, Springer Science & Business Media

Bijma P, Muir WM, Ellen ED, Wolf JB, Van Arendonk JA (2007) Multilevel selection 2: estimating the genetic parameters determining inheritance and response to selection. Genetics 175(1):289–299

Bijma P, Muir WM, Van Arendonk JA (2007) Multilevel selection 1: quantitative genetics of inheritance and response to selection. Genetics 175(1):277–288

Blackstone NW, Blackstone SR, Berg AT (2020) Variation and multilevel selection of sars-cov-2 Evolution

Boni MF, Nguyen TD, de Jong MD, van Doorn HR (2013) Virulence attenuation during an influenza a/h5n1 pandemic. Philosophical Transactions of the Royal Society B: Biological Sciences 368(1614):20120207

Böttcher MA, Nagler J (2016) Promotion of cooperation by selective group extinction. New J Phys 18(6):063008

Boyd R, Gintis H, Bowles S, Richerson PJ (2003) The evolution of altruistic punishment. Proc Natl Acad Sci 100(6):3531–3535

Boza G, Számadó S (2010) Beneficial laggards: multilevel selection, cooperative polymorphism and division of labour in threshold public good games. BMC Evol Biol 10(1):336

Bratus AS, Posvyanskii VP, Novozhilov AS (2017) Solutions with a bounded support promote permanence of a distributed replicator equation. Appl Anal 96(15):2652–2668

Chuang JS, Rivoire O, Leibler S (2009) Simpson's paradox in a synthetic microbial system. Science 323(5911):272–275

Cooney DB (2019a) Assortment and reciprocity mechanisms for promotion of cooperation in a model of multilevel selection. arXiv preprint arXiv:1910.06933

Cooney DB (2019) The replicator dynamics for multilevel selection in evolutionary games. J Math Biol 79(1):101–154

Cooney DB (2020) Analysis of multilevel replicator dynamics for general two-strategy social dilemma. Bull Math Biol 82:1–72

Cooney DB (2020b) PDE models of multilevel selection: The evolution of cooperation and the shadow of lower-level selection. PhD thesis, Princeton University

Cooney DB, Rossine FW, Morris DH, Levin SA (2021) A PDE model for protocell evolution and the origin of chromosomes via multilevel selection. arXiv preprint arXiv:2109.09357

Dawson DA (2018) Multilevel mutation-selection systems and set-valued duals. J. Math. Biol. p 1–84

Dwyer G, Levin SA, Buttel L (1990) A simulation model of the population dynamics and evolution of myxomatosis. Ecol Monogr 60(4):423–447

Evans LC, Gariepy RF (2015) Measure theory and fine properties of functions. CRC Press

Fontanari JF, Serva M (2013) Solvable model for template coexistence in protocells. EPL (Europhysics Letters) 101(3):38006

Fontanari JF, Serva M (2014) Effect of migration in a diffusion model for template coexistence in protocells. Bull Math Biol 76(3):654–672

Fontanari JF, Serva M (2014) Nonlinear group survival in kimura's model for the evolution of altruism. Math Biosci 249:18–26

Fontanari JF, Santos M, Szathmáry E (2006) Coexistence and error propagation in pre-biotic vesicle models: a group selection approach. J Theor Biol 239(2):247–256

Freedman HI, Moson P (1990) Persistence definitions and their connections. Proceedings of the American Mathematical Society 109(4):1025–1033

Fu F, Kocher SD, Nowak MA (2015) The risk-return trade-off between solitary and eusocial reproduction. Ecol Lett 18(1):74–84

Gilchrist MA, Coombs D (2006) Evolution of virulence: interdependence, constraints, and selection using nested models. Theor Popul Biol 69(2):145–153

Gilchrist MA, Coombs D, Perelson AS (2004) Optimizing within-host viral fitness: infected cell lifespan and virion production rate. J Theor Biol 229(2):281–288

Grafen A (1979) The hawk-dove game played between relatives. Anim Behav 27(4):905–907

Gueron S, Levin SA (1995) The dynamics of group formation. Math Biosci 128(1–2):243–264

Henriques GJ, Simon B, Ispolatov Y, Doebeli M (2019) Acculturation drives the evolution of intergroup conflict. Proc Natl Acad Sci 116(28):14089–14097

Hofbauer J (1981) A general cooperation theorem for hypercycles. Monatshefte für Mathematik 91(3):233–240

Hofbauer J, Schreiber SJ (2004) To persist or not to persist? Nonlinearity 17(4):1393

Hofbauer J, Sigmund K (1998) Evolutionary games and population dynamics. Cambridge University Press

Kaznatcheev A, Vander Velde R, Scott JG, Basanta D (2017) Cancer treatment scheduling and dynamic heterogeneity in social dilemmas of tumour acidity and vasculature. Br J Cancer 116(6):785–792

Levin S (2010) Crossing scales, crossing disciplines: collective motion and collective action in the global commons. Philosophical Transactions of the Royal Society B: Biological Sciences 365(1537):13–18

Levin S, Pimentel D (1981) Selection of intermediate rates of increase in parasite-host systems. Am Nat 117(3):308–315

Luo S (2014) A unifying framework reveals key properties of multilevel selection. J Theor Biol 341:41–52

Luo S, Mattingly JC (2017) Scaling limits of a model for selection at two scales. Nonlinearity 30(4):1682

MacLean RC, Fuentes-Hernandez A, Greig D, Hurst LD, Gudelj I (2010) A mixture of "cheats" and "co-operators" can enable maximal group benefit. PLoS Biol 8(9):e1000486

Markvoort AJ, Sinai S, Nowak MA (2014) Computer simulations of cellular group selection reveal mechanism for sustaining cooperation. J Theor Biol 357:123–133

Maynard Smith J (1982) Evolution and the Theory of Games. Cambridge University Press

Maynard Smith J, Száthmary E (1993) The origin of chromosomes I. selection for linkage. J Theor Biol 164(4):437–446

McLoone B, Fan WTL, Pham A, Smead R, Loewe L (2018) Stochasticity, selection, and the evolution of cooperation in a two-level Moran model of the snowdrift game. Complexity 2018

Meizis R (2020) Convergence of metric two-level measure spaces. Stochastic Processes and their Applications 130(6):3499–3539

Muir WM (1996) Group selection for adaptation to multiple-hen cages: selection program and direct responses. Poult Sci 75(4):447–458

Nowak MA (2006) Five rules for the evolution of cooperation. Science 314(5805):1560–1563

Nowak MA, Tarnita CE, Wilson EO (2010) The evolution of eusociality. Nature 466(7310):1057–1062

Ogura Y, Shimakura N (1987) Stationary solutions and their stability for Kimura's diffusion model with intergroup selection. J. Math. of Kyoto University 27(2):305–347

Ogura Y, Shimakura N (1987) Stationary solutions and their stability for Kimura's diffusion model with intergroup selection II. J. Math. of Kyoto University 27(4):635–655

Osorio L, Winter A (2020) Two level branching model for virus population under cell division. arXiv preprint arXiv:2004.14352

Ostrom E (2010) Beyond markets and states: polycentric governance of complex economic systems. American Economic Review 100(3):641–72

Pacheco JM, Santos FC, Souza MO, Skyrms B (2009) Evolutionary dynamics of collective action in n-person stag hunt dilemmas. Proceedings of the Royal Society of London B: Biological Sciences 276(1655):315–321

Paulsson J (2002) Multileveled selection on plasmid replication. Genetics 161(4):1373–1384

Pichugin Y, Traulsen A (2018) Reproduction costs can drive the evolution of groups. bioRxiv p 325670

Pichugin Y, Gokhale CS, Garcia J, Traulsen A, Rainey PB (2015) Modes of migration and multilevel selection in evolutionary multiplayer games. J Theor Biol 387:144–153

Pokalyuk C, Goerzer I (2019) Diversity patterns in parasite populations capable for persistence and reinfection with a view towards the human cytomegalovirus. bioRxiv p 512970

Pokalyuk C, Wakolbinger A (2019) Maintenance of diversity in a hierarchical host–parasite model with balancing selection and reinfection. Stochastic Processes and their Applications

Rainey PB, Rainey K (2003) Evolution of cooperation and conflict in experimental bacterial populations. Nature 425(6953):72–74

Rose CJ, Hammerschmidt K, Pichugin Y, Rainey PB (2020) Meta-population structure and the evolutionary transition to multicellularity. Ecol Lett 23(9):1380–1390

Sachs JL, Bull JJ (2005) Experimental evolution of conflict mediation between genomes. Proc Natl Acad Sci 102(2):390–395

Santos FC, Chalub FA, Pacheco JM (2007) A multi-level selection model for the emergence of social norms. In: European Conference on Artificial Life, Springer, p 525–534

Schlüter M, Tavoni A, Levin S (2016) Robustness of norm-driven cooperation in the commons. Proceedings of the Royal Society B: Biological Sciences 283(1822):20152431

Shaffer Z, Sasaki T, Haney B, Janssen M, Pratt SC, Fewell JH (2016) The foundress's dilemma: group selection for cooperation among queens of the harvester ant, pogonomyrmex californicus. Sci Rep 6:29828

Simon B (2010) A dynamical model of two-level selection. Evol Ecol Res 12(5):555–588

Simon B, Nielsen A (2012) Numerical solutions and animations of group selection dynamics. Evol Ecol Res 14(6):757–768

Simon B, Pilosov M (2016) Group-level events are catalysts in the evolution of cooperation. J Theor Biol 410:125–136

Simon B, Fletcher JA, Doebeli M (2013) Towards a general theory of group selection. Evolution 67(6):1561–1572

Staps M, van Gestel J, Tarnita CE (2019) Emergence of diverse life cycles and life histories at the origin of multicellularity. Nature Ecology & Evolution 3(8):1197–1205

Szathmáry E, Demeter L (1987) Group selection of early replicators and the origin of life. J Theor Biol 128(4):463–486

Szathmáry E, Smith JM (1995) The major evolutionary transitions. Nature 374(6519):227–232

Tarnita CE, Taubes CH, Nowak MA (2013) Evolutionary construction by staying together and coming together. J Theor Biol 320:10–22

Tavoni A, Schlüter M, Levin S (2012) The survival of the conformist: social pressure and renewable resource management. J Theor Biol 299:152–161

Taylor C, Nowak MA (2007) Transforming the dilemma. Evolution: International Journal of Organic Evolution 61(10):2281–2292

Tilman AR, Watson JR, Levin S (2017) Maintaining cooperation in social-ecological systems. Thyroid Res 10(2):155–165

Traulsen A, Nowak MA (2006) Evolution of cooperation by multilevel selection. Proc Natl Acad Sci 103(29):10952–10955

Traulsen A, Sengupta AM, Nowak MA (2005) Stochastic evolutionary dynamics on two levels. J Theor Biol 235(3):393–401

Traulsen A, Shoresh N, Nowak MA (2008) Analytical results for individual and group selection of any intensity. Bull Math Biol 70(5):1410

Turchin P (2010) Warfare and the evolution of social complexity: A multilevel-selection approach. Structure and Dynamics 4(3)

Van Vliet S, Doebeli M (2019) The role of multilevel selection in host microbiome evolution. Proc Natl Acad Sci 116(41):20591–20597

van Veelen M, Luo S, Simon B (2014) A simple model of group selection that cannot be analyzed with inclusive fitness. J Theor Biol 360:279–289

Velleret A (2019) Two level natural selection with a quasi-stationarity approach. arXiv preprint arXiv:1903.10161

Velleret A (2020) Individual-based models under various time-scales. ESAIM: Proceedings and Surveys 68:123–152

Wade MJ, Bijma P, Ellen ED, Muir W (2010) Group selection and social evolution in domesticated animals. Evol Appl 3(5–6):453–465

Xue KS, Hooper KA, Ollodart AR, Dingens AS, Bloom JD (2016) Cooperation between distinct viral variants promotes growth of h3n2 influenza in cell culture. Elife 5:e13974