Condensation in stochastic many-particle models

Applications to non-equilibrium statistical physics and microbial populations dynamics

Johannes S. Knebel

A dissertation submitted

to the Faculty of Physics at the

Ludwig–Maximilians–Universität München

for the degree of

Doctor rerum naturalium

Munich, May 19, 2017

First referee (Erstgutachter): Prof. Dr. Erwin Frey
Second referee (Zweitgutachter): Prof. Dr. Richard A. Blythe
Day of the oral examination: June 30, 2017

Zusammenfassung (Summary in German)

Die gegenseitige Wechselwirkung vieler Teilchen kann kollektives Verhalten hervorrufen, welches von der Betrachtung einzelner Teilchen unerwartet ist. Solch qualitativ neues, emergentes Verhalten umfasst zum Beispiel nichtlineare Rückkopplung, Phasenübergänge und Kondensation. In meiner Promotionsarbeit präsentiere ich unsere Analysen zweier stochastischer Vielteilchensysteme, in denen die Wechselwirkung vieler Teilchen zu Kondensationsphänomenen führt.

I **Kondensation in evolutionären Nullsummenspielen und in offenen, bosonischen Quantensystemen** *mit Markus F. Weber, Torben Krüger und Erwin Frey.*
Das erste Projekt ist motiviert durch Nullsummenspiele in der evolutionäeren Spieltheorie und deren Anwendung auf die Kondensation von nicht-interagierenden Bosonen in periodisch getriebenen und dissipativen Quantensystemen. Wir haben erklärt, wie es in dem stochastischen Vielteilchenprozess, der dem Nullsummenspiel zugrunde liegt, zur Kondensation in mehrere Zustände kommen kann. Diese Kondensation kann als eine Verallgemeinerung der Bose-Einstein Kondensation ins Nichtgleichgewicht verstanden werden. Aus mathematischer Sicht ist der Kondensationsprozess durch die antisymmetrische Lotka-Volterra Gleichung beschrieben, deren Langzeitverhalten wir charakterisiert haben. Die Ergebnisse dieses Projektes haben zu zwei Publikationen geführt, die dieser Doktorarbeit angefügt sind.

II **Kontrolle von Heterogenität in Bakterienpopulationen durch chemische Kommunikation** *mit Matthias Bauer, Matthias Lechner, Peter Pickl und Erwin Frey.*
Bakterien können mit Hilfe chemischer Moleküle, so-genannter Autoinduktoren, kommunizieren und damit kollektives Verhalten bei der Virulenz oder Biolumineszenz steuern. Mit bisherigen Erklärungsmodellen war es konzeptionel schwierig zu erklären, dass Bakterien in einer Population solche Autoinduktoren in unterschiedlichen Mengen produzieren können und trotzdem kollektives Verhalten in der Population gesteuert werden kann. Im zweiten Projekt schlagen wir daher einen Mechanismus vor, der experimentelle Beobachtungen zur heterogenen Produktion von Autoinduktoren in Bakterienpopulationen erklären könnte. Aus theoretischer Sicht kommt es in unserem stochastischen Vielteilchenmodell zu einer Aufteilung der Population in zwei stabil koexistierende Unterpopulationen (zwei Kondensatzustände). Verursacht wird der Kondensationsprozess durch die Veränderung der Umwelt, die durch die Autoinduktoren-Produktion der Bakterien bestimmt ist, und der Anpassung der Bakterien an diese Umwelt. Die mathematischen Aspekte dieses Modells liefern darüber hinaus einen Beitrag zur qualitativen Rolle von langreichweitigen Wechselwirkungen in der statistischen Physik. Die Ergebnisse dieses Projektes befinden sich derzeit im Peer-Review-Prozess.

Overview of the thesis

My thesis deals with two stochastic many-particle models in which the creation and annihilation of particles induce condensation phenomena if the number of particles is large. Even though a broad set of states is accessible to the particles, interactions lead them into only a subset of states, the condensate states. In other words, an initially complex system evolves into fragments over time. This thesis is organized into two parts according to the two projects that I worked on during my doctoral studies. The two theoretical models are motivated by their respective application to (I) nonequilibrium statistical physics of open quantum systems and evolutionary game theory, and to (II) microbial population dynamics. An overview of the two projects is provided in the following.

I Evolutionary games of condensates in stochastic many-particle processes
with Markus F. Weber, Torben Krüger, and Erwin Frey.
In this project, we explained condensation phenomena that arise in evolutionary game theory of zero-sum games and open quantum systems. Our results explain the formation of multiple condensates in driven-dissipative systems of non-interacting bosons, which constitutes a generalization of Bose-Einstein condensation to nonequilibrium. From a mathematical point of view, this condensation is captured by the antisymmetric Lotka-Volterra equation. Our results are published in the two publications "Coexistence and survival in conservative Lotka-Volterra networks", *Phys. Rev. Lett.* **110**(16), 168106 (2013) and "Evolutionary games of condensates in coupled birth-death processes", *Nat. Commun.* **6**, 6977 (2015). I contributed to both publications as first author. The publications are reprinted in sections 4 and 5 of chapter I (also reprinted in the PhD thesis of Markus F. Weber). Sections 1–3 of chapter I provide an overview of our results.

II Ecological feedback in quorum-sensing microbial populations
with Matthias Bauer, Matthias Lechner, Peter Pickl, and Erwin Frey.
In this project, we proposed a theoretical model that accounts for the existence and stability of phenotypic heterogeneity in the production of autoinducers in microbial populations. Autoinducers mediate the density-dependent collective response of microbial populations known as quorum sensing. We showed that the coupling between ecological and population dynamics through quorum sensing facilitates phenotypic heterogeneity. At present, one manuscript incorporating our results is under review for publication, for which Matthias Bauer and I share co-first authorship. Our main results are presented in sections 1–4 of chapter II.

Abstracts of the projects

When individual entities form a collective, the whole may evolve qualitatively differently than expected for the sum of its isolated entities ("more is different" [1]). The interaction between individual entities may lead to novel, and sometimes counter-intuitive, properties at the level of the collective such as nonlinear feedback, phase transitions, and condensation phenomena. How collective behavior emerges in systems with many entities and how it can be mathematically described are central to biological and statistical physics.

During my doctoral studies, I worked on two stochastic many-particle models in which particles condense into a subset of the accessible states of the system. The first of these models (chapter I) was formulated in the framework of zero-sum games in evolutionary game theory. We showed that this model explains the condensation of non-interacting bosons in a periodically driven and dissipative set-up, which was proposed theoretically only recently. The second stochastic many-particle model (chapter II) was developed in the context of microbial population dynamics and can also be formulated in the framework of evolutionary game theory. For both models, the model-specific interactions between the particles guide the particles into a certain subset of states during a realization of the stochastic process provided that the number of particles is large. The congregation of particles into certain states can be interpreted as a condensation process. We explored the conditions under which thermodynamic equilibrium is approached in the respective stochastic processes, and under which conditions the processes remain out of equilibrium and approach so-called nonequilibrium steady states. For each of the two models we determined the condensate states and answered how they are selected from the set of accessible states of the system. In the first model, the number of accessible states is finite, whereas it is a continuum in the second model. We explained the condensation processes in both models by a rigorous mathematical analysis and identified the specific interactions as the driver of condensation. For the first model, an antisymmetric rule in the pairwise interactions determines condensation: the amount of gain for one state equals the loss of another state. For the second model, effective global interactions between the particles are responsible for condensation: every particle communicates with every other particle in the population.

Beyond their applications to physics and microbiology that are described below, our two models show that simple microscopic interactions between particles can yield emergent — and to some degree surprising — condensation behavior on a macroscopic level of description.

I Evolutionary games of condensates in stochastic many-particle processes

with Markus F. Weber, Torben Krüger, and Erwin Frey.

Summary

The first project of my thesis deals with a stochastic many-particle model with which we explained a generic class of condensation phenomena that arise in both classical and quantum systems. Interestingly, this condensation does not proceed into a single, but into multiple states in a single realization of the stochastic process. Only recently has it been shown that this stochastic many-particle process governs the coarse-grained dynamics of a periodically driven and dissipative system of non-interacting bosons. This process can be understood as a generalization of Bose-Einstein condensation to nonequilibrium. We showed that the condensation of bosons in such a driven-dissipative, nonequilibrium system corresponds to the selection of strategies in evolutionary game theory. By combining analytical concepts from the theory of stochastic processes, nonlinear dynamics, and linear programming theory applied to antisymmetric matrices, we explained how multiple condensates form in the stochastic process. Furthermore, we proposed the possibility of condensates with oscillating occupations. We found that the condensation dynamics follow a simple physical guiding principle: the vanishing of relative entropy production guides the selection of condensates. How this condensation of bosons in nonequilibrium and the design of oscillating condensates may be realized with an open quantum system is an interesting challenge for future experiments.

Background information

A quantum particle is either a fermion or a boson. Whereas two fermions cannot occupy the same quantum state (Pauli exclusion principle), many bosons may cluster in the same quantum state. One dramatic consequence of these different quantum statistics was predicted on theoretical grounds by Bose and Einstein in 1924 upon studying a system of non-interacting bosons in thermodynamic equilibrium [2–4]. They showed that a macroscopic fraction of the bosons can cluster into the ground state of the system if the temperature of the heat bath is tuned below a critical value. This phase of matter, commonly referred to as a Bose-Einstein condensate, was experimentally realized with ultra-cold atoms for the first time in 1995 [5, 6].

How the concept of Bose-Einstein condensation extends to nonequilibrium systems has remained an intriguing question. Only recently has it been proposed theoretically that a system of non-interacting bosons driven by a time-periodic potential and weakly coupled to a heat bath may provide a suitable set-up for such a study [7]. The dynamics of this quantum system is effectively incoherent, that is, it can be described in terms of a stochastic many-body process on a coarse-grained time scale. Interestingly, it was observed in numerical studies that condensation in such a nonequilibrium set-up may not only proceed into a single condensate state, but into *multiple* condensates [7].

Research question

In this project we wanted to understand how multiple condensates form in the stochastic many-particle process that governs the dynamics of non-interacting bosons in driven-dissipative systems. How does this condensation proceed and what is the physical principle guiding the selection of condensates?

Summary of our findings

Correspondence between evolutionary game theory and bosonic condensation out of equilibrium. As a first result, we established the correspondence between the dynamics of the bosonic condensation and the selection of strategies in evolutionary zero-sum games. Strategies that prevail in an evolutionary zero-sum game correspond to the quantum Floquet states that become the condensates in a periodically driven and dissipative system of non-interacting bosons. Both phenomena are described by the same stochastic birth-death process, which is also known as the inclusion process. Furthermore, we characterized the condensation dynamics in this stochastic process. We explained how multiple condensates form, and which of the states become the condensates. Our results were published in "Evolutionary games of condensates in coupled birth-death processes", *Nat. Commun.* **6**, 6977 (2015), and are reprinted in section 5 of chapter I.

Condensation is described by the antisymmetric Lotka-Volterra equation (ALVE). We found that the dynamics of the stochastic many-particle process evolves on two separate time scales if the number of particles is large. The dynamics on the first time scale, which grows proportionally with the particle number, determines the outcome of the condensation. On this time scale, the temporal evolution of the stochastic process is governed by a set of nonlinearly coupled, ordinary differential equations that we refer to as the antisymmetric Lotka-Volterra equation (ALVE). The ALVE is characterized by antisymmetric, pairwise interactions between the states. The rates with which mass is exchanged between the states are determined by an antisymmetric matrix. On the second time scale, the properties of the eventual stationary state of the stochastic process are determined. Depending on the antisymmetric matrix, absorbing states, thermodynamic equilibrium, or nonequilibrium steady states are approached at long times.

Condensation in the ALVE. We found that condensation in the dynamics of the ALVE is fully determined by the defining antisymmetric matrix. We formulated an algebraic characterization of antisymmetric matrices to determine whether the ALVE allows the stable coexistence of all states, or whether and how it breaks up into a reduced subset of states. This subset of states is the set of condensates that is approached independently of the choice of initial conditions. In other words, the manifold spanned by the set of condensates is a global attractor of the dynamics defined by the ALVE. We proved that the condensation dynamics in the ALVE generically proceed exponentially fast, meaning that the manifold of condensates is approached exponen-

tially fast. This fragmentation of the state space is caused by the vanishing of relative entropy production. The vanishing of this collective quantity guides the selection of condensates as the underlying physical principle. Furthermore, we demonstrated how one can design systems that evolve into specific sets of condensates. For example, we formulated conditions on the antisymmetric matrix such that the resulting set of condensates exhibits a rock-paper-scissors network topology. In this set-up, the dynamics never come to rest, but perpetuate in an oscillating manner.
Our algebraic method of determining the condensates can be formulated as a problem in linear programming theory for antisymmetric matrices, and identifies so-called condensate vectors of antisymmetric matrices. These condensate vectors can be thought of as generalized kernel vectors that are specific to antisymmetric matrices; they determine which of the states become the condensates. By numerically exploiting our analytical insight into the condensation dynamics, we were able to study the ALVE for ensembles of large random networks of states. We found that the average number of condensates reaches a maximum at a critical value of the connectivity of the random network. This critical connectivity decreases as a power law when the number of states in the network is increased. The criticality of random networks thereby translates into a critical behavior of the condensation dynamics of the ALVE.

Coexistence of all states in the ALVE. In a previous study, we also characterized the conditions under which all states coexist as condensates in the ALVE, that is, none of the states becomes depleted. These conditions can be formulated in terms of Pfaffians of submatrices of the antisymmetric matrix occurring in the ALVE. The Pfaffian of an antisymmetric matrix is a determinant-like function tailored to antisymmetric matrices. We also showed how linearly independent kernel elements of the antisymmetric matrix give rise to independent conserved quantities. These conserved quantities restrict the trajectory defined by the ALVE onto a manifold that is diffeomorphic to a sphere of odd dimension. Using these conserved quantities, we demonstrated how one can construct periodic orbits in the ALVE for networks with an arbitrary number of states. Finally, we discussed the implications of our findings for the stochastic process of zero-sum games in evolutionary game theory without mutations in finite populations. We investigated how bifurcations in the ALVE affect the stochastic process. Bifurcations in the ALVE can be evoked by tuning entries of the antisymmetric matrix. For finite populations, the distance to a bifurcation affects the time until the first state has become depleted. We found that this depletion time diverges in the vicinity of critical parameter values for which coexistence of all states is obtained in the ALVE. Away from a bifurcation, the depletion time increases logarithmically with the number of particles as expected from a deterministic analysis, whereas it increases linearly with the number of particles at the bifurcation. We discussed these effects by using the rock-paper-scissors-lizard-spock network topology as an example. The analysis of coexistence scenarios in the ALVE and their consequences for the divergence of timescales in the vicinity of bifurcations were

published in "Coexistence and survival in conservative Lotka-Volterra networks", *Phys. Rev. Lett.* **110**(16), 168106 (2013), and are reprinted in section 4 of chapter I.

Relevance of our findings and prospects for future research
In the context of non-interacting bosons in periodically driven and dissipative set-ups, our results may be of experimental relevance. Periodic energy input and dissipation in open quantum systems may balance such that bosons aggregate into multiple quantum states. Given the experimental parameters of such a set-up, our results provide a theoretical method to determine which of the quantum states become the condensates. Moreover, we predicted the possibility of condensates with an oscillating occupation of bosons. We showed how one can tune the transition rates of bosons between states such that a rock-paper-scissors game of condensates emerges. How such oscillating condensates may be realized in an experimentally controlled open quantum system remains, at present, a challenge for future research.

II Ecological feedback in quorum-sensing microbial populations

with Matthias Bauer, Matthias Lechner, Peter Pickl, and Erwin Frey.

Summary

The second project of my thesis deals with a stochastic many-particle model whose applications range from microbial population dynamics to statistical physics and mathematical physics. We showed that the coupling between ecological and population dynamics through quorum sensing may induce phenotypic heterogeneity in the production of autoinducers in microbial populations. Our theoretical model and our mathematical analysis qualitatively explain recent experimental observations in microbial population dynamics and challenge currently accepted views on the origin of phenotypic heterogeneity in quorum-sensing microbial populations. Future experiments are needed to verify or falsify our model assumptions and predictions, and to clarify the role of the identified ecological feedback for phenotypic heterogeneity in quorum-sensing microbial populations.

Background information

Microbes can communicate with each other by making use of a chemical language [8, 9]. They produce so-called autoinducers, which are small signaling molecules, and secrete those molecules into the environment. These molecules are sensed by other microbes in the population, in turn. Upon responding to the sensed level of autoinducers in the environment, a coordinated gene expression of all cells of the population can be triggered. Such collective behavior of microbes is commonly referred to as "quorum sensing" and comprises, for example, the coordinated and collective expression of genes for virulence, biofilm formation, and bioluminescence. Elucidating the mechanisms that control the production of autoinducers and the level of autoinducers in the environment is, thus, pertinent to understanding such collective microbial behavior.

Recent experiments suggest that the production of autoinducers may vary between genetically identical cells in a population in that some cells of the population expressed autoinducer synthase genes during microbial growth, while others did not [10, 11]. This behavior is referred to as phenotypic heterogeneity [12]. The stable coexistence of different phenotypes in one population may serve the division of labor or act as a bet-hedging strategy and, thus, may be beneficial for the survival and resilience of a microbial species at long time scales. However, the experimentally observed phenotypic heterogeneity in the autoinducer production is not expected to occur in well-mixed populations if currently favored threshold models for quorum-sensing response are adopted.

Research question

In this work, we asked how phenotypic heterogeneity in the production of autoinducers is established in quorum-sensing microbial populations. How can a microbial

population control phenotypic heterogeneity of autoinducer production and, concomitantly, tightly adjust the average production level in the population to trigger quorum-sensing functions such as virulence?

Summary of our findings

The quorum-sensing model. We studied the collective behavior of a stochastic many-particle model of quorum sensing, in which cells produce autoinducers to different degrees and secrete those into a well-mixed environment. Autoinducers become thereby shared amongst all individuals. On the one hand, production of autoinducer molecules and accompanied gene expression are assumed to be metabolically costly such that non-producers reproduce faster than producing cells. On the other hand, we assume that cells can up-regulate their autoinducer production by a sense-and-response mechanism through quorum sensing. That is, individuals can increase their production in response to the sensed average production level in the population. The central feature of the quorum-sensing model is that individuals shape their environment (through the production of autoinducers) and respond to this self-shaped environment (by changing the individual production of autoinducers), in turn. Thus, ecological and population dynamics are coupled in the quorum-sensing model.

Phenotypic heterogeneity in the autoinducer production through the coupling of ecological and population dynamics in quorum-sensing microbial populations. We found that the coupling between ecological and population dynamics through quorum sensing can control a heterogeneous production of autoinducers in microbial populations. The population may split into two subpopulations: one with a low, and a second with a high production of autoinducers. This phenotypic heterogeneity in the autoinducer production is stable for many generations. At the same time, the overall autoinducer level in the environment is robustly self-regulated by how cellular production is up-regulated. Thus, further quorum-sensing functions such as virulence or bioluminescence can be triggered in the population. If cellular response to the environment is absent or too frequent, phase transitions occur from heterogeneous to homogeneous populations in which all individuals produce autoinducers to the same degree. Our results show that, if microbes sense and respond to their self-shaped environment, the population may not only respond as a homogeneous collective as is typically associated with quorum sensing, but may also become a robustly controlled collective of two different subpopulations.

Mean-field analysis of the quorum-sensing model explains phenotypic heterogeneity at the population level. To mathematically capture the numerically observed phenotypic heterogeneity in the autoinducer production, we developed a kinetic theory of the quorum-sensing model in three steps.

(i) First, we heuristically derived the mean-field equation of the quorum-sensing model that captures the microbial population dynamics on a macroscopic level. We refer to this mean-field equation as the autoinduction equation, which describes how

the distribution of production degrees in the population evolves in time.
(ii) Second, we proved that and quantified how the stochastic process converges to this mean-field equation as the population size grows to infinity. To this end, we introduced an auxiliary stochastic mean-field process that mimics the temporal evolution of the mean-field equation and updates the individuals' production degrees in an independent manner. This way, the law of large numbers is separated from the propagation of errors that build up due to correlations between the production degrees of the individuals. The proof for the autoinduction equation gives concrete error bounds for the quality and the speed of convergence of the stochastic process to the mean-field equation. These error bounds depend on the population size and initial correlations. Furthermore, our proof shows that the convergence to the mean-field equation is robust against changes of microscopic details in the definition of the stochastic process.
(iii) Third, we analyzed the mean-field equation and explained both homogeneous and heterogeneous states of the population. Depending on how growth rate differences between producers and non-producers balance with the response rate to the environment, homogeneous (unimodal) or heterogeneous (bimodal) stationary distributions are approached at long times. These stationary distributions can be interpreted as condensation-like solutions of the mean-field equation. The existence and the stability of heterogeneous stationary distributions is a consequence of the feedback between ecological and population dynamics, which effectively introduces long-range interactions between the individuals of the population. Furthermore, we explained the transitions from heterogeneous to homogeneous distributions in terms of nonequilibrium phase transitions.
In total, our mathematical analysis shows that phenotypic heterogeneity arises dynamically in the quorum-sensing model and that it is robust both against changes in the definition of the stochastic many-particle process (how up-regulation and growth rate differences are implemented) and against perturbations and demographic noise of the stochastic dynamics.

Relevance of our findings and prospects for future research
We expect that our results will have applications in the field of microbiology at the interface between phenotypic heterogeneity and quorum sensing in microbial populations, and in the field of theoretical physics with focus on nonequilibrium statistical mechanics and biological physics.

Relevance for microbiology. In recent years, a deeper mechanistic understanding of phenotypic heterogeneity has been achieved by exploring how the presence of different phenotypes in a population of genetically identical cells depends upon the stochastic expression of bistable gene regulatory circuits at the cellular level (so called "threshold models") [12–14]. Nonetheless, bistable regulation of autoinducer production is not always verified experimentally and, thus, this standard view of

phenotypic heterogeneity can be questioned. Our findings suggest an alternative, robust mechanism to stochastic gene expression in bistable gene regulatory circuits that may be realized through the coupling of ecological and population dynamics in quorum-sensing microbial populations.

We expect that our theoretical work will stimulate research on quorum sensing in microbial populations and alternative mechanisms to threshold models that generate phenotypic heterogeneity. Thus far, quorum sensing has only been shown to induce homogeneous or synchronized populations. In our work, we showed that quorum sensing might also play a decisive role for how phenotypic heterogeneity is controlled at the population level as opposed to control at the cellular level (as assumed in threshold models). We demonstrated that simple intuitions about quorum sensing might be misguiding if one neglects the interaction of microbes with their self-shaped environment.

Our conceptual model of quorum-sensing microbial populations describes a robust mechanism to induce phenotypic heterogeneity in the expression of autoinducers through the coupling of ecological and population dynamics. It will be highly interesting to conduct experiments that verify or falsify this proposed mechanism, that discriminate between the ecological feedback mechanism and the noisy bistable regulation mechanism in threshold models, and that reveal the possible interplay between the two mechanisms in quorum-sensing microbial populations.

Relevance for statistical physics. From a statistical physics point of view, the key feature of our quorum-sensing model is how a population evolves in time when its constituents respond to an environment that is being shaped by their own activities. The feedback between ecological and population dynamics is mediated by quorum sensing and creates an effective global coupling between the individuals in the population. Such a global coupling is reminiscent of long-range interactions in models of statistical mechanics, such as in the classical XY spin model with infinite range interactions [15]. Our analysis suggests that such global feedback can result in bimodal nonequilibrium steady-states in cases for which one might have naively expected unimodal equilibrium states at long times. We believe that this insight could inspire others to study the role of global feedback loops and long-range interactions that arise, for example, through the coupling of population and ecological dynamics. From a mathematical point of view, we believe that our developed method of an auxiliary stochastic mean-field process could be helpful to prove the convergence of mean-field equations for other stochastic many-particle processes. Our derivation and analysis of the mean-field equation and the mathematical characterization of phenotypic heterogeneity provide a general framework to study stochastic many-particle processes that arise through ecological feedback.

Contents

I Evolutionary games of condensates in stochastic many-particle processes

Abstract of the chapter

This chapter presents our results on a stochastic many-particle process that arises for example, in the study of zero-sum games in evolutionary game theory (where it describes the selection of strategies), and in the study of driven-dissipative systems of non-interacting bosons. In the latter context, the process describes the formation of multiple condensates as a generalization to Bose-Einstein condensation in thermodynamic equilibrium. A graphical summary of our work is provided in figure 1. The stochastic process is generically out of equilibrium and may exhibit condensation into multiple states if the number of particles is large. Thus far, it has remained elusive from an analytical point of view, which of the system's states become the condensates given the parameters and initial conditions of the stochastic process.

In our work, we showed that a nonlinearly coupled system of ordinary differential equations, the so-called *antisymmetric Lotka-Volterra equation* (17), describes the condensation into multiple quantum states. Moreover, we developed an algebraic method to determine the condensates. A suitably defined relative entropy guides the selection of condensates. The condensation dynamics generically proceed exponentially fast. However, the dynamics within the selected set of condensates may not come to rest, but perpetuate in an oscillating manner instead. To demonstrate this phenomenon, we proposed the design of a rock–paper–scissors game of condensates.

The results of our work are published in two publications, which are reprinted in sections 4 and 5. Acknowledgments related to this project can be found in the manuscripts. Sections 1 and 2 provide a fast-forward introduction and an overview of our results; section 3 serves as an outlook to study further properties of the stochastic process.

1 Introduction of the stochastic many-particle process with coupled birth and death

1.1 A stochastic process on a network of states

In the following, we define the stochastic process whose dynamics is analyzed in this chapter. We consider a system of S non-degenerate states $E_i, i = 1, \ldots, S < \infty$, each of which is occupied by $N_i \geq 0$ indistinguishable particles; figure 1(A). The physical

interpretation of the states E_i is described in detail in sections 1.3 and 1.4. The configuration of the system at time t is fully characterized by the vector of occupation numbers $\mathbf{N} = (N_1, N_2, \ldots, N_S) \in \mathbb{N}_0^S$. For later purposes, we also introduce the state concentrations $\mathbf{x} = (x_1, \ldots, x_S)$ with $x_i := N_i/N$. Configurations may change due to jumps of particles between two states E_i and E_j according to the following scheme:

$$\begin{array}{c} (N_1, \ldots, N_i - 1, \ldots, N_j + 1, \ldots, N_S) = \mathbf{N} - \mathbf{e}_i + \mathbf{e}_j \\ \updownarrow \\ (N_1, \ldots, N_i, \ldots, N_j, \ldots, N_S) = \mathbf{N} \\ \updownarrow \\ (N_1, \ldots, N_i + 1, \ldots, N_j - 1, \ldots, N_S) = \mathbf{N} + \mathbf{e}_i - \mathbf{e}_j \,. \end{array}$$

Here, the vector $\mathbf{e}_i \in \mathbb{Z}^S$ denotes the unit vector in direction i (equal to one at index i, otherwise zero). Jumps of particles between the states and, thus, transitions between different configurations, occur continuously in time and are modelled in terms of a continuous-time Markov process. The jump process conserves the total number of particles $N = \sum_i N_i$, or equivalently $\sum_i x_i = 1$, over time and, thus, can be interpreted as a coupled birth-death process. We are interested in the probability $P(\mathbf{N}, t)$ of finding the system in configuration $\mathbf{N}$ at time t, given that it was initially in configuration $\mathbf{N}_0$ at time t_0. The temporal evolution of the probability distribution $P(\mathbf{N}, t)$ is governed by the classical master equation [16–18]:

$$\partial_t P(\mathbf{N}, t) = \sum_{\substack{i,j=1 \\ j \neq i}}^{S} \Big(\Gamma_{i \leftarrow j}(N_i - 1, N_j + 1) P(\mathbf{N} - \mathbf{e}_i + \mathbf{e}_j, t) - \Gamma_{i \leftarrow j}(N_i, N_j) P(\mathbf{N}, t) \Big) \,. \tag{1}$$

We consider the following transition rate from configuration $\mathbf{N}$ to $\mathbf{N} + \mathbf{e}_i - \mathbf{e}_j$:

$$\Gamma_{i \leftarrow j}(N_i, N_j) = r_{ij}(N_i + s_{ij}) N_j, \text{ with } s_{ij} \geq 0 \text{ and } r_{ij} \geq 0 \,. \tag{2}$$

A stochastic process with the above choice of the transition rates is sometimes also referred to as the inclusion process [19]. It governs the dynamics of both the selection of strategies in zero-sum games in evolutionary game theory (see section 1.3) and the condensation of non-interacting bosons into quantum states in driven-dissipative systems (see section 1.4). While the values of the parameters $\{s_{ij}\}_{i,j}$ can attain any non-negative value in the context of evolutionary game theory, they are equal to 1 for all i and j in the context of bosonic condensation. The parameters $\{r_{ij}\}_{i,j} =: R$ turn out to be the crucial bifurcation parameters of the dynamics as we describe in section 2; see also figure 1.

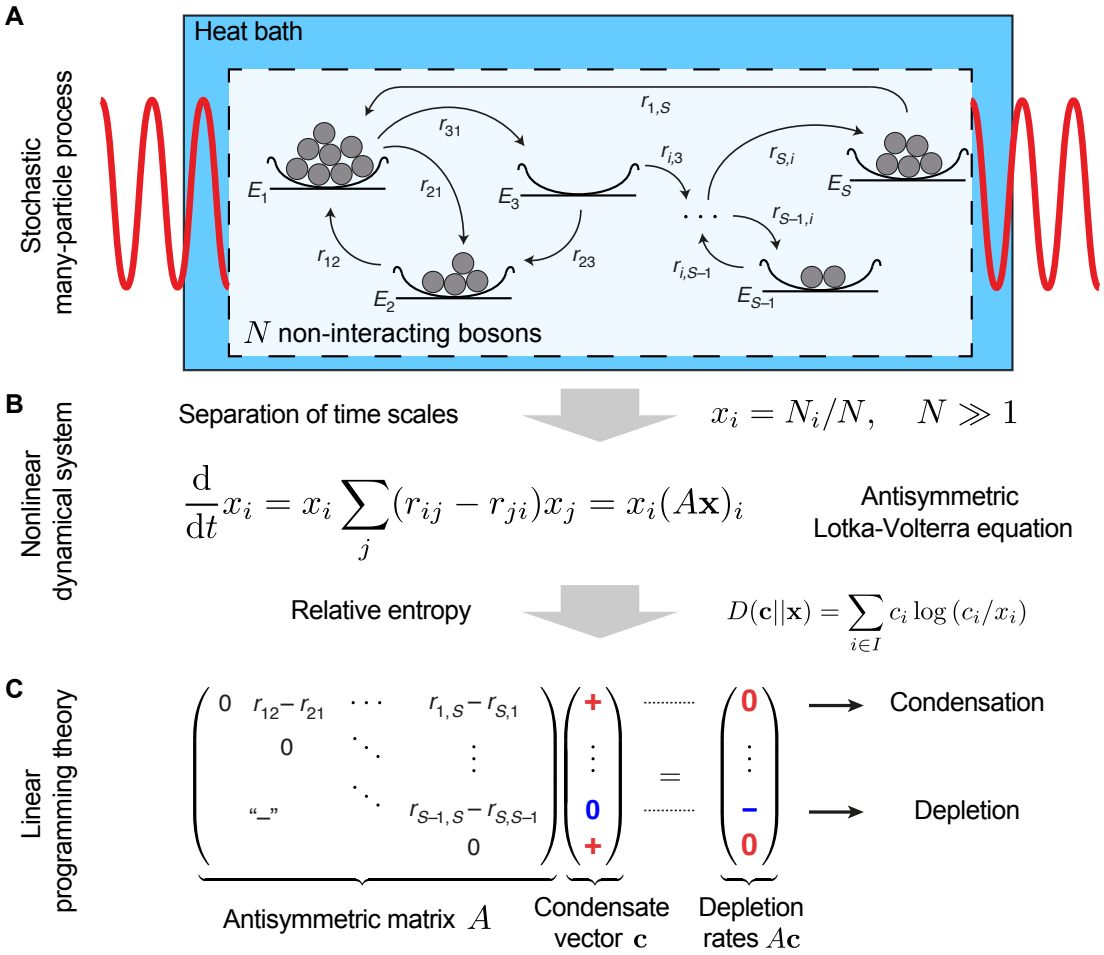


Figure 1. Graphical summary of our work on evolutionary games of condensates in stochastic many-particle processes. (A) In our work, we consider the effective dynamics of a periodically driven-dissipative system of N non-interacting bosons; see section 1.4. Such an open quantum system becomes incoherent over time and can be described in terms of a classical many-particle process (1). The total rate for the transition of a boson from state E_j to E_i does not only depend on the number of bosons in the departure state (N_j), but also on the arrival state ($N_i + 1$) reflecting the quantum statistics of bosons (16). Moreover, the rate constants $\{r_{ij}\}_{i,j}$ characterize the transition between two states. Remarkably, condensation can proceed into *multiple* states depending upon the values of the rate constants. (B) We are interested in how a single realization of the stochastic process proceeds if $N \gg 1$. We showed that the condensation process is described by the *antisymmetric Lotka-Volterra equation* (17); see section 1.2. The entries of the occurring antisymmetric matrix A are determined from the rate constants as $a_{ij} = r_{ij} - r_{ji}$. (C) We developed an algebraic method to determine the condensates; see section 2. By determining a so-called condensate vector $\mathbf{c}$ of the antisymmetric matrix A, and defining the relative entropy of $\mathbf{c}$ to the vector of concentrations $\mathbf{x}$, we identified the states that become the condensates and characterized the condensation process. The dynamics within the selected set of condensates may not come to rest, but perpetuate in an oscillating manner. In particular, we proposed a rock–paper–scissors game of condensates [20].

1.2 Separation of time scales in large populations

In our work, we are interested in how condensation proceeds in a single realization of the stochastic process defined by the master equation (1) with rates (2). Note that, in general, this question is conceptionally different from asking how an ensemble evolves in time.

To address this question on a mathematical level, one may approximate the master equation (1) by performing a Kramers-Moyal expansion if the particle number $N \gg 1$ is large. Truncating this expansion at second order in $1/N$ leads to the Fokker-Planck equation, which can be written as a set of Itō stochastic differential equations for the state concentrations x_i:

$$\mathrm{d}x_i = \left(\alpha_{i,0}(\mathbf{x}) + \frac{1}{N}\alpha_{i,1}(\mathbf{x})\right)\mathrm{d}t + \frac{1}{\sqrt{N}}\sum_{j=1}^{S}\zeta_{ij}(\mathbf{x})\,\mathrm{d}W_j\,. \tag{3}$$

Here, the deterministic drift terms are obtained as:

$$\alpha_{i,0}(\mathbf{x}) = x_i\sum_{j=1}^{S}(r_{ij} - r_{ji})x_j\,, \tag{4}$$

$$\alpha_{i,1}(\mathbf{x}) = \frac{1}{N}\sum_{j=1}^{S}(r_{ij}s_{ij}x_j - r_{ji}s_{ji}x_i)\,, \tag{5}$$

and the stochastic term is represented by the Wiener increment $\mathrm{d}W_j$ of zero mean and with covariances $\mathbb{E}(\mathrm{d}W_i\,\mathrm{d}W_j) = \delta_{ij}\mathrm{d}t$ (accordingly, the unconditional probability density function of the Wiener process W_j is a normal distribution with zero mean and variance t at time t). The matrix $\zeta(\mathbf{x})$ is a square root of the positive semi-definite diffusion matrix $\beta_{ij} = \beta_{ij,0} + \frac{1}{N}\beta_{ij,1}$ in the sense that $\zeta\zeta^T = \beta$. The decomposition of β implies that ζ can be written as $\zeta_{ij} = \zeta_{ij,0} + \mathcal{O}(\frac{1}{N})$. A derivation of these findings and the explicit form of β are reprinted in section 5 on pages 75–77. For the above set of stochastic differential equations (3), a leading-order time scale and subleading-order time scales can be identified. On the leading-order time scale ($t \sim \mathcal{O}(1)$), only the drift term $\alpha_{i,0}$ is relevant, whereas on the first subleading-order time scale ($t \sim \mathcal{O}(N)$), the terms $\alpha_{i,1}$ and $\zeta_{ij,0}$ compete. Here "subleading" means that the latter two terms (and also the term $\zeta_{ij,1}$) cause only slow changes of the leading-order dynamics. In other words, the dynamics on the subleading-order time scales causes only slow changes of the $\mathcal{O}(1)$-trajectory.

The leading-order dynamics of the stochastic process (1) for $N \gg 1$ is, thus, governed by the following set of ordinary differential equations:

$$\frac{\mathrm{d}}{\mathrm{d}t} x_i^{(\mathrm{lead})} = x_i^{(\mathrm{lead})} \sum_j (r_{ij} - r_{ji}) x_j^{(\mathrm{lead})} = x_i^{(\mathrm{lead})} (A\mathbf{x}^{(\mathrm{lead})})_i \,, \tag{6}$$

which we refer to as the *antisymmetric Lotka-Volterra equation* (ALVE). The ALVE is characterized by the antisymmetric matrix A with entries $a_{ij} \coloneqq r_{ij} - r_{ji}$; see equation (17) for a complete mathematical definition.

As we show in section 2, the ALVE describes the condensation into a subset of the system's states. Because condensation in the ALVE proceeds exponentially fast for generic choices of antisymmetric matrices A, the condensation dynamics of the ALVE determine the condensation dynamics of the stochastic process (1), and the temporal evolution up to a time $t \sim \mathcal{O}(N)$ [20]. It is noteworthy that only the parameters $\{r_{ij}\}_{i,j}$ characterize the dynamics at the leading-order time scale; the parameters $\{s_{ij}\}_{i,j}$ influence the dynamics only on subleading-order time scales.

To see that both the deterministic drift term $\alpha_{i,1}$ and the stochastic term $\zeta_{ij,0}$ contribute only on the time scale $t \sim \mathcal{O}(N)$, one may introduce the dilated time $t' = t/N$ and the Wiener process $V_{t'} = W_t/\sqrt{N}$ (with zero mean and variance t'; thus, $\mathbb{E}(\mathrm{d}V_i\, \mathrm{d}V_j) = \delta_{ij}\mathrm{d}t'$). On this dilated time scale t', the corrections due to the terms $\alpha_{i,1}$ and $\zeta_{ij,0}$ read as follows:

$$\mathrm{d}x_i^{(\mathrm{sublead})} = \alpha_{i,1}(\mathbf{x})\, \mathrm{d}t' + \sum_{j=1}^{S} \zeta_{ij,0}(\mathbf{x})\, \mathrm{d}V_j \,. \tag{7}$$

The influence of the subleading-order time scales on the characteristics of the stochastic process are briefly discussed as an outlook in section 3.

1.3 Selection of strategies in zero-sum games in evolutionary game theory

We began our study of the master equation (1) in the context of strategy selection in zero-sum games in evolutionary game theory, for which some background is provided in the following (see [21–23] and references therein for detailed and pedagogical introductions to evolutionary game theory). First, we introduce the nonlinear dynamical framework within which interaction terms are derived from game theoretical concepts in a second step. In essence, the antisymmetric Lotka-Volterra equation is derived as the replicator equation of zero-sum games.

Evolutionary game theory has been developed as a mathematical concept to study mechanisms that underlie the temporal evolution of biological populations [24, 25].

Not only has evolutionary game theory contributed to the understanding of the evolution of cooperation [26], but it has also initiated theoretical discussions and experiments on mechanisms that may stabilize coexistence of species and biodiversity [27–31].

At the heart of evolutionary game theory stands the idea that interactions between different phenotypes, traits, or behavioral programs of species can be understood as interacting strategies in a population of individuals. To connect the notion of evolutionary game theory to the previous section, let us denote such a set of strategies in a population as $\{E_1, E_2, \ldots, E_S\}$. Whether a strategy spreads or goes extinct in the long run depends on the abundance of all other strategies in the population (hence, manifesting the principle of frequency-dependent selection) and on whether this strategy entails a relative advantage over the other strategies at a specific time.

1.3.1 Deterministic replicator dynamics

On a mathematical level, evolutionary game theory was formalized first in the framework of the celebrated replicator equation [21, 23, 32, 33]:

$$\partial_t x_i = x_i(f_i(\mathbf{x}) - \bar{f}(\mathbf{x})) , \quad i = 1, \ldots, S . \tag{8}$$

The replicator equation describes the temporal evolution of the composition of a population in terms of S nonlinearly coupled differential equations assuming a well-mixed, infinitely large population. The variable x_i denotes the fraction of individuals of strategy E_i in the population. For the sake of readability and simplicity, we use here the same symbol x_i to denote a deterministic variable (while it denotes a random variable in the previous section). The relative growth rate ($\partial_t x_i / x_i = \partial_t \ln(x_i)$) of the fraction of individuals with strategy E_i is given by the difference of the strategy's fitness f_i and the average fitness of the population $\bar{f} = \sum_{i=1}^{S} x_i f_i$. The initial conditions are chosen such that $x_{i,0} > 0$ for all $i = 1, \ldots, S$ and $\sum_{i=1}^{S} x_{i,0} = 1$. The latter two properties are conserved over time by the dynamics of the replicator equation (8) [21]. For this reason, the normalization $\bar{f}$ appears in the replicator equation in the first place.

Possible long-time scenarios and stability properties of the replicator equation (8) such as coexistence of all strategies, survival, and extinction are determined by the choice of the fitness function $f_i(\mathbf{x})$. The fitness of a strategy of type E_i may be prescribed by several functional relations, depending upon the biological situation that one might want to model. In evolutionary game theory, as already indicated above, the fitness function is derived from models in classical game theory to capture the interactions between phenotypes, traits or behavioral programs in terms of the interactions between strategies. These interactions are characterized, for example, by a symmetric two-player game [34] and its according payoff matrix $\mathcal{P}$; see below for

an example. It is assumed that the fitness of strategy E_i is derived from the weighted payoff $(\mathcal{P}\mathbf{x})_i$ it receives when playing against all other strategies [35]. Common choices of the fitness function are given by:

(1) $f_i = (\mathcal{P}\mathbf{x})_i$, assuming that fitness equals the weighted payoff;

(2) $f_i = 1 + s(\mathcal{P}\mathbf{x})_i$, introducing the basal fitness of 1 and the selection strength s, which scales the influence of the underlying game on the replicator dynamics and effectively sets the time scale of the selection process described by the replicator equation (8); see chapter II for an example (the fitness function is denoted as ϕ_i that example);

(3) $f_i = e^{s(\mathcal{P}\mathbf{x})_i}$, defined in order to ensure positive fitness values f_i. For small selection constants, one obtains case (2) as an approximation.

Instead of defining classical game theory from first principles, let us here use the example of the prominent children's game Rock-Paper-Scissors to illustrate the concept of a payoff matrix. In an interaction of this game, two players independently choose amongst the three strategies E_1 = Rock, E_2 = Paper, and E_3 = Scissors. The payoff in a single round of the game depends upon the strategies played and is distributed as follows: Paper wraps Rock and, thus, is assigned the payoff 1 (while Rock looses that same amount against Paper), Scissors cuts Paper (thus, Paper pays the payoff 1 to Scissors), and Rock crushes Scissors (payoff 1 for Rock and payoff -1 for Scissors). When the same strategies meet in a round of the game, the payoff is 0 for both players. In summary, one defines the Rock-Paper-Scissors game through the following payoff diagram:

	Rock	Paper	Scissors
Rock	0	-1	1
Paper	1	0	-1
Scissors	-1	1	0

Accordingly, the payoff matrix is given as:

$$\mathcal{P}_{\mathrm{RPS}} = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}, \tag{9}$$

which can be illustrated by a weighted network because of the antisymmetry of the matrix $\mathcal{P}_{\mathrm{RPS}}$, see figure 2(A).

Because the payoff of the winner equals the loss of the defeated, the total payoff in one round of the game is zero. Therefore, the Rock-Paper-Scissors game is also referred to as a zero-sum game [34]. One may, of course, generalize the Rock-Paper-Scissors game to a zero-sum game with four, five, or more interacting strategies

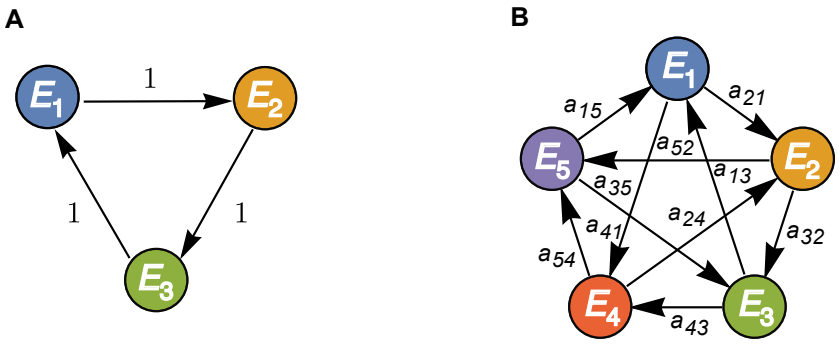


Figure 2. Network representation of zero-sum games. The antisymmetric payoff matrix of a zero-sum game can be depicted as a weighted graph. The nodes of the network correspond to the strategies of the zero-sum game. A link connecting $E_i \to E_j$ denotes that strategy E_j dominates E_i because the payoff in an interaction is transferred from E_i to E_j. The transferred amount of payoff is denoted by the weight attached to a link. (A) The children's game Rock-Paper-Scissors with payoff 1 for every dominance relation. (B) A zero-sum game with five strategies, which may be interpreted as a Rock-Paper-Scissors-Lizard-Spock game, and arbitrary payoff values [36]. Every strategy dominates two other strategies and is dominated by two others.

and allow for arbitrary payoff values; see 2(B). As long as the payoff matrix is an antisymmetric matrix A (that is, $A = -A^T$), the game is a zero-sum game.

In the framework of evolutionary game theory, the choice of a zero-sum game for the definition of the fitness $f_i(\mathbf{x}) = (A\mathbf{x})_i$ simplifies the replicator equation (8). For an antisymmetric matrix, the average fitness of the population is zero: $\bar{f} = \sum_{i=1}^{S} x_i(A\mathbf{x})_i = 0$, and the replicator equation is the antisymmetric Lotka-Volterra equation that we study here (see mathematical definition in equation (17)) [20, 37].

1.3.2 Stochastic dynamics

One may also formulate an evolutionary game as a stochastic process in a finite population (that is, $N < \infty$ individuals in the population) and directly obtain the master equation (1) with transition rates (2); see [20]. In such a set-up, every individual plays a fixed strategy out of the set of S strategies $\{E_1, E_2, \ldots, E_S\}$. Individuals update their strategy due to pairwise interactions with individuals of different strategies according to the dominance rules of a prescribed game. For example, if strategy E_j is dominated by strategy E_i, the rate with which any individual of strategy E_j adopts strategy E_i is given by $\Gamma_{i\leftarrow j} = r_{ij}N_iN_j$ (if there are N_i individuals playing strategy E_i and N_j individuals of strategy E_j). Here, the rate constant r_{ij} characterizes the dominance relation between the pair of strategies. Furthermore, one may also allow for spontaneous switching of individuals between strategies. If a single individual playing

strategy E_j randomly switches to strategy E_i at rate $r_{ij}s_{ij}$, the rate with which any individual of strategy switches from strategy E_j to E_i is given by $\Gamma_{i\leftarrow j} = r_{ij}s_{ij}N_j$. In total, these two processes (interaction between strategies and switching of strategies) result in the master equation (1) with total transition rates (2).

1.4 Condensation into quantum states in driven-dissipative systems of non-interacting bosons

Interestingly, the master equation (1) is also encountered in the theory of open quantum systems [7, 38]. In this section, we first provide a brief overview of the recently proposed set-up of a driven-dissipative many-body quantum system whose dynamics is governed by the master equation (1). Afterwards, we sketch some steps of its derivation by outlining both the Floquet-Born-Markov theory of a single quantum particle and the derivation of the stochastic many-body dynamics that is obtained for non-interacting particles.

The following notes are not a rigorous derivation of the master equation from the quantum many-particle system; they are rather meant to give a flavor of the methods and assumptions involved in deriving the master equation. All details of the derivation can be found in the manuscript of Vorberg *et al.* [7] and the references therein. For a pedagogical introduction and a detailed background, the reader is referred to the PhD thesis of Waltraut Wustmann [39]. The general theory of master equations in open quantum systems is discussed, for example, in the textbooks of Breuer and Petruccione [38], and Gardiner and Zoller [40].

1.4.1 Overview

The set-up and the effective picture of the many-particle open quantum system that we consider here are sketched in figure 1(A): A system of N *non-interacting* bosons is weakly coupled to a heat bath and driven by an external time-periodic force (a so-called Floquet system) [41–43]. For such an open quantum system with periodic energy input, one can systematically eliminate the degrees of freedom of the bath. On a coarse-grained time scale (upon employing the Born and Markov approximations [38, 44]), the density matrix of the system becomes diagonal in the time-periodic Floquet basis; see below. In other words, in the suitable set of Floquet states $\{E_1, \ldots, E_S\}$, the dynamics becomes incoherent and is described by a classical stochastic process. In this effective description, the bosons transition incoherently between the quantum Floquet states. Nevertheless, the transition rates still reflect the quantum statistics of the bosons. The more bosons occupy a quantum Floquet state, the higher is the rate for a boson to transition into this state, reflecting the fact that bosons tend to congregate. The rate constants of these transitions are determined by microscopic properties of the system, the heat bath, and the coupling between the two. These

rate constants determine how condensation proceeds in the nonequilibrium set-up. In effect, energy inflow into the system via the periodic driving force and dissipation of energy into the weakly-coupled heat bath may guide the bosons not only into a single, but also into multiple condensates [7, 20].

Let us now explain in greater detail some of the steps to derive the master equation (1) from the driven-dissipative system of non-interacting bosons.

1.4.2 Floquet-Born-Markov theory of a single quantum particle

As a first step, one derives the master equation for a *single* quantum particle (boson or fermion), which is driven periodically in time and weakly coupled to a heat bath. Its derivation is outlined below in point (i) and (ii). As a result, the master equation for a single quantum particle is obtained as follows:

$$\partial_t p_i(t) = \sum_{j=1}^{S} \Big(r_{ij} p_j(t) - r_{ji} p_i(t) \Big) \,. \tag{10}$$

Here, the occupation probabilities $p_i(t)$ are the diagonal elements of the single-particle density operator, $p_i(t) = \langle E_i(t)|\rho_{ij}|E_i(t)\rangle$, in the basis of the so-called Floquet states $|E_i(t)\rangle$. The parameters $\{r_{i,j}\}_{i,j}$ characterize the rate of the quantum particle to transition between two states. The off-diagonal elements of the density matrix decay in the Floquet basis such that the dynamics becomes incoherent. To emphasize the fact that the quantum dynamics becomes incoherent, the master equation (10) is typically called a "classical" master equation in the context of open quantum systems. The derivation of equation (10) is sketched in the following.

(i) To derive the single-particle master equation (10), it is suitable to formulate the equation of motion for the system coupled to the bath in terms of the von-Neumann equation [38]:

$$i\hbar\, \partial_t \rho_{\mathrm{tot}}(t) = [H_{\mathrm{tot}}(t), \rho_{\mathrm{tot}}(t)] \,, \tag{11}$$

where $[A, B] = AB - BA$ denotes the commutator of the two operators A and B. The total Hamiltonian $H_{\mathrm{tot}} = H_{\mathrm{s}} + H_{\mathrm{b}} + \lambda H_{\mathrm{sb}}$ is a sum of:

- H_{s}, which denotes the Hamiltonian of the free particle with a time-periodic external potential such that $H_{\mathrm{s}}(t) = H_{\mathrm{s}}(t + \tau)$ (τ denotes the period of the driving potential),

- H_{b}, which is the free Hamiltonian of the bath that is typically assumed to be composed of free harmonic quantum oscillators, and

– H_{sb}, which denotes the interaction Hamiltonian between system and bath (the weak coupling parameter $\lambda \ll 1$ serves as a perturbation parameter in the Born approximation; see below).

Typically, one is only interested in how the system's degrees of freedom evolve in time and not in how the total system (system combined with the coupled heat bath) evolves. To this end, the reduced density operator of the system is introduced by integrating out the bath's degrees of freedom:

$$\rho_{\mathrm{s}}(t) = \mathrm{Tr}_{\mathrm{b}}\left(\rho_{\mathrm{tot}}(t)\right) \ . \tag{12}$$

From the equation of motion of the total system (11), one can derive a closed integro-differential equation that governs the temporal evolution of ρ_{s}. This equation can be simplified by employing the Born-Markov approximation. Upon assuming a weak interaction between bath and system (Born approximation, $\lambda \ll 1$), it is assured that the system does not feed back on the heat bath. Further assuming that correlations in the bath decay on a much faster time scale than significant changes in ρ_{s} take place (Markov approximation), the equation of motion for ρ_{s} becomes local in time. In other words, the temporal evolution of ρ_{s} on a suitably chosen coarser time scale depends only on the current state and not on its past. The resulting equation for ρ_{s} under the Born-Markov approximation is sometimes also referred to as a quantum master equation. Note that the derivation starting out from the von-Neumann equation of the density operator of the total system is systematic in that degrees of freedom of the bath are systematically integrated out from the total system and, thus, is more feasible than starting out from the Schrödinger equation of the total system.

We also note that the framework outlined above is similar, for example, to the derivation of the Brownian motion of a classical test particle suspended in a heat bath of harmonic oscillators; see, for example, [45].

(ii) The next step lies in representing the reduced density operator of the system ρ_{s} in the Floquet basis of the system's Hamiltonian H_{s}. The Floquet basis is obtained as follows. The single-particle Schrödinger equation with a time-periodic Hamiltonian $H_{\mathrm{s}}(t) = H_{\mathrm{s}}(t+\tau)$,

$$i\hbar\partial_t|\psi(t)\rangle = H_{\mathrm{s}}(t)|\psi(t)\rangle \ , \tag{13}$$

admits a complete set of solutions of the form $|\psi_i(t)\rangle = e^{-iE_i t/\hbar}|E_i(t)\rangle$ with $i = 1,\ldots,S$ [46]. The values $E_i \in \mathbb{R}$ are referred to as the quasi-energies. The so-called Floquet states $|E_i(t)\rangle = |E_i(t+\tau)\rangle$ are time-periodic, the set $\{|E_i(t)\rangle\}_{i=1,\ldots,S}$ is a complete orthonormal basis at every instant in time. Flo-

quet states can be regarded as the temporal analogue to Bloch states in a spatially periodic quantum system.

Upon formulating the dynamics of ρ_s in this Floquet basis of the time-periodic single-particle Hamiltonian,

$$\rho_s(t) = \sum_{i,j=1}^{S} p_{ij}(t)|E_i(t)\rangle\langle E_j(t)| \,, \tag{14}$$

one derives a closed equation of motion for the matrix elements of the reduced density operator $\{p_{ij}(t)\}_{i,j}$. It turns out that, in this Floquet basis, the off-diagonal elements decay to zero. Note that the off-diagonal elements of the density matrix are sometimes called the coherences because they describe the superposition of quantum states. Because the decay of the off-diagonal elements is much faster than the relaxation dynamics of the diagonal elements in this set-up, coherence in the quantum system studied here becomes effectively negligible. In total, only the diagonal entries of the reduced density operator are relevant; and the classical, single-particle master equation (10) governs the dynamics of the system. All microscopic properties of the set-up (for example, the period of the driving potential, the temperature and spectral density of the bath, and the coupling strength between system and bath) are encoded in the rate constants $\{r_{ij}\}_{i,j}$ that characterize the transition of a single particle between two states [7].

Classical master equations such as in equation (10) are sometimes referred to as Pauli master equations and arise, for example, in models of quantum optics [40, 47, 48]. In this context, one typically needs to assume the rotating-wave approximation to obtain decoherence from a theoretical point of view [38, 40], while, notably, the off-diagonal elements of the density matrix decay without further ad-hoc assumptions in a Floquet system as considered above.

1.4.3 Stochastic many-body system of non-interacting particles

The equation of motion (10) for the diagonal elements of the reduced density operator of the single quantum particle readily generalizes to the many-particle master equation (1). However, care has to be taken in generalizing the single-particle transition rates to the N-particle system. For a single particle, the transition rate for a single particle from state E_j to state E_i is given by:

$$\Gamma_{i\leftarrow j} = r_{ij} \geq 0 \,. \tag{15}$$

For N *non-interacting* particles, the total transition rate for a transition of one particle from state E_j to state E_i is obtained as:

$$\Gamma_{i\leftarrow j}(N_i, N_j) = r_{ij}(1 + qN_i)N_j \,, \tag{16}$$

where the factor q characterizes the different quantum statistics of the particles under consideration:

- The case $q = 1$ occurs for bosons and follows from the bosonic cummutator relation $[b_i, b_j^\dagger] = b_i b_j^\dagger - b_j^\dagger b_i = \delta_{ij}$. Heuristically speaking, if a particle transitions from state E_j to state E_i, a particle is created in state E_i and annihilated in state E_j. Thus, the relevant transition operator with non-vanishing overlap is given by $b_j^\dagger b_j b_i b_i^\dagger$, which can be written as $N_j(1 + N_i)$ by employing the definition of the particle number operator $N = b^\dagger b$ and the bosonic cummutator relation. Therefore, the bosonic quantum statistics is the reason why the factor $(1 + N_i)$ occurs in the total transition rate (16), and the resulting stochastic many-particle process can be understood as an inclusion process.
- With the same reasoning, it follows $q = -1$ for fermions because of the fermionic anticummutator relation $\{f_i, f_j^\dagger\} = f_i f_j^\dagger + f_j^\dagger f_i = \delta_{ij}$. The operator characterizing the relevant particle transition is given by $f_j^\dagger f_j f_i f_i^\dagger = N_j(1-N_i)$. The resulting stochastic many-particle process is an exclusion process because a transition to a state that is already occupied is forbidden.
- The case $q = 0$ captures the characteristics of classical particles, and a Poisson process is obtained for the stochastic many-particle process (1) in this case.

1.4.4 Summary

In essence, within the framework of Floquet-Born-Markov theory, the dynamics of the quantum many-particle system becomes incoherent on a coarse-grained time scale. The effective dynamics of the periodically driven-dissipative system of non-interacting bosons can be described in terms of the classical master equation (1) with $\{s_{ij}\}_{i,j} = 1$ in the transition rates (2). States correspond to quantum Floquet states (labeled by $E_1, \ldots, E_S$) and the fraction of bosons in a specific quantum state E_i is given by the concentration x_i. The total rate $\Gamma_{i\leftarrow j}$ for the transition of a boson from state E_j to E_i depends linearly on the number of bosons in the departure state (N_j) and the arrival state ($N_i + 1$), and on the rate constants $\{r_{ij}\}_{i,j}$. These rate constants determine the condensation of the many-particle process as we show next.

2 The leading-order time scale of the stochastic process: Condensation and coexistence in the antisymmetric Lotka-Volterra equation (ALVE)

In this section, we summarize our results on the long-time behavior of the antisymmetric Lotka-Volterra equation (see definition in equation (17) below) with focus on condensation and coexistence scenarios. All details can be found in our manuscripts that are reprinted in sections 4 and 5.

2.1 Mathematical definition of the ALVE

The *antisymmetric Lotka-Volterra equation* (ALVE) is mathematically defined for a system of S dynamical variables as motivated in section 1.2. The concentration or mass in state E_i is denoted as x_i and the vector of state concentrations as $\mathbf{x}(t) = (x_1(t), \ldots, x_S(t))$. These concentrations evolve in time according to a system of nonlinearly coupled ordinary differential equations of first order in time, the ALVE:

$$\frac{\mathrm{d}}{\mathrm{d}t} x_i(t) = x_i(t) \sum_{j=1}^{S} a_{ij} x_j(t) = x_i(t)(A\mathbf{x}(t))_i \,, \quad \text{for all } i = 1, \ldots, S \,. \tag{17}$$

The matrix $A = \{a_{ij}\}_{i,j} \in \mathbb{R}^{S\times S}$ is antisymmetric (or skew-symmetric), that is $a_{ij} = -a_{ji}$. We are interested in the temporal evolution of initial concentrations that are strictly positive and normalized, that is, $\mathbf{x}(t = 0) =: \mathbf{x}_0$ lies in the open simplex $\in \Delta_{S-1}$. In other words, $\mathbf{x}_0 \in \mathbb{R}^S$ such that $x_{i,0} > 0$ for all $i = 1, \ldots, S$ and $\sum_{i=1}^{S} x_{i,0} = 1$. For brevity, the time variable t is not explicitly written out in most of the following derivations.

First intuition of the long-time behavior

The antisymmetric matrix A defines the set of control parameters of the ALVE. It specifies how mass is exchanged between the S states through pairwise interactions. Mass in state E_i changes through interaction with state E_j as $\partial_t x_i \sim a_{ij} x_i x_j$. A negative matrix entry $a_{ij} < 0$ means that mass is transported from state E_i to E_j. Concomitantly, state E_j gains this mass through $\partial_t x_j \sim -a_{ij} x_j x_i = a_{ji} x_j x_i$ with $a_{ji} > 0$. Since no other interactions are defined by the ALVE (17), the total mass is conserved over time:

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_{i=1}^{S} x_i = \sum_{i,j=1}^{S} a_{ij} x_i x_j \overset{A=-A^T}{=} -\sum_{i,j=1}^{S} a_{ji} x_i x_j \overset{\text{relabel } i,j}{=} -\sum_{i,j=1}^{S} a_{ij} x_i x_j = 0 \,. \tag{18}$$

Consequently, the ALVE (17) defines a trajectory not leaving the open simplex, that is $\mathbf{x}(t) \in \Delta_{S-1}$ for all times [21]. If the dynamics is started on the boundary of the simplex, $\mathbf{x}_0 \in \partial\Delta_{S-1} = \overline{\Delta}_{S-1} \backslash \Delta_{S-1}$, it remains restricted to the boundary. Thus, the natural question about the long-time behavior of a state concentration x_i is whether it remains bounded away from zero, whether it approaches zero, or whether it expresses any other qualitatively different behavior (such as, for example, a heteroclinic orbit). All properties of the ALVE (17) such as mass conservation, coexistence, and condensation that are discussed in this thesis, can be ultimately traced back to its quadratic interaction structure and the antisymmetry of the matrix A.

An antisymmetric matrix can also be interpreted as the antisymmetric adjacency matrix of a weighted network (also referred to as the skew adjacency matrix of a weighted directed graph), see figure 2. States in the ALVE correspond to nodes of the weighted network and entries of the antisymmetric matrix A characterize the links between nodes. Some basic facts from linear algebra on antisymmetric matrices that are relevant to our study are collected in appendix A.

2.2 Condensation and coexistence in the ALVE

2.2.1 Overview

Here we show that, depending on the entries of the antisymmetric matrix, a state concentration may either vanish for long times ($x_i(t) \to 0$ as $t \to \infty$), in which case E_i is referred to as a *depleted state* ("depletion" or "extinction"), or it may remain bounded away from zero for all times ($x_i(t) \geq Const > 0$ for all times t), in which case E_i is called a *condensate* ("condensation" or "survival"); see below and [20, 37]. Other cases than being a condensate state or a depleted state do not occur in the ALVE. Whether a state is a condensate or becomes depleted is independent of the initial conditions $\mathbf{x}_0$ and depends only on the antisymmetric matrix A. However, the dynamics within the attractive manifold that is spanned by the condensate states depends on both the initial conditions and the antisymmetric matrix A.

2.2.2 Coexistence of all states

Before we discuss condensation, let us first discuss situations in which none of the states become depleted. We refer to these scenarios as the *coexistence of all states* in the ALVE (17) [20, 36, 49, 50]. All states are condensates in this case. In mathematical terms, the trajectory of all state concentrations stays away from the boundary of the $(S-1)$-simplex by a non-vanishing distance for all times. In the context of evolutionary game theory, coexistence of all states translates to a situation in which none of the strategies goes extinct. Despite the interactions between the agents of the population, all strategies remain present for all times.

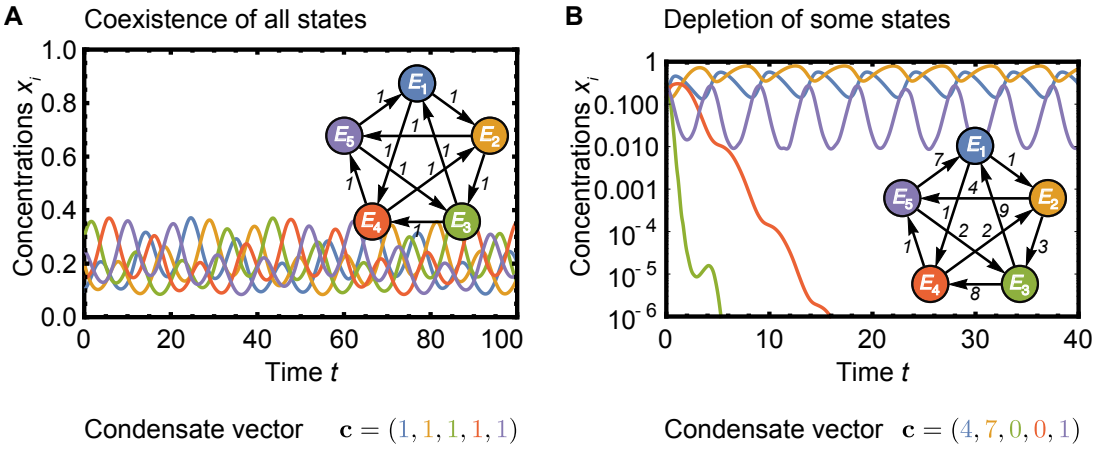


Figure 3. Coexistence and condensation in the antisymmetric Lotka-Volterra equation are determined by the condensate vector of the antisymmetric matrix. (A) Coexistence of all states is obtained if and only if the condensate vector of the antisymmetric matrix A is a strictly positive kernel element. (B) The condensate vector determines the set of states that become the condensates. In the depicted example, states E_3 and E_4 become depleted exponentially fast as indicated by the zero-entries of the condensate vector of the antisymmetric matrix.

We now show that a strictly positive kernel of the matrix A implies coexistence of all states in the ALVE. First, let $\mathbf{p}$ be any element of the kernel (or nullspace) of A, that is, $\mathbf{p} \in \mathrm{Ker}(A)$ such that $A\mathbf{p} = 0$. It follows that the collective quantity defined as the Kullback-Leibler divergence (or relative entropy) of the kernel element $\mathbf{p}$ to the trajectory $\mathbf{x}$,

$$D(\mathbf{p}||\mathbf{x}(t)) = \sum_{\substack{i=1 \\ (p_i \neq 0)}}^{S} p_i \log\left(\frac{|p_i|}{x_i(t)}\right) , \tag{19}$$

is conserved under the dynamics of the ALVE (17) as one confirms straightforwardly ($\frac{\mathrm{d}}{\mathrm{d}t} D(\mathbf{p}||\mathbf{x}) = \sum_i (A\mathbf{p})_i x_i = 0$). Furthermore, independent kernel elements of A correspond to independent conserved quantities (see pages 27–29 for details). Therefore, the dimension of the kernel, $\dim \mathrm{Ker}(A)$, determines how many such conserved quantities of form $D(\mathbf{p}||\mathbf{x})$ exist. Note that the relative entropy defined in equation (19) is essentially the logarithm of the conserved quantity used in our publication [36]. This definition makes a comparison possible with the case of depletion of some of the states (see equation (22)).

We call a kernel element $\mathbf{p}$ with $\mathbf{p} \in \mathrm{Ker}(A)$ and $p_i > 0$ for all $i = 1, \ldots, S$ a "strictly positive kernel element". For such a kernel element, the relative entropy (19) is conserved and positive for all times: $0 < D(\mathbf{p}||\mathbf{x}(t)) = D(\mathbf{p}||\mathbf{x}(0)) < \infty$. Therefore,

none of the state concentrations vanishes (otherwise, D would diverge in contradiction to the boundedness of D). In other words, if there exist strictly positive elements in the kernel of the antisymmetric matrix A, none of the states becomes depleted and all states coexist in the ALVE (17).

We found that the dynamics of the ALVE is restricted to a manifold whose dimension is at most $S-1-\dim \mathrm{Ker}(A)$ (the term -1 arises because of the conservation of the total concentration $\sum_i x_i = 1$, see equation (18)). Because the rank of an antisymmetric matrix is always even, S and $\dim \mathrm{Ker}(A)$ have the same parity. Thus, $S - 1 - \dim \mathrm{Ker}(A)$ is odd; see also appendix A. When the kernel of A contains a strictly positive element, we showed that the manifold on which the dynamics take place is diffeomorphic to an odd-dimensional sphere; see our manuscript in section 4. With these insights, periodic orbits of the ALVE can be constructed in arbitrary dimensions S if the kernel of the antisymmetric matrix A has dimension $S - 2$ and contains strictly positive kernel elements [36].

2.2.3 Condensation into some of the states

To determine the set of condensates in the case of a general antisymmetric matrix (that is, if the kernel of A does not contain strictly positive kernel elements), we proceed in two steps: (i) First, we exploit an algebraic property of antisymmetric matrices. (ii) Second, we connect this algebraic property to the long-time dynamics of the ALVE via a collective quantity that has the same functional form as the conserved quantity in equation (19).

(i) *Condensate vectors of antisymmetric matrices.*
Given an antisymmetric matrix A with real entries, there exist so-called condensate vectors $\mathbf{c}$ fulfilling the following properties for a unique index set $I \subseteq \{1, \ldots, S\}$ (see linear programming theory by Kuhn and Tucker [51], and figure 1(C) for an illustration):

$$c_i > 0 \text{ and } (A\mathbf{c})_i = 0\,, \qquad \text{for all } i \in I \tag{20}$$

$$c_i = 0 \text{ and } (A\mathbf{c})_i < 0\,, \qquad \text{for all } i \in \bar{I} = \{1, ..., S\}\backslash I\,. \tag{21}$$

Condensate vectors can be thought of as generalized positive and attractive kernel elements. They are strictly positive kernel elements on the index set I of the submatrix A_I (matrix built from A by only including rows and columns whose indices lie in I), and they are zero vectors on the index set $\bar{I}$ and at the same time attractive in that $(A\mathbf{c})_{\bar{I}} < 0$. The existence of such condensate vectors is not intuitive at first sight and is, indeed, special to antisymmetric matrices [51]. Even upon normalization (that is, $\sum_i c_i = 1$), there may exist linearly independent condensate vectors for a given antisymmetric matrix A, depending on the degeneracy of the kernel of A_I. However,

the index set of positive entries of $\mathbf{c}$, that is, the set $I = \{i \in \{1, ..., S\} : c_i > 0\}$, is unique for a given antisymmetric matrix A.

(ii) *Lyapunov function of the ALVE.*
To connect condensate vectors with the long-time dynamics of the ALVE, we introduce a collective quantity in the spirit of the conserved quantities (19) from above. In particular, we define the Kullback-Leibler divergence (or relative entropy) $D(\mathbf{c}||\mathbf{x})$ of an arbitrarily chosen condensate vector $\mathbf{c}$ of A to the state concentrations $\mathbf{x}(t)$ as:

$$D(\mathbf{c}||\mathbf{x}(t)) = \sum_{\substack{i=1 \\ (c_i \neq 0)}}^{S} c_i \log\left(\frac{c_i}{x_i(t)}\right) = \sum_{i \in I} c_i \log\left(\frac{c_i}{x_i(t)}\right) . \tag{22}$$

Note the asymmetry in the definition of D: We consider the relative entropy of $\mathbf{c}$ to $\mathbf{x}$ and *not* the relative entropy of $\mathbf{x}$ to $\mathbf{c}$ as one might try first from a statistical mechanics point of view. The collective quantity D in equation (22) is not conserved over time, but it is a Lyapunov function of the ALVE (17). In other words, the value of D decreases over time as one confirms directly ($\frac{\mathrm{d}}{\mathrm{d}t} D(\mathbf{c}||\mathbf{x}) = \sum_{i \in \bar{I}} (A\mathbf{c})_i x_i < 0$). Due to the definition of D as a relative entropy and due to the Lyapunov property, D is bounded from below by 0 and from above by $D(t = 0)$. Similarly to the coexistence case, one concludes that all states with index $i \in I$ remain bounded away from 0 for all times, that is, $x_i(t) \geq Const > 0$ for all $i \in I$ and for all t (otherwise, D would diverge in contradiction to the boundedness of D). With further arguments from analysis exploiting the boundedness of D, it is possible to show that all other states with index $i \in \bar{I}$ become depleted, that is, $x_i(t) \to 0$ as $t \to \infty$ for all $i \in \bar{I}$ [20]; see section 5 for details.

In summary, condensation and depletion in the ALVE (17) are determined by an algebraic property of the antisymmetric matrix A. The set of condensates I is uniquely determined by the antisymmetric matrix alone in terms of its condensate vectors. All states with indices $i \in I$ become condensates, while all states with indices $i \in \bar{I}$ become depleted for $t \to \infty$. This selection of condensates in the ALVE is, notably, independent of the initial conditions and proceeds exponentially fast in generic cases; see the Methods section of [20] for details. Furthermore, the Lyapunov function $D(\mathbf{c}||\mathbf{x})$ becomes a conserved quantity of the form $D(\mathbf{p}||\mathbf{x})$ for long times. When all states coexist, the set of condensates is given by $I = \{1, \ldots, S\}$ and all condensate vectors $\mathbf{c}$ are strictly positive kernel elements of the antisymmetric matrix A, see equation (20). Therefore, coexistence of all states occurs if and only if the kernel of the antisymmetric matrix A contains strictly positive elements.

3 Thermodynamic equilibrium of the stochastic process

In this section, we provide an algebraic characterization of the conditions under which the stationary state of the stochastic process defined by the master equation (1) with transition rates (2) is in thermodynamic equilibrium. The presentation of these results are meant to stimulate and motivate further research on the properties of the stationary state and on its approach when the stochastic process is *out of* thermodynamic equilibrium.

A stochastic process is said to be in thermodynamic equilibrium when the stationary process is reversible. In other words, the behavior of the stochastic process does not change upon reversing the direction of time. Thermodynamic equilibrium holds if and only if the detailed balance conditions are fulfilled [52, 53]. The implications of detailed balance can be inferred most easily at the level of the master equation (1). Kolmogorov's criterion is equivalent to the detailed balance condition and involves the transition rates alone. Kolmogorov's criterion requires that forward and backward processes must be equally likely for all possible cycles of configurations in a system [52, 53].

For simple cycles in configuration space,

$$E_1 \to E_2 \to \ldots \to E_{S'} \to E_1 \,, \tag{23}$$

Kolmogorov's criterion reads as follows:

$$\begin{aligned}&\Gamma_{1\leftarrow S'}(N_1, N_{S'}+1)\Gamma_{S'\leftarrow S'-1}(N_{S'}, N_{S'-1}+1)\cdots\Gamma_{3\leftarrow 2}(N_3, N_2+1)\Gamma_{2\leftarrow 1}(N_2, N_1+1)\\ &= \Gamma_{1\leftarrow 2}(N_1, N_2+1)\Gamma_{2\leftarrow 3}(N_2, N_3+1)\cdots\Gamma_{S'-1\leftarrow S'}(N_{S'-1}, N_{S'}+1)\Gamma_{S'\leftarrow 1}(N_{S'}, N_1+1)\,.\end{aligned} \tag{24}$$

For the stochastic process (1) with transition rates of the form $\Gamma_{i\leftarrow j} = r_{ij}(N_i+1)N_j$ (that is, the bosonic case in equation (16)), the above condition simplifies to:

$$r_{1S'}r_{S',S'-1}\cdots r_{32}r_{21} = r_{12}r_{23}\cdots r_{S'-1,S'}r_{S',1}\,. \tag{25}$$

Via induction in the length of cycles in configuration space, one can show that Kolmogorov's criterion holds if and only if condition (25) holds for any finite sequence $(1, 2, \ldots, S')$ and if $r_{ij} = 0$ implies $r_{ji} = 0$. If the matrix R of rate constants fulfils the latter conditions, the stationary process is reversible and, hence, in thermodynamic equilibrium. The equilibrium distribution of the stochastic process can be expressed in a Gibbs-Boltzmann form in that case [52].

In linear algebra, a matrix R, for which condition (25) is fulfilled for any finite sequence $(1, 2, \ldots, S')$ and for which $r_{ij} = 0$ implies $r_{ji} = 0$, is called symmetrizable. A matrix is symmetrizable if and only if there exists an invertible (that is, non-singular) diagonal matrix V such that VR is symmetric. In general, for every square

matrix R there exists an invertible matrix W such that WR is symmetric [54]. If such an invertible matrix W can be chosen in a diagonal form V, Kolmogorov's criterion for the corresponding coupled birth-death process with bosonic transition rates holds and the stationary process is in thermodynamic equilibrium. On the other hand, if no such diagonal matrix V can be found, Kolmogorov's criterion is violated and the stochastic process does not approach thermodynamic equilibrium (that is, the probability current between different configurations does not vanish). Therefore, thermodynamic equilibrium holds for the stationary process if and only if the rate matrix R is symmetrizable.

This result generalizes previous statements by Vorberg *et al.* about thermodynamic equilibrium in this stochastic process [7]. In their study, a system was considered with single-particle energies $E_1 < E_2 < E_3 \ldots < E_S$. Furthermore, the system was assumed to be in weak contact with a thermal bath of inverse temperature β. For such a set-up, the entries of the rate matrix obey $r_{ji}/r_{ij} = \exp(\beta(E_i - E_j))$ and it was shown that detailed balance holds in this case. As one easily checks, conditions (25) are fulfilled for this choice of rate constants.

How the stationary distribution looks like if the stochastic process is out of thermodynamic equilibrium, how the stationary distribution is approached once condensation has set in, and on which timescales this approach takes place may be interesting questions for future research.

4 Publication

Coexistence and survival in conservative Lotka-Volterra networks

by

J. Knebel[1], T. Krüger[2], M. F. Weber[1], and E. Frey[1]

[1]Department of Physics, Arnold Sommerfeld Center for Theoretical Physics and Center for NanoScience, Ludwig-Maximilians-Universität München, Theresienstraße 37, 80333 München, Germany,

[2]Department of Mathematics, Ludwig-Maximilians-Universität München, Theresienstraße 38, 80333 München, Germany

Reprinted on pages 22–26.

Published in *Physical Review Letters* **110(16), 168106 (2013)**,

DOI: 10.1103/PhysRevLett.110.168106;

also available on arXiv:1303.7116.

Supplemental Material reprinted on pages 27–29.

Coexistence and Survival in Conservative Lotka-Volterra Networks

Johannes Knebel[1], Torben Krüger[2], Markus F. Weber[1], and Erwin Frey[1*]
[1] *Arnold Sommerfeld Center for Theoretical Physics and Center for NanoScience, Department of Physics, Ludwig-Maximilians-Universität München, Theresienstrasse 37, 80333 München, Germany*
[2] *Department of Mathematics, Ludwig-Maximilians-Universität München, Theresienstrasse 38, 80333 München, Germany*

Analyzing coexistence and survival scenarios of Lotka-Volterra (LV) networks in which the total biomass is conserved is of vital importance for the characterization of long-term dynamics of ecological communities. Here, we introduce a classification scheme for coexistence scenarios in these conservative LV models and quantify the extinction process by employing the Pfaffian of the network's interaction matrix. We illustrate our findings on global stability properties for general systems of four and five species and find a generalized scaling law for the extinction time.

PACS numbers: 87.23.Cc, 02.50.Ey, 05.40.-a, 87.10.Mn

Understanding the stability of ecological networks is of pivotal importance in theoretical biology [1, 2]. Coexistence and extinction of species depend on many factors such as inter- and intra-species interactions [3, 4], population size [5–9], and mobility of individuals [10–16]. An intriguing question is how the stability of ecosystems depends on the interaction network between species. Is it the topology of the network (whose links may arise through predation, competition over common resources, or mutual cooperation) that sets the level of biodiversity? And how important is the strength of a single interaction link? Stable coexistence can, for example, be observed for natural populations in non-hierarchical networks that are comprised of species that interact in a competitive and predator-prey like manner [17, 18]. By understanding the interplay between the structure of the interaction network and the strengths of its links, it is possible to reveal mechanisms that underlie this stability.

A paradigm in addressing these ecologically important questions from a theoretical perspective are Lotka-Volterra (LV) models [19, 20] in which the total biomass of species is conserved. These conservative LV systems [12, 21, 22] originate in the well-mixed limit from agent-based formulations of reaction-diffusion systems, where individuals of S different species $A_1, A_2, \ldots, A_S$ compete directly with each other following the simplified reaction scheme [23]: $A_i + A_j \longrightarrow A_i + A_i$. Species A_i beats species A_j with rate k_{ij} and immediately replaces an individual of species A_j with an own offspring. Species A_j is thus degraded at the same rate such that the interaction matrix $G_S = \{k_{ij}\}_{i,j}$ is *skew-symmetric*. The interaction network can be visualized by a graph; see Fig. 1. Neglecting demographic fluctuations [24], the deterministic dynamics for the species' *concentration* vector $\mathbf{x} = (x_1, \ldots, x_S)^T$ is given by the rate equations (REs):

$$\partial_t x_i = x_i \cdot (G_S \mathbf{x})_i \ , \quad \text{for all } i = 1, \ldots, S \ . \tag{1}$$

This conservative LV model has been investigated as a prototype to understand principles of biodiversity from a theoretical point of view [8, 25]. While these systems are also of central importance to many other fields of science (e.g., plasma physics [26], evolutionary game theory [27, 28], and chemical kinetics [29]), no general scheme to classify coexistence, survival, and extinction of species has been established so far. It is frequently assumed that the topology of the interaction network alone determines coexistence of species [30, 31], i.e., that such systems can be regarded as Boolean networks [32]. Recent investigations of specific topologies indicate, however, that knowledge about the network topology may not suffice to conclude whether all species coexist or if some of them go extinct [33–35]. These questions on global stability properties have been previously addressed successfully for various particular LV systems [27, 36] and for hierarchical networks [37, 38].

In this letter, we present a general classification of coexistence scenarios in conservative LV networks with an arbitrary number of species. We elucidate the consequences of the interplay between the network structure and the strengths of its interaction links on global stability. By analyzing conserved quantities, we find conditions on the reaction rates that yield coexistence of all species. In our mathematical framework this amounts to the characterization of positive kernel elements of the interaction matrix: By employing the algebraic concept of the Pfaffian of a skew-symmetric matrix, we are able to

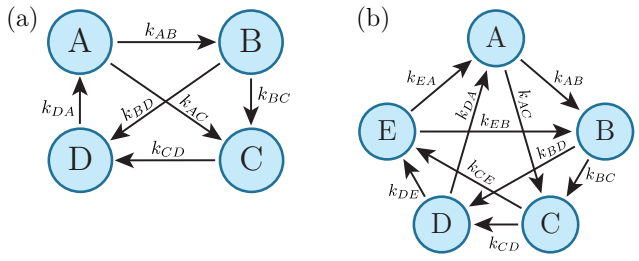


FIG. 1. (Color online) Two interaction topologies specifying the conservative LV systems. (a) The general cyclic four species systems (4SS). (b) The general cyclic five species system (5SS) as a natural extension of the rock-papers-scissors configuration (RPS) [39].

generalize previous approaches [34, 40] and to quantify the extinction process when no conserved quantities exist. We illustrate our general results for coexistence and survival scenarios of four and five species systems (4SS and 5SS), cf. Fig. 1. Moreover, we demonstrate the implications of our findings for the stability of stochastic systems: We show how the extinction time diverges with the distance to the critical rate at which coexistence of all species is observed.

First, we discuss some general results for the REs (1) before the specific interaction topologies in Fig. 1 are analyzed. In order to characterize the stability of the generic LV system, we study conserved quantities. We elaborate on the form of conserved quantities, under which conditions they exist at all, and how many conserved quantities there are for a given interaction network. Since the interaction matrix G_S is skew-symmetric, the REs (1) conserve the sum over all species' concentrations $\tau_0 = x_1 + \dots x_S$, independent of the interaction scheme. Hence, the dynamics can be normalized onto the $(S-1)$-dimensional simplex where all concentrations are non-negative and add up to 1. The vertices of the simplex correspond to the extinction of all but one species, its edges reflect the extinction of all but two species, and so on. Further conserved quantities have previously been derived as $\tau = x_1^{p_1} \dots x_S^{p_S}$ [20, 27, 40]. Interestingly, these conserved quantities can be obtained from solutions of the linear problem $G_S \mathbf{p} = \mathbf{0}$ because $\dot{\tau} = -\tau \langle G_S \mathbf{p}, \mathbf{x} \rangle$, with $\mathbf{p} = (p_1, \dots, p_S)^T$. One infers that τ is conserved if the exponent vector $\mathbf{p}$ is an eigenvector corresponding to eigenvalue 0 [40], or in other words, if $\mathbf{p}$ lies in the kernel of the matrix G_S.

Coexistence means that all concentrations stay away from the boundary of the simplex by a finite distance for all times. Since the species' concentrations are bounded to the interval $[0, 1]$, one concludes from the structure of the conserved quantity τ that *all* S species coexist if the kernel of the interaction matrix is positive, i.e., one finds an element $\mathbf{p}$ in the kernel of G_S with positive entries $p_i > 0$ for all i. Hence, to reveal coexistence scenarios in the conservative LV model, one has to characterize the kernel of the interaction matrix G_S and identify its positive elements. Note that this conclusion goes beyond stating that a positive kernel element corresponds to a stationary point in the inside of the simplex; see REs (1).

The existence of conserved quantities constrains the dynamics to a submanifold of the simplex whose dimension D_c is determined as follows. The rank of a skew-symmetric matrix is always even, because its non-zero eigenvalues are purely imaginary, conjugate pairs. The rank-nullity theorem [41] then implies that the dimension of the kernel of G_S is odd whenever S is odd, and even whenever S is even. Each linearly independent kernel element gives rise to an independent conserved quantity τ which constrains the degrees of freedom of the trajectory. Together with τ_0, one finds that the dynamics

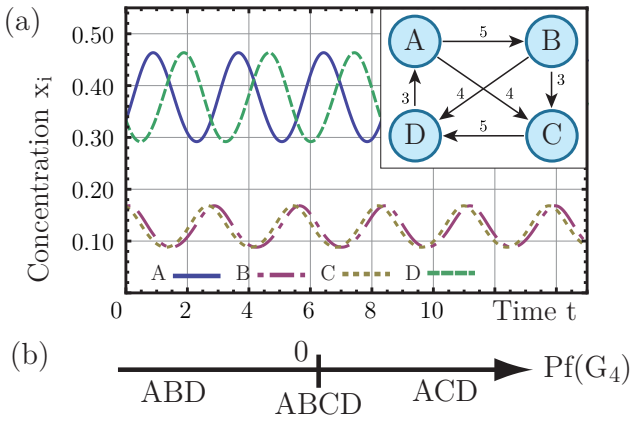


FIG. 2. (Color online) Coexistence and survival in the general cyclic 4SS are controlled by the Pfaffian of the interaction matrix. (a) For $\mathrm{Pf}(G_4) = 0$, one obtains coexistence of all species on periodic orbits. (b) Deterministic survival diagram: for $\mathrm{Pf}(G_4) < 0$, species A, B, and D survive in a stable RPS configuration, whereas A, C, and D survive for $\mathrm{Pf}(G_4) > 0$.

in case of non-stationary motion is constrained to a deformed sphere of dimension $D_c = S - 1 - \dim \mathrm{Ker}\,(G_S)$ for a positive kernel; see the Supplemental Material (SM) for mathematical details. Thus, coexistence in high-dimensional systems is generically observed on non-periodic trajectories ($D_c > 1$); see Movie M1 of SM. Only if the reaction rates are fine-tuned to a positive and maximal kernel of dimension $S-2$, the dynamics is restricted to periodic orbits ($D_c = 1$); see Fig. 2(a) and Movie M2 of SM. In particular, for $S = 3$ or 4, a positive kernel immediately implies coexistence on periodic orbits. This follows from the fact that with three species, the kernel is always one-dimensional. For the general 4SS, the dimension of the kernel of the interaction matrix is either 0 or 2. A two-dimensional, positive kernel yields coexistence on periodic orbits; see Fig. 2(a). If $\dim \mathrm{Ker}\,(G_S) = 0$, i.e., if the kernel is trivial, one observes extinction of species as detailed below.

Next, we focus on the mapping between the reaction rates in G_S and its kernel elements in order to find the stationary points. To this end, we apply the concepts of the Pfaffian and of the adjugate matrix [41, 42]. The Pfaffian is a simpler form of the determinant tailored to skew-symmetric matrices with the property that its square equals the value of the determinant. In contrast to the non-negative determinant of skew-symmetric matrices, the Pfaffian carries a sign which will turn out to be crucial for our purposes. For a skew-symmetric matrix, the Pfaffian can be computed recursively as:

$$\mathrm{Pf}\,(G_S) = \sum_{i=2}^{S} (-1)^i \cdot k_{1i} \cdot \mathrm{Pf}\,(G_{\hat{1}\hat{i}}) \ , \qquad (2)$$

with $G_{\hat{1}\hat{i}}$ being the matrix where both the first and i-th column and row have been removed from the matrix

G_S. The Pfaffian of a 2×2 skew-symmetric matrix $G_2 = \{k_{AB}\}$, is given by $\mathrm{Pf}(G_2) = k_{AB}$. For the interaction matrix corresponding to the LV network in Fig. 1(a),

$$G_4 = \begin{pmatrix} 0 & k_{AB} & k_{AC} & -k_{DA} \\ -k_{AB} & 0 & k_{BC} & k_{BD} \\ -k_{AC} & -k_{BC} & 0 & k_{CD} \\ k_{DA} & -k_{BD} & -k_{CD} & 0 \end{pmatrix} ,$$

the Pfaffian is $\mathrm{Pf}(G_4) = k_{AB}k_{CD} - k_{AC}k_{BD} - k_{DA}k_{BC}$.

The Pfaffian always vanishes for odd S as opposed to systems with an even number of species [42]. In the latter case, the Pfaffian is zero only if a constraint on the reaction rates is fulfilled. If the Pfaffian vanishes, one finds more kernel elements than just the null vector and, thus, conserved quantities of form τ exist. In the following, we distinguish between even and odd S.

For an even number of species and a two-dimensional kernel, positive kernel elements can be identified via the adjugate matrix R_S which is a generalized inverse of the interaction matrix such that $G_S \cdot R_S = -\mathrm{Pf}(G_S) \cdot \mathbb{I}_S$, with $\mathbb{I}_S$ being the identity matrix [42]. The adjugate matrix can be computed as $(R_S)_{ij} = (-1)^\sigma \mathrm{Pf}\big(G_{\hat{i}\hat{j}}\big)$ where $(-1)^\sigma$ denotes the sign of the permutation $\sigma = (i\ j\ 1 \dots \hat{i} \dots \hat{j} \dots S)$, and the columns of R_S give two independent kernel elements of G_S.

As an example, consider again the general cyclic 4SS depicted in Fig. 1(a). By setting all reaction rates equal to each other (e.g., to 1), the Pfaffian does not vanish and, therefore, not all species can coexist. Only when the rates are chosen such that $\mathrm{Pf}(G_4) = 0$, do we obtain two independent kernel elements of G_4: From its adjugate matrix, R_4, we identify $\mathbf{p}_1 = (k_{CD}, 0, k_{DA}, k_{AC})^T$ and $\mathbf{p}_2 = (k_{BD}, k_{DA}, 0, k_{AB})^T$. We infer the two conserved quantities $\tau_1 = x_A^{k_{CD}} x_C^{k_{DA}} x_D^{k_{AC}}$ and $\tau_2 = x_A^{k_{BD}} x_B^{k_{DA}} x_D^{k_{AB}}$, and conclude that the kernel is positive and coexistence occurs on periodic orbits; see Fig. 2(a). Hence, classifying LV networks in terms of their topology is incomplete; the strengths of the interaction links are crucial in general.

In general, if the Pfaffian for a system with even S is non-zero, i.e., when only the null vector lies in the kernel, coexistence of all species is not possible. Still one can quantify the extinction process by generalizing an approach of Durney et al. [34] for a system with $S = 4$ to systems composed of an arbitrary even number of species. We define the function $\rho = x_1^{q_1} \dots x_S^{q_S}$ in the same way as the conserved quantity τ, but this time choosing the exponent vector $\mathbf{q}_S = -R_S\mathbf{1}$ with $\mathbf{1} = (1, \dots, 1)^T$. It is straightforward to show that this function evolves exponentially in time:

$$\rho(t) = \rho(0) \cdot e^{-\mathrm{Pf}(G_S)\cdot t} , \qquad (3)$$

generalizing previous investigations [24, 33–35]. It is quite remarkable that ρ quantifies the global collective dynamics of systems with an arbitrary interaction topology and even S. Depending on the sign of the Pfaffian, ρ grows or decays exponentially fast with the Pfaffian of the interaction matrix as rate. Since the system's dynamics is driven towards the boundary of the simplex, one can conclude on the extinction of some species. This feature of ρ is reminiscent of a Lyapunov function; note also that ρ becomes a conserved quantity τ if the Pfaffian is zero. An interesting question for future investigations is to ask whether further quantities exist that characterize the dynamics of conservative LV networks.

For the general 4SS shown in Fig. 1(a), we find $\mathbf{q}_4 = (-k_{CD} + k_{BD} - k_{BC}, k_{CD} + k_{DA} + k_{AC}, -k_{BD} - k_{DA} - k_{AB}, k_{BC} - k_{AC} + k_{AB})^T$. The fact that $(\mathbf{q}_4)_2$ is always positive suggests that species B goes extinct for a positive Pfaffian, and that the converse holds true for $(\mathbf{q}_4)_3$ and species C for a negative Pfaffian. In both cases, the system tends to a stable rock-paper-scissors (RPS) configuration. In summary, we derive the survival diagram shown in Fig. 2(b). Interestingly, A and D always survive in this topology although D can be easily tuned to be the weakest species. We emphasize that this result depends on the sign of the Pfaffian and cannot be obtained from applying the concept of the determinant. Again, since the Pfaffian of the interaction matrix characterizes the dynamics of this 4SS, its topology alone does not determine the long-time dynamics. These findings unify previous results for other 4SS [24, 33, 34], and show that rules like "survival of the strongest" or "survival of the weakest" [25, 43] cannot be formulated in general.

For an odd number of species, the kernel of G_S is always nontrivial. In general, if $\dim\ker G_S = 1$, we determine the independent kernel element via the adjugate vector [42], $\mathbf{r}_S = \big(\mathrm{Pf}(G_{\hat{1}}), -\mathrm{Pf}(G_{\hat{2}}), \dots, \mathrm{Pf}\big(G_{\hat{S}}\big)\big)^T$, which enables us to investigate the influence of the reaction rates on the survival scenarios. For $S = 3$, only the well-studied RPS topology [8, 27] leads to a positive adjugate vector $\mathbf{r}_3$. In other words, coexistence of all three species depends only on the topology of the network. This behavior is unique to $S = 3$ and changes dramatically for systems with more than three species.

We illustrate the importance of the reaction rates for a system of five interacting species; see Fig. 1(b). This interaction topology where each species dominates two species and is outperformed by the two remaining species, recently gained attention as a natural extension of the RPS game [30, 39, 44]. For specificity, we investigate the dependence of the survival scenarios on the rate k_{AB} with which species A beats species B and chose the other rates (see Fig. 3(b), left inset) such that either five or four species survive depending on the value of k_{AB}; see Fig. 3(a). The kernel of the interaction matrix depends on k_{AB} and is characterized by the adjugate vector $\mathbf{r}_5 = (0, 0, 3k_{AB} - 15, 5 - k_{AB}, 5k_{AB} - 25)^T$. For $k_{AB} \neq 5$, the kernel is one-dimensional and non-positive, and four species survive. In contrast, for $k_{AB} = 5$, $\mathbf{r}_5$ equals the null vector which in turn means that the kernel becomes three-dimensional [42]. Since we have ensured that the

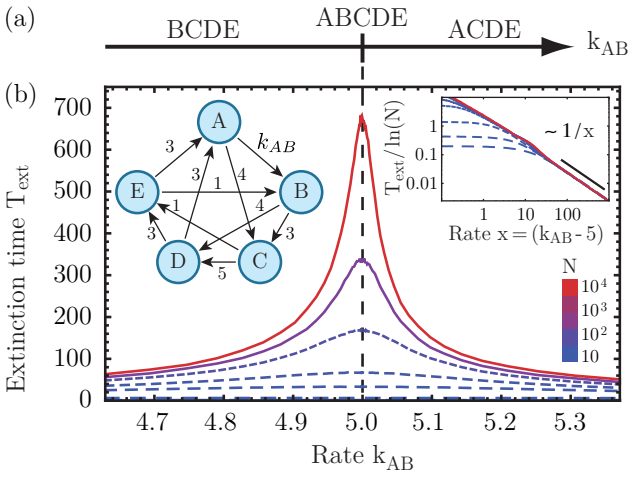


FIG. 3. (Color online) Stability of the cyclic 5SS. (a) For the interaction scheme (left inset of (b)), one obtains coexistence of all species for the critical rate $k_{AB} = 5$. (b) Stability of the stochastic system, reflected by the extinction time T_{ext}, peaks at the critical rate, which becomes more pronounced as $N \to \infty$. We find a scaling law for T_{ext} in the distance to the critical rate (right inset). Initial conditions were chosen as $\mathbf{x}(0) = 1/5 \cdot \mathbf{1}$. Larger line gap corresponds to smaller N.

kernel is also positive, we obtain coexistence of all five species on periodic orbits ($D_c = 1$).

Finally, we discuss the implications of our findings by asking how demographic noise affects the stability of stochastic LV systems. We analyze ecological LV systems with a finite number N of interacting individuals in the eye of the knowledge gained from the deterministic analysis. It has been shown that due to demographic fluctuations the system ultimately reaches an absorbing state that is characterized by the extinction of all but one species [45–48]. Moreover, the scaling behavior of the mean extinction time with the system size N characterizes the stability of the interaction network [14, 47].

As an example, we continue the discussion of the 5SS from Fig. 3(b), left inset. We have carried out extensive computer simulations employing the Gillespie algorithm [49] to measure the time T_{ext} until the first species has become extinct for different system sizes N and different reaction rates k_{AB}. The results are displayed in Fig. 3(b) and highlight the significance of the deterministic drift underlying the stochastic dynamics. We observe a peak in the extinction time as the reaction rate k_{AB} approaches the critical value $k_{\text{cr}} = 5$ for which we obtain coexistence of all species in the deterministic case. The divergence of the extinction time for $k_{AB} \to k_{\text{cr}}$ becomes more pronounced for larger system sizes as the system reaches the deterministic limit for $N \to \infty$.

A scaling analysis reveals how the extinction time peaks in the vicinity of the coexistence scenario. Near the critical rate, the extinction time scales linearly with the system size leading to neutrally stable interaction networks [8, 24, 50]. At larger distance from the critical rate, the deterministic driving force to the absorbing boundary becomes more dominant than the demographic fluctuations; see Fig. 3(b), right inset. The interplay between the stochastic and deterministic forces is reflected by the scaling law:

$$T_{\text{ext}} \propto \begin{cases} N & \text{for } k_{AB} = k_{\text{cr}} \ , \\ \frac{\ln N}{|k_{AB} - k_{\text{cr}}|} & \text{for } k_{AB} \neq k_{\text{cr}} \ , \end{cases} \tag{4}$$

which extends the linear scaling $T_{\text{ext}} \propto N$ of neutral coexistence. We observe a power-law dependence in the distance of the reaction rates to the critical rate and logarithmic scaling with N for attracting boundaries [8, 51].

The observed scaling law (4) for $k_{AB} \neq k_{\text{cr}}$ can be attributed to the exponentially fast extinction of species $x_i = x_i(0) \exp(-\alpha_i t)$; see Eq. (1). The extinction rate α_i is computed via the temporal average over the trajectory $\langle \mathbf{x} \rangle$ as $\alpha_i = -(G_S \langle \mathbf{x} \rangle)_i$, which becomes linear in the distance to the critical rate $|k_{AB} - k_{\text{cr}}|$ for large times. The logarithmic dependence on N follows by defining that a species with concentration x_i less than $1/N$ has become extinct. With this scaling behavior at hand, we are able to compare the ecological stability of different interaction networks based on our analysis of the REs (1).

In this Letter, we investigated global stability properties of conservative LV networks. By employing the Pfaffian of the interaction matrix, we revealed the relation between the reaction rates and the conditions for coexistence, and exemplified the implications for the stability of ecological networks with finite populations. We expect that our results will also stimulate further progress for the investigation of extinction scenarios. Beyond analyzing whether an ecosystem is stable or unstable, it would be highly interesting to actually predict *which* of its species ultimately survive for a general conservative LV system. This would, for example, allow us to predict the eventual outcome of an unstable version of the five species system shown in Fig. 1(b), and to formulate the conditions under which 3- or 4-species cycles are attained. First insights into these extinction dynamics will be outlined in a future publication [52]. We believe that a full characterization of general conservative LV dynamics is possible.

We would like to thank Steffen Rulands, Lukas Darnstädt, Falk Töppel, and Michal Oszmaniec for helpful discussions. This project was supported by the Deutsche Forschungsgemeinschaft in the framework of the SFB TR 12 "Symmetry and Universality in Mesoscopic Systems", and the German Excellence Initiative via the program "Nanosystems Initiative Munich" (NIM). J.K. acknowledges funding by the Studienstiftung des Deutschen Volkes.

* frey@lmu.de

[1] R. M. May, *Stability and complexity in model ecosystems* (Princeton University Press, Princeton, NJ, 1973).
[2] J. M. Montoya, S. L. Pimm, and R. V. Solé, Nature **442**, 259 (2006).
[3] G. Szabó and G. Fáth, Phys. Rep. **446**, 97 (2007).
[4] J. Mathiesen, N. Mitarai, K. Sneppen, and A. Trusina, Phys. Rev. Lett. **107**, 188101 (2011).
[5] C. Taylor, D. Fudenberg, A. Sasaki, and M. Nowak, Bull. Math. Biol. **66**, 1621 (2004).
[6] A. J. McKane and T. J. Newman, Phys. Rev. Lett. **94**, 218102 (2005).
[7] A. Traulsen, J. C. Claussen, and C. Hauert, Phys. Rev. Lett. **95**, 238701 (2005).
[8] T. Reichenbach, M. Mobilia, and E. Frey, Phys. Rev. E **74**, 051907 (2006).
[9] A. Melbinger, J. Cremer, and E. Frey, Phys. Rev. Lett. **105**, 178101 (2010).
[10] M. Nowak and R. May, Nature **359**, 826 (1992).
[11] R. Durrett and S. Levin, Theor. Popul. Biol. **46**, 363 (1994).
[12] L. Frachebourg, P. L. Krapivsky, and E. Ben-Naim, Phys. Rev. E **54**, 6186 (1996).
[13] M. Mobilia, I. T. Georgiev, and U. C. Täuber, J. Stat. Phys. **128**, 447 (2006).
[14] T. Reichenbach, M. Mobilia, and E. Frey, Nature **448**, 1046 (2007).
[15] R. Abta, M. Schiffer, A. Ben-Ishay, and N. M. Schnerb, Theoretical population biology **74**, 273 (2008).
[16] S. Rulands, T. Reichenbach, and E. Frey, J. Stat. Phys. **2011**, L01003 (2011).
[17] L. W. Buss and J. B. C. Jackson, Am. Nat. **113**, 223 (1979).
[18] B. Sinervo and C. Lively, Nature **380**, 240 (1996).
[19] A. Lotka, J. Am. Chem. Soc. **42**, 1595 (1920).
[20] V. Volterra, *Leçons sur la Théorie Mathématique de la Lutte pour la Vie*, 1st ed. (Gauthier-Villars, Paris, 1931).
[21] N. S. Goel, S. C. Maitra, and E. W. Montroll, Rev. Mod. Phys. **43** (1971).
[22] Y. Itoh, Proc. Jpn. Acad. **47**, 854 (1971).
[23] K. I. Tainaka, Phys. Rev. Lett. **63**, 2688 (1989).
[24] A. Dobrinevski and E. Frey, Phys. Rev. E **85**, 051903 (2012).
[25] M. Frean and E. R. Abraham, Proc. R. Soc. B **268**, 1323 (2001).
[26] V. Zakharov, S. Musher, and A. Rubenchik, JETP Lett. **19**, 151 (1974); S. Manakov, Sov. Phys. JETP **40**, 269 (1975).
[27] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics*, 1st ed. (Cambridge University Press, Cambridge, 1998).
[28] M. Nowak, *Evolutionary Dynamics* (Harvard University Press, Cambridge, MA, 2006).
[29] N. G. Van Kampen, *Stochastic Process in Physics and Chemistry*, 3rd ed. (Elsevier, Amsterdam, 2007).
[30] R. Laird and B. S. Schamp, J. Theor. Biol. **256**, 90 (2009).
[31] Y. Li, L. Dong, and G. Yang, Physica A **391**, 125 (2011).
[32] K. Klemm and S. Bornholdt, Phys. Rev. E **72**, 055101 (2005).
[33] S. O. Case, C. H. Durney, M. Pleimling, and R. K. P. Zia, Europhys. Lett. **92**, 58003 (2010).
[34] C. H. Durney, S. O. Case, M. Pleimling, and R. K. P. Zia, Phys. Rev. E **83**, 051108 (2011).
[35] C. H. Durney, S. O. Case, M. Pleimling, and R. K. P. Zia, J. Stat. Phys. **2012**, P06014 (2012).
[36] B. S. Goh, Am. Nat. **111**, 135 (1977); R. Redheffer, J. Differ. Equations **263**, 245 (1984); M. Zeeman, P. Am. Math. Soc. **123**, 87 (1995); Y. Takeuchi, *Global Dynamical Properties of Lotka-Volterra Systems* (World Scientific, 1996).
[37] E. Chauvet, J. Paullet, J. Previte, and Z. Walls, Math. Magazine **75**, 243 (2002).
[38] T. Cheon, Phys. Rev. Lett. **90**, 258105 (2003).
[39] The Big Bang Theory, "*The Lizard-Spock Expansion*" Season 2, Episode 8 (2008).
[40] R. K. P. Zia, e-print arXiv:1101.0018 (2010).
[41] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, 3rd ed. (SIAM, Philadelphia, 2000).
[42] C. E. Cullis, *Matrices and Determinoids*, 1st ed., Vol. I and II (Cambridge University Press, Cambridge, 1913).
[43] M. Berr, T. Reichenbach, M. Schottenloher, and E. Frey, Phys. Rev. Lett. **102**, 048102 (2009).
[44] K. Hawick, *Cycles, Diversity and Competition in Rock-Paper-Scissors-Lizard-Spock Spatial Game Agent Simulations*, Tech. Rep. (Massey University, 2011).
[45] P. L. Krapivsky, S. Redner, and E. Ben-Naim, *A Kinetic View of Statistical Physics*, 1st ed. (Cambridge University Press, Cambridge, 2010).
[46] L. Frachebourg and P. L. Krapivsky, J. Phys. A: Math. Gen. **31**, L287 (1998).
[47] T. Antal and I. Scheuring, Bull. Math. Biol. **68**, 1923 (2006).
[48] M. Parker and A. Kamenev, Phys. Rev. E **80**, 021129 (2009).
[49] D. T. Gillespie, J. Comp. Phys. **22**, 403 (1976).
[50] J. Cremer, T. Reichenbach, and E. Frey, New J. Phys. **11**, 093029 (2009).
[51] E. Frey, Physica A **389**, 4265 (2010).
[52] Manuscript in preparation.

Supplemental Material
Coexistence and Survival in Conservative Lotka-Volterra Networks

Johannes Knebel[1], Torben Krüger[2], Markus F. Weber[1], and Erwin Frey[1]

[1]*Arnold Sommerfeld Center for Theoretical Physics (ASC) and Center for NanoScience (CeNS), Department of Physics, Ludwig-Maximilians-Universität München, Theresienstrasse 37, 80333 München, Germany*

[2]*Department of Mathematics, Ludwig-Maximilians-Universität München, Theresienstrasse 38, 80333 München, Germany*

Characterization of coexistence scenarios of conservative LV networks

In this supplemental material, we extend the characterization of the qualitative behavior of coexistence scenarios for the conservative LV model as defined in the main text,

$$\partial_t x_i = x_i \cdot (G_S \mathbf{x})_i \ , \tag{1}$$

where G_S is a skew-symmetric matrix. We show that trajectories lie on odd dimensional, deformed spheres. In case of a positive and maximal kernel of the interaction matrix this behavior translates to periodic orbits.

In the main text of this letter, we define coexistence of all species for the deterministic case if the species' concentrations retain a finite distance to the absorbing boundary for all times and a given set of initial conditions. Furthermore, we show that a positive kernel of the interaction matrix implies this coexistence. We call the kernel of the interaction matrix positive if there exists an element $\mathbf{p}$ in the kernel of G_S with $p_i > 0$ for all i.

First, we show that the motion of the LV system defined in Eq. (1) is either stationary or restricted to a $(S-1-n)$-dimensional manifold M, where n denotes the dimension of the kernel of G_S. Let us assume a LV system characterized by an interaction matrix G_S with a positive, n-dimensional kernel. Note that n has the same parity as S since the rank of a skew-symmetric matrix is always even. In a sufficiently small neighborhood of a positive element $\mathbf{p}$ of the kernel of the interaction matrix, the kernel is still positive and contains n

linearly independent vectors $\mathbf{p}^{(1)}, \ldots, \mathbf{p}^{(n)}$. We can assume these vectors to be normalized such that $\sum_{i=1}^{S} p_i^{(l)} = 1$ holds true for $l = 1, \ldots, n$. They give rise to n constants of motion of the form

$$\tau_l = x_1^{p_1^{(l)}} \ldots x_S^{p_S^{(l)}} \ ,$$

as shown in the main text. In addition, the trivially conserved quantity

$$\tau_0 = \sum_{i=1}^{S} x_i = 1 \ ,$$

always exists. In order to prove linear independence of these conserved quantities, we compute:

$$0 = \sum_{l=1}^{n} c_l \cdot \mathrm{d} \log \tau_l + c_0 \cdot \mathrm{d}\tau_0 = \sum_{i=1}^{S} \left(\sum_{l=1}^{n} c_l p_i^{(l)} + c_0 x_i \right) \frac{\mathrm{d}x_i}{x_i} \ ,$$

with arbitrary real constants $c_0, c_1, \ldots, c_n$. In case of $\mathbf{x} \notin \ker G_S$, the latter equation holds true only if $c_l = 0$ for all $l = 0, \ldots, n$. This result shows the linear independence of $\mathrm{d}\tau_0, \mathrm{d}\tau_1, \ldots, \mathrm{d}\tau_n$. For $\mathbf{x} \in \ker G_S$, the motion of the system is stationary as can be seen from Eq. (1). Hence, we conclude that the motion is either stationary or restricted to a $(S-1-n)$-dimensional manifold M and the dimension of M is always odd.

In the following, we elucidate that this manifold is diffeomorphic to a sphere of dimension $(S-1-n)$. We introduce the coordinates

$$u_1 = \log x_1, \ldots, u_S = \log x_S \ ,$$

and note that the manifold M can be characterized in these new coordinates by the intersection $U \cap \{f(\mathbf{u}) = 1\}$. The latter U denotes the set

$$U = \left\{ \mathbf{u} \in \mathbb{R}^S : \sum_i p_i^{(l)} u_i = \alpha_l; \ l = 1, \ldots, n \right\} \ ,$$

which is an $(S-n)$-dimensional affine subspace given by the intersection of the hypersurfaces defined by the conserved quantities $\tau_1, \ldots, \tau_n$. The constants $\alpha_1, \ldots, \alpha_S < 0$ are determined by the initial conditions. The function

$$f : \mathbf{u} \mapsto \sum_i e^{u_i}$$

corresponds to the sum over the species' concentrations x_i and is a strictly convex function. Provided that $\min_{\mathbf{u}\in U} f(\mathbf{u}) < 1$, this property implies that the set $\{\mathbf{u} \in U : f(\mathbf{u}) \leq 1\}$ is a strictly convex bounded subset of U with open interior and smooth boundary which is, therefore, diffeomorphic to a sphere. If on the other hand $\min_{\mathbf{u}\in U} f(\mathbf{u}) = 1$, then U consists of only one point and the motion is stationary.

As a consequence, the motion along solutions of Eq. (1) is either stationary or takes place on an odd dimensional, deformed sphere in case of a positive kernel. From the observation that M cannot contain an element of the kernel of G_S for non-stationary motion, it follows via the equations of motion (1) that $\partial_t x_i \neq 0$ and together with the compactness of M, we conclude that $|\partial_t x_i| \geq const > 0$ for all i and all times. In other words, the dynamics on these odd-dimensional spheres does not come to rest.

In summary, if the kernel is positive, quasi-periodic and non-periodic trajectories are typically observed as can be seen from the corresponding Fourier spectrum; see also Movie M1 of the Supplemental Material (SM). If the dimension of the kernel is maximal, that is if $\dim\ker G_S = S - 2$, the non-stationary trajectories are restricted to deformed circles and they never come to rest. Therefore, the motion occurs on periodic orbits; see Movie M2 of the SM.

5 Publication

Evolutionary games of condensates in coupled birth-death processes

by

J. Knebel[*,1], M. F. Weber[*,1], T. Krüger[*,2], and E. Frey[1]

[*]contributed equally to this work,

[1]Arnold Sommerfeld Center for Theoretical Physics and Center for NanoScience, Department of Physics, Ludwig-Maximilians-Universität München, Theresienstraße 37, 80333 München, Germany,

[2]IST Austria, Am Campus 1, 3400 Klosterneuburg, Austria

Reprinted on pages 32–69.

Published in *Nature Communications* **6, 6977 (2015)**,

DOI: 10.1038/ncomms7977;

also available on arXiv:1504.07816.

Supplementary Information reprinted on pages 70–83.

Evolutionary games of condensates in coupled birth-death processes

Johannes Knebel[1†], Markus F. Weber[1†], Torben Krüger[2†] & Erwin Frey[1⋆]

[1] *Arnold Sommerfeld Center for Theoretical Physics and Center for NanoScience, Department of Physics, Ludwig-Maximilians-Universität München, Theresienstraße 37, 80333 München, Germany.*

[2] *IST Austria, Am Campus 1, 3400 Klosterneuburg, Austria.*

[†]These authors contributed equally to this work.

[⋆]Corresponding author: frey@lmu.de.

Final manuscript with high-resolution figures and electronic supplementary material published in: Nature Communications 6:6977 (2015), doi:10.1038/ncomms7977

www.nature.com/ncomms/2015/150424/ncomms7977/full/ncomms7977.html

(open access)

Condensation phenomena arise through a collective behaviour of particles. They are observed in both classical and quantum systems, ranging from the formation of traffic jams in mass transport models to the macroscopic occupation of the energetic ground state in ultracold bosonic gases (Bose-Einstein condensation). Recently, it has been shown that a driven and dissipative system of bosons may form multiple condensates. Which states become the condensates has, however, remained elusive thus far. The dynamics of this condensation are described by coupled birth-death processes, which also occur in evolutionary game theory. Here, we apply concepts from evolutionary game theory to explain the formation of multiple condensates in such driven-dissipative bosonic systems. We show that vanishing of relative entropy production determines their selection. The condensation proceeds exponentially fast, but the system never comes to rest. Instead, the occupation numbers of condensates may oscillate, as we demonstrate for a rock-paper-scissors game of condensates.

Condensation phenomena occur in a broad range of contexts in both classical and quantum systems. Networks such as the World-Wide-Web or the citation network perpetually grow by the addition of nodes or links and they evolve by rewiring. Over time, a finite fraction of the links of a network may be attached to particular nodes. These nodes become hubs and thereby dominate the dynamics of the whole network; they become condensate nodes[1–3]. Condensation also occurs in models for the jamming of traffic[4–7] and in related mass transport models in which particles hop between sites on a lattice[3,8,9]. A condensate forms when a finite fraction of all particles aggregates into a cluster that dominates the total particle flow. Bose-Einstein condensation, on the other hand, is a quintessentially quantum mechanical phenomenon. When an equilibrated, dilute gas

of bosonic particles is cooled to a temperature near absolute zero, a finite fraction of bosons may condense into the energetic ground state[10–12]. Long-range phase coherence builds up and quantum physics becomes manifest on the macroscopic scale[13,14].

In both the classical and the quantum mechanical context, condensation occurs when one or multiple states become macroscopically occupied (they become condensates), whereas the other states become depleted[15,16]. However, the physical origins of condensation in the above examples differ from each other. Why and how condensation arises in a particular system remains a topic of general interest and vivid research.

Here, we study condensation in two systems from different fields of research: incoherently driven-dissipative systems of non-interacting bosons and evolutionary games of competing agents. As we show below, the physical principle of vanishing entropy production governs the formation of condensates in both of these systems. The entities that constitute the respective system shall be called particles. They may be quantum or classical particles (bosons or agents). The dynamics of these particles eventually lead to condensation into particular states (quantum states or strategies). Before describing the above two systems, we now introduce the mathematical framework of our study.

On an abstract level, we consider a system of S (non-degenerate) states $E_i, i = 1, \ldots, S$, each of which is occupied by $N_i \geq 0$ indistinguishable particles, see Fig. 1a. The configuration of the system at time t is fully characterized by the occupation numbers $\mathbf{N} = (N_1, N_2, \ldots, N_S)$. This configuration changes continuously in time due to the transition of particles between states.

The total number of particles in this coupled birth-death process is conserved ($N = \sum_i N_i$). We are interested in the probability $P(\mathbf{N}, t)$ of finding the system in configuration $\mathbf{N}$ at time t. The temporal evolution of the probability distribution $P(\mathbf{N}, t)$ is governed by the classical master equation[17,18]:

$$\partial_t P(\mathbf{N}, t) = \sum_{\substack{i,j=1 \\ j \neq i}}^{S} \Big(\Gamma_{i \leftarrow j}(N_i - 1, N_j + 1) P(\mathbf{N} - \mathbf{e}_i + \mathbf{e}_j, t) - \Gamma_{i \leftarrow j}(N_i, N_j) P(\mathbf{N}, t) \Big) , \tag{1}$$

where $\mathbf{e}_i \in \mathbb{Z}^S$ denotes the unit vector in direction i (equal to one at index i, otherwise zero). The rate for the transition of particles from state E_j to E_i depends linearly on the number of particles in the departure and in the arrival state:

$$\Gamma_{i \leftarrow j} = r_{ij}(N_i + s_{ij}) N_j , \tag{2}$$

with rate constant $r_{ij} \geq 0$ and constants $s_{ij} \geq 0$.

Condensation in this framework is understood as the macroscopic occupation of one or multiple states[15,16]: We consider a state E_i as a condensate when the long-time average of the number of particles in this state scales linearly with the system size ($\langle N_i \rangle_t \sim \mathcal{O}(N)$ for large t). Hence, a condensate harbours a finite fraction of the total number of particles for large systems ($N \gg 1$). We refer to a state as depleted when its average occupation number scales less than linearly with the system size. Therefore, the fraction of particles in a depleted state vanishes in the limit of large systems.

Depending on the values of the rate constants r_{ij}, numerical simulation of equations (1) with

rates (2) reveals that all states, multiple states, or only one state become condensates when the particle density N/S is large enough to detect condensation[19]. Thus far, various questions about condensation have remained elusive for the coupled birth-death process defined by equation (1): Which of the states become condensates? How does this selection of condensates proceed? Is it possible to construct systems that condense into a specific set of condensates?

In the following, we answer these questions by illuminating the physical principle that governs the formation of multiple condensates on the leading order timescale. We show that the vanishing of relative entropy production determines the selection of condensates (see equations (3) and (4) below). We elaborate how condensate selection is determined by the rate constants r_{ij}. The condensation proceeds exponentially fast into a dynamic, metastable steady state within which the occupation numbers of condensates may oscillate. By applying our general results to systems with many states, we show that the interplay between critical properties of such networks of states[20] and dynamically stable network motifs[21] determines the selection of condensates. The results of our analysis apply to any system whose dynamics are described by the coupled birth-death processes (1) with rates (2). Before proceeding to the mathematical and numerical analysis of condensation in these processes, we now give a brief overview of such systems.

Results

Non-interacting bosons in driven-dissipative systems. The classical master equation (1) has recently been derived by Vorberg et al. in the study of bosonic systems that are dissipative and

driven by external sources[19]. For a system of non-interacting bosons that is weakly coupled to a reservoir and driven by an external time-periodic force (a so-called Floquet system)[22–24], one can eliminate the reservoir degrees of freedom (Born and Markov approximation)[25,26] and the density matrix of the system becomes diagonal (see the Supplement of the work of Vorberg et al.[19]). The effective dynamics of the bosons become incoherent and are captured on a macroscopic level in terms of the coupled birth-death processes (1) with rates $\Gamma_{i\leftarrow j} = r_{ij}(N_i + 1)N_j$ (that is all $s_{ij} = 1$ in the rates (2)). These non-equilibrium setups may not only lead the bosons into a single, but also into multiple condensates[19].

For the incoherently driven-dissipative systems described above, the state E_i denotes a time-dependent Floquet state[22–24]. The total rate $\Gamma_{i\leftarrow j}$ for the transition of a boson from state E_j to E_i depends linearly on the number of bosons in the departure state (N_j) and the arrival state ($N_i + 1$). The latter factor stems from the indistinguishability of bosons and reflects their tendency to congregate. Although, we refer to equation (1) as a classical master equation and coherence does not build up, the quantum statistics of bosons is still encoded in the functional form of $\Gamma_{i\leftarrow j}$. The rate constant r_{ij} is determined by microscopic properties of the system and the reservoir.

Condensation in the above setup is to be distinguished from Bose-Einstein condensation. Typically, studies on Bose-Einstein condensation focus on the existence of long-range phase coherence in thermal equilibrium[10–15], its kinetic formation[13,14,27–32], and the fragmentation of a coherent condensate into multiple condensates (for example when the equilibrium ground state is degenerate)[13,16]. In contrast, the classical birth-death processes (1) with rates (2) describe conden-

sation in bosonic systems that are externally driven by a continuing supply of energy, dissipate into the environment, and exhibit decoherence.

Equations of type (1) may also arise in atomic physics and quantum optics and are known as Pauli master equations[33–35]. They describe how the population of S non-degenerate energy levels changes over time when a system harbours N indistinguishable, non-interacting bosonic atoms. Such changes may occur by interactions with a radiation field that induces transitions between energy levels. A theoretical description of these transitions in terms of a Pauli master equation is appropriate if coherence is negligible. As in the previous example, the system then approaches a state in which some of the energy levels are macroscopically occupied (condensates) whereas others are depleted. More generally, whenever a rate constant r_{ij} governs the transition of a single boson from a state E_j to E_i, the rates (2) with $s_{ij} = 1$ for all i and j apply if N non-interacting bosons are brought into the system[19].

Strategy selection in evolutionary game theory. The classical master equation (1) also occurs in evolutionary game theory (EGT). Historically, EGT was developed to study evolutionary processes that are driven by selection and mutation[36,37] and seeks to identify optimal strategies for competitive interactions. For example, EGT has been applied in the study of the prominent "rock-paper-scissors" game, which was proposed as a facilitator of species coexistence and has inspired both experimental and theoretical research[38–42]. Furthermore, the "prisoner's dilemma" game serves as a paradigmatic model to explore the evolution and maintenance of cooperation[43,44]. The interplay between non-linear and stochastic effects underlies the dynamics of such evolutionary games[45–50].

In EGT, one typically considers a system of N interacting agents (classical particles) who repeatedly play one fixed strategy E_i out of the S possible choices $E_1, E_2, \ldots, E_S$. In each succeeding interaction, the defeated agent adopts the strategy of its opponent. Since N_j agents playing strategy E_j can potentially be defeated by one of the N_i agents playing strategy E_i, the rate of change is $\Gamma_{i \leftarrow j} = r_{ij} N_i N_j$. If an agent who plays E_j can also spontaneously mutate into an agent who plays E_i (with rate $\mu_{ij} = r_{ij} s_{ij}$), one recovers the classical master equation (1) with rates (2).

Thus, there exists a correspondence between condensation in incoherently driven-dissipative bosonic systems and strategy selection in evolutionary game theory: the transition of bosons between states can be interpreted in terms of the interaction and mutation of agents employing evolutionary strategies. In effect, the states in an incoherently driven-dissipative setup play an evolutionary game and the winning states form the condensates.

After having introduced the above examples, we now proceed with the mathematical and numerical analysis of the classical master equation (1). We show that the dynamics of condensation change on two distinct timescales. At the leading order timescale, the dynamics are described by a set of nonlinearly coupled, ordinary differential equations (see equation (3) below), which determine the states that become condensates. We identify these states by applying concepts from EGT. After an exposition of the physical principles that underlie the condensation dynamics, implications of our general results for incoherently driven-dissipative systems are discussed.

The antisymmetric Lotka-Volterra equation. The total number of particles needed for condensation phenomena to occur is large ($N \gg 1$). In order to detect macroscopic occupancies, it is also

assumed that the particle density N/S is large. Therefore, one may approximate the classical master equation (1) by a Langevin equation for the state concentrations $x_i(t) = N_i(t)/N$ (details of the derivation are provided in Supplementary Note 1). Originally proposed for Brownian particles suspended in a liquid, the Langevin equation decomposes the dynamics of a sample trajectory of the random process into two contributions: into a deterministic drift and into noise stemming from the discreteness of particle numbers ("demographic fluctuations"). Both the demographic fluctuations and the contribution to the deterministic drift that corresponds to mutations in the EGT setting are suppressed by a small prefactor $1/N$. Therefore, these terms change the dynamics only slowly. The deterministic drift that corresponds to interactions between agents is, however, not suppressed. It thus governs the dynamics to leading order.

Hence, we find that the leading order dynamics of the condensation process (1)-(2) are described by the differential equations:

$$\frac{\mathrm{d}}{\mathrm{d}t} x_i = x_i (A\mathbf{x})_i \,. \tag{3}$$

The matrix A is antisymmetric and encodes the effective transition rates between states ($a_{ij} = r_{ij} - r_{ji}$). The constants s_{ij} that occur in the definition of the rates (2) do not change the leading order dynamics, but they become relevant on subleading order timescales.

We refer to equation (3) as the antisymmetric Lotka-Volterra equation (ALVE). It provides a description of pairwise interactions that preserve the total number of particles. Therefore, the ALVE finds a broad range of applications in diverse fields of research, in addition to the aforementioned condensation of bosons far from equilibrium. It was first studied by Volterra[51] in the context

of predator-prey oscillations in population biology[47,52,53]. In plasma physics, the ALVE describes the spectra of plasma oscillations (Langmuir waves)[54,55], and in chemical kinetics it captures the dynamics of bimolecular autocatalytic reactions[18,56–58]. In EGT, the ALVE is known as the replicator equation of zero-sum games such as the rock-paper-scissors game[47,59–61]. Table 1 summarizes all of the above analogies.

Despite the simple structure of the ALVE, it exhibits a rich and complex behaviour. In the following, we show how the mathematical analysis of the ALVE explains condensation into multiple states (condensate selection). To this end, we extend an approach for the analysis of the ALVE that was introduced in the context of EGT[59,60].

Production of relative entropy and condensate selection. Our analysis starts from a theorem in linear programming theory[62]. Given an antisymmetric matrix A, it is always possible to find a vector $\mathbf{c}$ that fulfils the following conditions: its entries are positive for indices in $I \subseteq \{1,\dots,S\}$ and zero for indices in $\bar{I} = \{1,\dots,S\} - I$, whereas the entries of $A\mathbf{c}$ are zero for indices in I and negative for indices in $\bar{I}$ (Fig. 1b). Although several vectors $\mathbf{c}$ with these properties may exist, the index set I is unique and, thus, determined by the antisymmetric matrix A. Finding such a "condensate vector" $\mathbf{c}$ is crucial for the understanding of condensate selection and of the condensation dynamics. The condensate vector has the following physical interpretations.

All condensate vectors yield fixed points of the ALVE (3). Because of the antisymmetry of matrix A, a linear stability analysis of these fixed points does not yield insight into the global dynamics (Supplementary Note 1). However, global stability properties can be inferred by showing

that the relative entropy of a condensate vector to the state concentrations,

$$D(\mathbf{c}||\mathbf{x}) = \sum_{i \in I} c_i \log\left(c_i/x_i\right), \tag{4}$$

is a Lyapunov function (note that we do not consider the relative entropy of the state concentrations to the condensate vector, but define the relative entropy vice versa). The relative entropy (4) decreases with time and is bounded from below (see Methods and Supplementary Fig. 1). Therefore, the dynamics relax to a subsystem in which relative entropy production is zero. The relaxation of relative entropy production is reminiscent of Prigogine's study of open systems in non-equilibrium thermodynamics. Indeed, we find that the system, to cite Prigogine's phrase, "settles down to the state of least dissipation"[63].

This state of least dissipation is characterized even further by the condensate vector $\mathbf{c}$. Considering the definition of the relative entropy (4) and its boundedness, it follows that every concentration x_i with $i \in I$ remains larger than a positive constant. On the other hand, states with indices in $\bar{I}$ become depleted for long times (see Methods). Therefore, we find that the condensate vector determines the selection of condensates. Positive entries of $\mathbf{c}$ correspond to states that become condensates, whereas zero entries of $\mathbf{c}$ correspond to states that become depleted. Both the set of condensates and the set of depleted states are unique (Fig. 1b) and independent of the initial conditions. Generically, the entries of the condensate vector are also unique upon normalization (its entries sum up to one) and yield the rate $|(A\mathbf{c})_i|$ at which a state E_i becomes depleted. The condensate selection occurs exponentially fast (see Fig. 2 and Methods).

After relaxation, the dynamics of the system are restricted to the condensates. In other words,

the condensates form the attractor of the dynamics. However, the dynamics in this subsystem do not come to rest. The state of least dissipation is a dynamic state with a perpetually changing number of particles in the condensates: periodic, quasiperiodic, and non-periodic oscillations are observed (Fig. 2b and Supplementary Fig. 1). In the generic case, the entries of the condensate vector represent the temporal average of condensate concentrations according to the ALVE (3). After condensate selection, the dynamics of these active condensates take place on a high-dimensional, deformed sphere[61].

An algebraic algorithm to find the condensates. Numerical integration of the ALVE (3) is neither a feasible nor a reliable method for identifying condensates (Fig. 2, Supplementary Figs. 1 and 2). Instead, we determine these states by numerically searching for a condensate vector $\mathbf{c}$. To this end, we reformulate the above conditions on $\mathbf{c}$ in terms of two linear inequalities[62]:

$$A\mathbf{c} \leq 0 \;\text{ and }\; \mathbf{c} - A\mathbf{c} > 0\,. \tag{5}$$

We solve these inequalities with a linear programming algorithm that is both reliable and efficient. The time to find a condensate vector scales only polynomially with the number of states S (see Supplementary Fig. 3 and Methods for details).

Condensation in large random networks of states. We used our combined analytical and numerical approach to study how the connectivity of a random network of states affects the selection of condensates under the dynamics of the ALVE (3). The connectivity specifies the percentage of states between which particle transitions occur[20,64]. After having generated a network with a given connectivity, the strength and direction of an allowed transition between states E_j and E_i were

determined by randomly sampling the corresponding effective rate constant $a_{ij} = r_{ij} - r_{ji}$.

Our results for condensation in large random networks of states are summarized in Fig. 3. When the connectivity of a network is zero, all of its states are isolated. Particles are not exchanged between states and none of the states becomes depleted. For an increased connectivity, isolated pairs of states are sampled in a random network. One state in an isolated pair is always depleted and the average number of condensates decreases rapidly. Upon approaching a critical connectivity, cycles and trees of all orders become embedded in a random network. This critical connectivity scales inversely with the number of states[20]. We observe that, under the dynamics of the ALVE (3), the average number of condensates becomes minimal for a connectivity that also scales inversely with the number of states (see Fig. 3g). We attribute this minimum to the interplay between the criticality of random networks and condensate selection on connected components of the network[61,65]. Embedded directed cycles are a recurring motif[21] in the remaining network of condensates after condensate selection. Above the critical connectivity, a single giant cluster is formed. On average, half the number of states in this giant cluster become condensates once the network is fully connected ($C = 1$) (Fig. 3a)[19,60,66]. Thus, our analysis emphasizes the importance of critical properties of random networks for condensate selection.

Design of active condensates. Our understanding of condensate selection can be used to design systems that condense into a particular network of states, a game of condensates. We exemplify this procedure by formulating conditions under which a system relaxes into a rock-paper-scissors (RPS) game of condensates[40]. Three particular states E_1, E_2, and E_3 in a system become a RPS

cycle of condensates if, and only if, the following two conditions are fulfilled (Fig. 4). First, the "RPS condition" requires that the rate constants between the three states form a RPS network (for example, $r_{12} > r_{21}, r_{23} > r_{32}$, and $r_{31} > r_{13}$). Second, the "attractivity condition" requires that the inflow of particles into the RPS cycle from any other state E_k is greater than the outflow to that state E_k (for all $k = 4, \ldots, S$). The values of the rate constants between the states that become depleted are irrelevant. More complex games of condensates can be designed by formulating similar conditions on the rate constants. These conditions are formulated as inequalities that depend on Pfaffians of the antisymmetric matrix A and its submatrices (see Methods)[61]. The flow of particles between states in these systems causes condensate concentrations to oscillate (Fig. 2b).

Discussion

Our findings thus suggest intriguing dynamics of condensates in systems whose temporal evolutions are captured by the classical master equation (1) with rates (2); for example in driven-dissipative systems of non-interacting bosons. Condensates observed on the leading order timescale are metastable. For longer times, relaxation into a steady state occurs[19,67]. When detailed balance is broken in the system of condensates, the net probability current between at least two states does not vanish and a non-equilibrium steady state is approached[68,69]. The simplest way of designing such condensates is illustrated by the above RPS game. In this game, detailed balance is broken, for example, when the transition of particles is unidirectional (with totally asymmetric rate constants $r_{12} > r_{21} = 0, r_{23} > r_{32} = 0$, and $r_{31} > r_{13} = 0$). For non-interacting bosons in driven-dissipative systems, the continuing supply with energy through the external time-periodic driving

force (Floquet system) and the dissipation of energy into the environment may, therefore, prevent the system from reaching equilibrium. How such systems may be realized in an experiment poses an interesting question for future research.

The transition of particles between condensates in the here studied coupled birth-death processes parallels the interaction and mutation of winning agents in evolutionary game theory, reflecting an "evolutionary game of condensates". Our results suggest the possibility of creating novel bosonic systems with an oscillating occupation of condensates. Non-interacting bosons in incoherently driven-dissipative systems are promising candidates. Since the antisymmetric Lotka-Volterra equation also arises in population biology, chemical kinetics, and plasma physics, all of our mathematical results apply to these fields as well.

Methods

Asymptotics of the antisymmetric Lotka-Volterra equation. The asymptotic behaviour of the ALVE (3) can be characterized as follows: For every antisymmetric matrix A there exists a unique subset of states $I \subseteq \{1, \ldots, S\}$ whose concentrations stay away from zero for all times, that is,

$$x_i(t) \geq Const(A, \mathbf{x}_0) > 0 \text{ for all } t \geq 0 \text{ and for every } i \in I \; . \tag{6}$$

The set I is the set of condensates. All of the other states with indices in $\bar{I} = \{1, \ldots, S\} - I$ become depleted as $t \to \infty$, that is,

$$x_i(t) \to 0 \text{ as } t \to \infty \text{ for every } i \in \bar{I} \; . \tag{7}$$

The set of condensates can be determined algebraically from the antisymmetric matrix A and does not depend on the initial conditions $\mathbf{x}_0 \in \Delta_{S-1} = \{\mathbf{x} \in \mathbb{R}^S \mid x_i > 0 \text{ for all } i,\ \sum_{i=1}^S x_i = 1\}$.

To show this result, the time-dependent entropy $D(\mathbf{c}||\mathbf{x})(t)$ of a condensate vector $\mathbf{c} = (c_1, \ldots, c_n) \in \overline{\Delta}_{S-1}$ ($c_i \geq 0$ for all i and $\sum_i c_i = 1$) relative to the trajectory $\mathbf{x}(t)$ is considered (that is, the Kullback-Leibler divergence of $\mathbf{x}(t)$ from $\mathbf{c}$), see equation (4). A condensate vector is defined via the properties (see Fig. 1b):

$$c_i > 0 \text{ and } (A\mathbf{c})_i = 0 \text{ for } i \in I, \text{ and} \tag{8}$$

$$c_i = 0 \text{ and } (A\mathbf{c})_i < 0 \text{ for } i \in \bar{I}\,. \tag{9}$$

Such a vector can always be found for an antisymmetric matrix[62]. Notably, the index set I is unique although more than one condensate vector may exist.

Considering the time derivative of the relative entropy $D(\mathbf{c}||\mathbf{x})(t)$ and employing equations (3) and (8) yields:

$$\frac{\mathrm{d}}{\mathrm{d}t} D(\mathbf{c}||\mathbf{x})(t) = -\sum_{i=1}^S c_i \frac{\partial_t x_i}{x_i} = -\sum_{i=1}^S c_i (A\mathbf{x})_i = \sum_{i=1}^S (A\mathbf{c})_i x_i = \sum_{i\in\bar{I}} (A\mathbf{c})_i x_i\,. \tag{10}$$

Since $(A\mathbf{c})_{\bar{I}} < 0$ and $\mathbf{x} > 0$, it follows that $\partial_t D(\mathbf{c}||\mathbf{x})(t) < 0$. Therefore, the relative entropy $D(\mathbf{c}||\mathbf{x})$ is a Lyapunov function if $\mathbf{c}$ is chosen in accordance with equations (8) and (9). Moreover, $D(\mathbf{c}||\mathbf{x})$ is bounded from above by $D(\mathbf{c}||\mathbf{x})(0)$ and from below by zero for all times. This can be seen from the definition of D, and from integration of equation (10) (using that $(A\mathbf{c})_{\bar{I}} < 0$ and $\mathbf{x} > 0$):

$$0 \leq D(\mathbf{c}||\mathbf{x})(t) = D(\mathbf{c}||\mathbf{x})(0) + \int_0^t ds\ \sum_{i\in\bar{I}} (A\mathbf{c})_i x_i(s) \leq D(\mathbf{c}||\mathbf{x})(0)\,. \tag{11}$$

From the definition of the relative entropy in equation (4), it follows that every concentration x_i with $i \in I$ remains larger than a positive constant, that is, $x_i(t) \geq Const(A, \mathbf{x}_0) > 0$ for all times t (if $x_i(t) \to 0$ for $i \in I$, it follows that $D \to \infty$, which contradicts the boundedness of D).

Furthermore, equation (11) implies that,

$$-\int_0^t ds\ (A\mathbf{c})_i x_i(s) \leq -\int_0^t ds\ \sum_{i\in\bar{I}} (A\mathbf{c})_i x_i(s) \leq D(\mathbf{c}||\mathbf{x})(0)\ , \tag{12}$$

for every $i \in \bar{I}$ and for all t. Therefore, concentration x_i is integrable for every $i \in \bar{I}$ ($x_i \in L^1(0,\infty)$) with the bound:

$$0 < \int_0^\infty ds\ x_i(s) \leq \frac{D(\mathbf{c}||\mathbf{x})(0)}{-(A\mathbf{c})_i} = Const(A, \mathbf{x}_0) \quad \text{for every } i \in \bar{I}\ . \tag{13}$$

Since the derivative of the concentrations is bounded from above and below, $|\partial_t x_i| = |x_i(A\mathbf{x})_i| \leq \|(A\mathbf{x})\|_\infty \leq \|A\|_{\infty\to\infty} \leq Const(A)$, one concludes that x_i is uniformly continuous ($\|A\|_{\infty\to\infty}$ denotes the operator norm of A induced by the maximum norm on $\mathbb{R}^S$). Together with the integrability (13), it follows that states with indices in $\bar{I}$ become depleted as $t \to \infty$, that is, $x_i(t) \to 0$ for $i \in \bar{I}$.

In conclusion, given an antisymmetric matrix $A = R - R^T$ via a rate constant matrix $R = \{r_{ij}\}_{i,j}$, one finds a condensate vector $\mathbf{c}$ that satisfies inequalities (8)-(9). The index set I, for which entries of $\mathbf{c}$ are positive, represents condensates. The index set $\bar{I}$, for which entries of $\mathbf{c}$ are zero, represents states that become depleted. Moreover, equation (10) implies that the relative entropy becomes a conserved quantity in the subsystem of condensates (Supplementary Fig. 1).

Temporal average of condensate concentrations. The ALVE (3) is solved implicitly by,

$$x_i(t) = x_i(0) e^{t \cdot (A \langle \mathbf{x} \rangle_t)_i} \ , \tag{14}$$

with the time average of the trajectory $\langle \mathbf{x} \rangle_t$ defined as:

$$\langle \mathbf{x} \rangle_t = \frac{1}{t} \int_0^t ds \ \mathbf{x}(s) \ . \tag{15}$$

It is shown above that $0 < Const(A, \mathbf{x}_0) \leq x_i(t) \leq 1$ holds for the states that become condensates ($i \in I$). By rearranging equation (14), one thus obtains:

$$|(A \langle \mathbf{x} \rangle_t)_i| \leq \frac{1}{t} \left| \log \left(\frac{x_i(t)}{x_i(0)} \right) \right| \leq \frac{Const(A, \mathbf{x}_0)}{t} \quad \text{for all } i \in I \ . \tag{16}$$

Note that $Const$ is used to denote arbitrary positive, time-independent constants. Therefore, the right hand side of equation (16) vanishes for $t \to \infty$. On the other hand, x_i is integrable for $i \in \bar{I}$ (equation (13)). Thus, the corresponding component of the time average converges to zero,

$$\langle x_i \rangle_t \leq \frac{Const(A, \mathbf{x}_0)}{t} \to 0 \text{ as } t \to \infty \text{ for every } i \in \bar{I} \ . \tag{17}$$

Hence, the distance of the time average $\langle \mathbf{x} \rangle_t$ to the kernel of the antisymmetric submatrix A^I converges to zero (the submatrix A^I corresponds to the system of condensates with indices in I).

Structure of a generic antisymmetric matrix. For systems with an even number of states S, the antisymmetric matrix $A = R - R^T$ generically has a trivial kernel, whereas for systems with an odd number of states, the kernel of A is generically one-dimensional. A higher dimensional kernel of A only occurs if the matrix entries are tuned[52,61,70]. As a consequence, when all of the entries above the diagonal of A are, for example, randomly drawn from a continuous probability

distribution (for example from a Gaussian distribution), all 2^S submatrices of A have a kernel with dimension of less than or equal to one.

The projection of $\mathbf{x} \in \mathbb{R}^S$ to the subspace $\mathbb{R}^{(J)} \subseteq \mathbb{R}^S$ for an arbitrary index set $J \subseteq \{1, \ldots, S\}$ is defined as $\mathbf{x}_J \coloneqq P_J \mathbf{x} \coloneqq (x_j)_{j \in J}$. In other words, entries of $\mathbf{x}_J$ are zero for indices in the complement $\bar{J}$. In the following, the short notation $A^J \coloneqq P_J A P_J$ is also used (see above). Furthermore, the set of antisymmetric matrices whose submatrices have a kernel with dimension less than or equal to one is defined:

$$\Omega \coloneqq \left\{ A \in \mathbb{R}^{S \times S} \mid A \text{ is antisymmetric and } \dim \ker A^J \leq 1 \text{ for all } J \subseteq \{1, \ldots, S\} \right\} . \tag{18}$$

The complement $\bar{\Omega}$ has measure zero with respect to the flat measure $\mathrm{d}A$ on antisymmetric matrices (the translation invariant measure, which is sigma-finite and not trivial).

For antisymmetric matrices $A \in \Omega$, the kernel can be characterized as follows[61,70]. If the number of states S is even, the kernel of A is trivial: $\ker A = \{\mathbf{0}\}$. If the number of states is odd, the kernel is one-dimensional: $\ker A = \{\mathbf{v}\}$. This kernel element can be computed analytically in terms of Pfaffians of submatrices of A:

$$\mathbf{v} = \left(\mathrm{Pf}\left(A_{\hat{1}}\right), -\mathrm{Pf}\left(A_{\hat{2}}\right), \ldots, \mathrm{Pf}\left(A_{\hat{S}}\right) \right) . \tag{19}$$

The submatrix $A_{\hat{k}} \in \mathbb{R}^{(S-1)\times(S-1)}$ denotes the matrix for which the k-th column and row are removed from A.

For antisymmetric matrices $A \in \Omega$, the normalized condensate vector $\mathbf{c}$ with $\sum_i c_i = 1$ and with properties (8), (9) is unique. The latter follows from $A^I \mathbf{c} = 0$ (equations (8) and (18)). Therefore, the condensate vector is the unique kernel vector of the subsystem of condensates whose

interactions are characterized by the matrix A^I. Furthermore, I contains an odd number of elements. To determine the condensate vector for $A \in \Omega$, one can proceed as follows. For each odd-dimensional submatrix A^I with $I \subseteq \{1, \ldots, S\}$, one computes the kernel element $\mathbf{v}$ according to equation (19) and defines the vector $\mathbf{w} \in \mathbb{R}^S$ by setting $\mathbf{w}_I = \mathbf{v}$ and $\mathbf{w}_{\bar{I}} = 0$. There exists exactly one set I for which $(A\mathbf{w})_{\bar{I}} < 0$. The corresponding vector $\mathbf{w}$ is the unique condensate vector upon normalization.

Temporal average of condensate concentrations (generic case). It was shown above that the temporal average of condensate concentrations $\langle \mathbf{x} \rangle_t$ converges to a non-negative kernel element of the antisymmetric matrix A^I. In the generic case, the condensate vector $\mathbf{c}$ is the unique kernel element of A^I upon normalization. Therefore, positive entries of $\mathbf{c}$ represent the asymptotic temporal average of condensate concentrations,

$$\| \langle \mathbf{x} \rangle_t - \mathbf{c} \|_\infty \leq \frac{Const(A, \mathbf{x}_0)}{t} \to 0 \text{ as } t \to \infty \, . \tag{20}$$

Exponentially fast depletion of states (generic case). Upon inserting equation (20) into the implicit solution (14) of the ALVE, the exponentially fast depletion of states with $i \in \bar{I}$ can be seen as follows (note that $(A\mathbf{c})_i < 0$ according to the choice of the condensate vector in equations (8)

and (9)):

$$x_i(t) = x_i(0)e^{t\cdot(A\langle\mathbf{x}\rangle_t)_i} \tag{21}$$

$$\leq x_i(0)e^{t\cdot((A\mathbf{c})_i+\|A(\langle\mathbf{x}\rangle_t-\mathbf{c})\|_\infty)} \tag{22}$$

$$\leq x_i(0)e^{t\cdot(A\mathbf{c})_i+Const(A,\mathbf{x}_0)} \tag{23}$$

$$= Const(A,\mathbf{x}_0)e^{t\cdot(A\mathbf{c})_i}\ , \tag{24}$$

and analogously,

$$x_i(t) = x_i(0)e^{t\cdot(A\langle\mathbf{x}\rangle_t)_i} \geq Const(A,\mathbf{x}_0)e^{t\cdot(A\mathbf{c})_i}\ . \tag{25}$$

Therefore, condensate selection occurs exponentially fast at depletion rate $|(A\mathbf{c})_i|$. The dynamics of cases for non-generic antisymmetric matrices are discussed in Supplementary Note 2.

Linear programming algorithm. For the numeric computation of condensate vectors $\mathbf{c}$, a finite threshold $\delta > 0$ was introduced into the inequalities (5): $A\mathbf{c} \leq 0$ and $\mathbf{c} - A\mathbf{c} \geq \delta > 0$. Its value was set to $\delta = 1$ by rescaling of $\mathbf{c}$. Numerical solution of the inequalities was performed by using the IBM ILOG CPLEX Optimization Studio 12.5 and its interface to the C++ language. The software Mathematica 9.0 from Wolfram Research was also found to be applicable. Further information on the calibration of the linear programming algorithm and a simplified Mathematica algorithm are provided in Supplementary Note 3.

References

1. Krapivsky, P. L., Redner, S. & Leyvraz, F. Connectivity of growing random networks. *Phys. Rev. Lett.* **85**, 4629–4632 (2000).

2. Bianconi, G. & Barabási, A.-L. Bose-Einstein condensation in complex networks. *Phys. Rev. Lett.* **86**, 5632–5635 (2001).

3. Evans, M. R. & Hanney, T. Nonequilibrium statistical mechanics of the zero-range process and related models. *J. Phys. A: Math. Gen.* **38**, R195–R240 (2005).

4. Evans, M. R. Bose-Einstein condensation in disordered exclusion models and relation to traffic flow. *Europhys. Lett.* **36**, 13–18 (1996).

5. Krug, J. & Ferrari, P. A. Phase transitions in driven diffusive systems with random rates. *J. Phys. A: Math. Gen.* **29**, L465–L471 (1996).

6. Chowdhury, D., Santen, L. & Schadschneider, A. Statistical physics of vehicular traffic and some related systems. *Phys. Rep.* **329**, 199–329 (2000).

7. Kaupužs, J., Mahnke, R. & Harris, R. J. Zero-range model of traffic flow. *Phys. Rev. E* **72**, 056125 (2005).

8. Spitzer, F. Interaction of Markov processes. *Advances in Mathematics* **5**, 246–290 (1970).

9. Evans, M. R. & Waclaw, B. Condensation in stochastic mass transport models: beyond the zero-range process. *J. Phys. A: Math. Theor.* **47**, 095001 (2014).

10. Bose, S. N. Plancks Gesetz und Lichtquantenhypothese. *Z. Phys.* **26**, 178–181 (1924).

11. Einstein, A. Quantentheorie des einatomigen idealen Gases. *Sitzb. d. Preuss. Akad. d. Wiss.* 261–267 (1924).

12. Einstein, A. Quantentheorie des einatomigen idealen Gases. Zweite Abhandlung. *Sitzb. d. Preuss. Akad. d. Wiss.* 3–14 (1925).

13. Griffin, A., Snoke, D. & Stringari, G. *Bose Einstein Condensation* (Cambridge University Press, Cambridge, 1995).

14. Anglin, J. R. & Ketterle, W. Bose-Einstein condensation of atomic gases. *Nature* **416**, 211–218 (2002).

15. Penrose, O. & Onsager, L. Bose-Einstein condensation and liquid helium. *Phys. Rev.* **104**, 576–584 (1956).

16. Mueller, E. J., Ho, T.-L., Ueda, M. & Baym, G. Fragmentation of Bose-Einstein condensates. *Phys. Rev. A* **74**, 033612 (2006).

17. Gardiner, C. *Stochastic Methods: A Handbook for the Natural and Social Sciences* (Springer, Berlin, 2009).

18. Van Kampen, N. G. *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam, 2007).

19. Vorberg, D., Wustmann, W., Ketzmerick, R. & Eckardt, A. Generalized Bose-Einstein condensation into multiple states in driven-dissipative systems. *Phys. Rev. Lett.* **111**, 240405 (2013).

20. Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47–97 (2002).

21. Milo, R. *et al.* Network motifs: simple building blocks of complex networks. *Science* **298**, 824–827 (2002).

22. Blümel, R. *et al.* Dynamical localization in the microwave interaction of Rydberg atoms: the influence of noise. *Phys. Rev. A* **44**, 4521–4540 (1991).

23. Kohler, S., Dittrich, T. & Hänggi, P. Floquet-Markovian description of the parametrically driven, dissipative harmonic quantum oscillator. *Phys. Rev. E* **55**, 300–313 (1997).

24. Breuer, H.-P., Huber, W. & Petruccione, F. Quasistationary distributions of dissipative nonlinear quantum oscillators in strong periodic driving fields. *Phys. Rev. E* **61**, 4883–4889 (2000).

25. Grifoni, M. & Hänggi, P. Driven quantum tunneling. *Phys. Rep.* **304**, 229–354 (1998).

26. Breuer, H.-P. & Petruccione, F. *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, 2002).

27. Gardiner, C. W. & Zoller, P. Quantum kinetic theory: a quantum kinetic master equation for condensation of a weakly interacting Bose gas without a trapping potential. *Phys. Rev. A* **55**, 2902–2921 (1997).

28. Kagan, Y. & Svistunov, B. V. Evolution of correlation properties and appearance of broken symmetry in the process of Bose-Einstein condensation. *Phys. Rev. Lett.* **79**, 3331–3334 (1997).

29. Bijlsma, M. J., Zaremba, E. & Stoof, H. T. C. Condensate growth in trapped Bose gases. *Phys. Rev. A* **62**, 063609 (2000).

30. Gardiner, C. W., Lee, M. D., Ballagh, R. J., Davis, M. J. & Zoller, P. Quantum kinetic theory of condensate growth: comparison of experiment and theory. *Phys. Rev. Lett.* **81**, 5266–5269 (1998).

31. Walser, R., Williams, J., Cooper, J. & Holland, M. Quantum kinetic theory for a condensed bosonic gas. *Phys. Rev. A* **59**, 3878–3889 (1999).

32. Kocharovsky, V. V., Scully, M. O., Zhu, S.-Y. & Suhail Zubairy, M. Condensation of N bosons. II. Nonequilibrium analysis of an ideal Bose gas and the laser phase-transition analogy. *Phys. Rev. A* **61**, 023609 (2000).

33. Pauli, W. *Festschrift zum 60. Geburtstage A. Sommerfeld* (Hirzel, Leipzig, 1928).

34. Mandel, L. & Wolf, E. *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, UK, 1995).

35. Gardiner, C. W. & Zoller, P. *Quantum Noise* (Springer, Berlin Heidelberg, 2004).

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

37. Nowak, M. A. & Sigmund, K. Evolutionary dynamics of biological games. *Science* **303**, 793–799 (2004).

38. Sinervo, B. & Lively, C. M. The rock-paper-scissors game and the evolution of alternative male strategies. *Nature* **380**, 240–243 (1996).

39. Kerr, B., Riley, M., Feldman, M. & Bohannan, B. Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature* **418**, 171–174 (2002).

40. Reichenbach, T., Mobilia, M. & Frey, E. Mobility promotes and jeopardizes biodiversity in rock-paper-scissors games. *Nature* **448**, 1046–1049 (2007).

41. Weber, M. F., Poxleitner, G., Hebisch, E., Frey, E. & Opitz, M. Chemical warfare and survival strategies in bacterial range expansions. *J. R. Soc. Interface* **11**, 20140172 (2014).

42. Szolnoki, A. *et al.* Cyclic dominance in evolutionary games: a review. *J. R. Soc. Interface* **11**, 20140735 (2014).

43. Nowak, M. A., Sasaki, A., Taylor, C. & Fudenberg, D. Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646–650 (2004).

44. Szolnoki, A., Antonioni, A., Tomassini, M. & Perc, M. Binary birth-death dynamics and the expansion of cooperation by means of self-organized growth. *EPL* **105**, 48001 (2014).

45. McKane, A. J. & Newman, T. J. Predator-prey cycles from resonant amplification of demographic stochasticity. *Phys. Rev. Lett.* **94**, 218102 (2005).

46. Traulsen, A., Claussen, J. C. & Hauert, C. Coevolutionary dynamics: from finite to infinite populations. *Phys. Rev. Lett.* **95**, 238701 (2005).

47. Reichenbach, T., Mobilia, M. & Frey, E. Coexistence versus extinction in the stochastic cyclic Lotka-Volterra model. *Phys. Rev. E* **74**, 51907 (2006).

48. Melbinger, A., Cremer, J. & Frey, E. Evolutionary game theory in growing populations. *Phys. Rev. Lett.* **105**, 178101 (2010).

49. Biancalani, T., Dyson, L. & McKane, A. J. Noise-induced bistable states and their mean switching time in foraging colonies. *Phys. Rev. Lett.* **112**, 038101 (2014).

50. Rulands, S., Jahn, D. & Frey, E. Specialization and bet hedging in heterogeneous populations. *Phys. Rev. Lett.* **113**, 108102 (2014).

51. Volterra, V. *Leçons sur la Théorie Mathématique de la Lutte pour la Vie* (Gauthier-Villars, Paris, 1931).

52. Goel, N. S., Maitra, S. C. & Montroll, E. W. On the Volterra and other nonlinear models of interacting populations. *Rev. Mod. Phys.* **43**, 231–276 (1971).

53. May, R. M. *Stability and Complexity in Model Ecosystems* (Princeton University Press, Princeton, NJ, 1973).

54. Zakharov, V., Musher, S. & Rubenchik, A. Nonlinear stage of parametric wave excitation in a plasma. *JETP Lett.* **19**, 151–152 (1974).

55. Manakov, S. Complete integrability and stochastization of discrete dynamical systems. *Sov. Phys.-JETP* **40**, 269–274 (1975).

56. Itoh, Y. Boltzmann equation on some algebraic structure concerning struggle for existence. *Proc. Japan Acad.* **47**, 854–858 (1971).

57. Di Cera, E., Phillipson, P. E. & Wyman, J. Chemical oscillations in closed macromolecular systems. *Proc. Natl. Acad. Sci. USA* **85**, 5923–5926 (1988).

58. Di Cera, E., Phillipson, P. E. & Wyman, J. Limit-cycle oscillations and chaos in reaction networks subject to conservation of mass. *Proc. Natl. Acad. Sci. USA* **86**, 142–146 (1989).

59. Akin, E. & Losert, V. Evolutionary dynamics of zero-sum games. *J. Math. Biol.* **20**, 231–258 (1984).

60. Chawanya, T. & Tokita, K. Large-dimensional replicator equations with antisymmetric random interactions. *J. Phys. Soc. Jpn.* **71**, 429–431 (2002).

61. Knebel, J., Krüger, T., Weber, M. F. & Frey, E. Coexistence and survival in conservative Lotka-Volterra networks. *Phys. Rev. Lett.* **110**, 168106 (2013).

62. Kuhn, H. & Tucker, A. *Linear Inequalities and Related Systems* (Princeton University Press, Princeton, NJ, 1956).

63. Prigogine, I. Time, structure, and fluctuations. *Science* **201**, 777–785 (1978).

64. Gardner, M. R. & Ashby, W. R. Connectance of large dynamic (cybernetic) systems: critical values for stability. *Nature* **228**, 784 (1970).

65. Durney, C. H., Case, S. O., Pleimling, M. & Zia, R. K. P. Saddles, arrows, and spirals: deterministic trajectories in cyclic competition of four species. *Phys. Rev. E* **83**, 051108 (2011).

66. Allesina, S. & Levine, J. M. A competitive network theory of species diversity. *Proc. Natl. Acad. Sci. USA* **108**, 5638–5642 (2011).

67. Eisert, J., Friesdorf, M. & Gogolin, C. Quantum many-body systems out of equilibrium. *Nature Phys.* **11**, 124–130 (2015).

68. Zia, R. K. P. & Schmittmann, B. Probability currents as principal characteristics in the statistical mechanics of non-equilibrium steady states. *J. Stat. Mech.* **2007**, P07012 (2007).

69. Kriecherbauer, T. & Krug, J. A pedestrian's view on interacting particle systems, KPZ universality and random matrices. *J. Phys. A: Math. Theor.* **43**, 403001 (2010).

70. Cullis, C. E. *Matrices and Determinoids*, vol. I and II (Cambridge University Press, Cambridge, 1913).

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Acknowledgements We are grateful for fruitful discussions with Peter Zoller, Ulrich Schollwöck, Immanuel Bloch, Wilhelm Zwerger, André Eckardt, Daniel Vorberg, Alexander Schnell, Marianne Bauer, Brendan Osberg, Jacob Halatek, Matthias Bauer, and Sebastian Koch. This work was supported by the Deutsche Forschungsgemeinschaftas as project A7 of the SFB TR 12 "Symmetry and Universality in Mesoscopic Systems", and by the German Excellence Initiative via the program "Nanosystems Initiative Munich" (NIM). J.K. acknowledges funding from the Studienstiftung des Deutschen Volkes.

Author Contributions J.K., M.F.W., T.K., E.F. designed, discussed, and planned the study. T.K., J.K., M.F.W. developed the analytical results. M.F.W., T.K., J.K. developed the numerical algorithms and generated the data. J.K., M.F.W., T.K., E.F. interpreted the results and wrote the manuscript.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

The authors declare no competing financial interests.

Correspondence and requests for materials should be addressed to E.F. (email: frey@lmu.de).

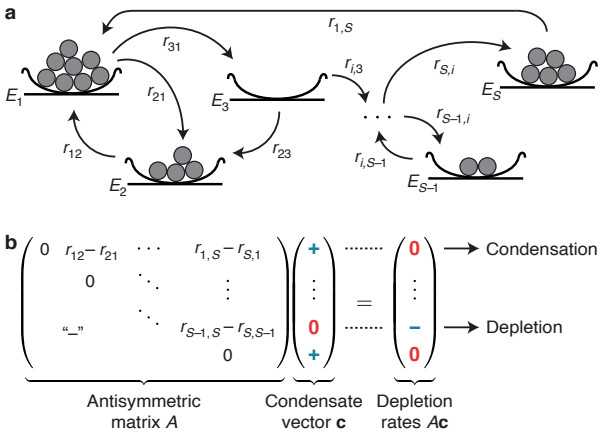


Figure 1 Condensation into multiple states due to particle transitions between states and mathematics of condensate selection. (**a**) With respect to condensation in an incoherently driven-dissipative quantum system, each bowl represents a state E_i that is occupied by N_i non-interacting bosons (filled circles). If indicated by an arrow, bosons may undergo transitions from state E_j to state E_i at a rate $\Gamma_{i\leftarrow j} = r_{ij}(N_i + 1)N_j$, with rate constant r_{ij}. In the language of evolutionary game theory, the figure depicts the interaction of N_i agents (filled circles) playing strategies E_i (bowls). An agent playing strategy E_j adopts strategy E_i at a rate $\Gamma_{i\leftarrow j} = r_{ij}N_iN_j$. The above rate of bosonic condensate selection is recovered if agents may also spontaneously mutate from E_j to E_i at a rate r_{ij}. (**b**) A condensate vector $\mathbf{c}$ for an antisymmetric matrix A has two properties: its entries are positive for indices for which $A\mathbf{c}$ is zero, and they are zero for indices for which $A\mathbf{c}$ is negative ("-" signifies the antisymmetry of matrix A). Temporal evolution of the relative entropy of the condensate vector to the state concentrations under the ALVE (3) relates positive entries of the condensate vector to condensates, and its zero entries to depleted states. Generically, positive

entries of $\mathbf{c}$ represent the asymptotic temporal average of oscillating condensate concentrations according to the ALVE (3), and negative entries of $A\mathbf{c}$ represent depletion rates.

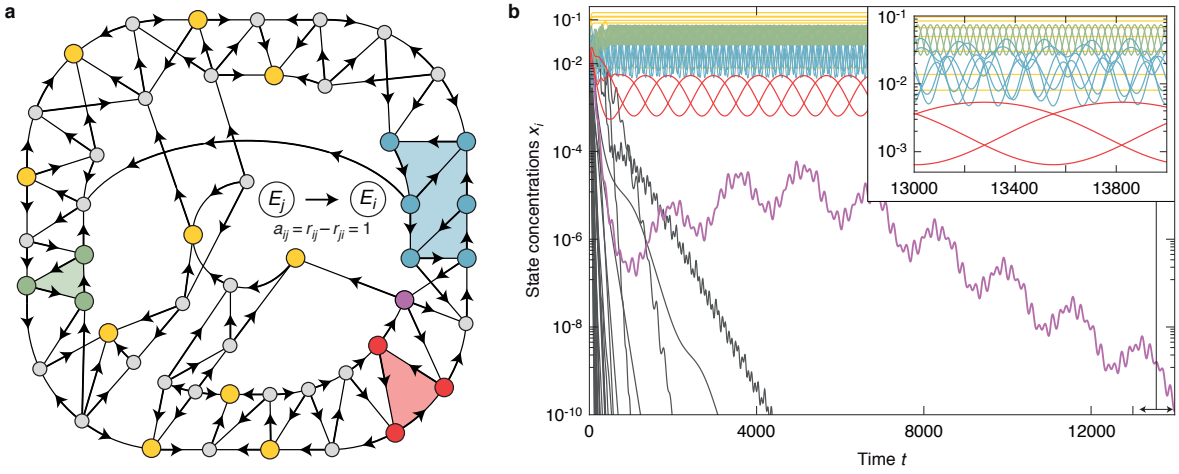


Figure 2 Fragmentation of an exemplary system into multiple condensates with oscillating state concentrations. (**a**) Randomly sampled network of 50 states. Disks represent states. An arrow from state E_j to state E_i represents an effective rate constant $a_{ij} = r_{ij} - r_{ji} = 1$ (a missing arrow indicates a forbidden transition with $a_{ij} = 0$). Computation of a condensate vector $\mathbf{c}$ predicted relaxation into ten isolated condensates (yellow), one interacting subsystem with six condensates (blue), and two rock-paper-scissors (RPS) cycles (red and green). All other states become depleted. The complete network also comprises RPS cycles of which some states become depleted. Knowledge of the network topology alone is thus insufficient to determine condensates. (**b**) Temporal evolution of state concentrations x_i (logarithmic scale). Colours in accordance with (**a**). Numerical integration of the ALVE (3) confirmed the selection of states based on the condensate vector $\mathbf{c}$. Subsystems with six (blue) and three (red and green) condensates exhibit oscillations of concentrations with non-vanishing particle flow. Depletion of states occurs exponentially fast. Identifying condensates from a condensate vector $\mathbf{c}$ is more reliable than through numerical inte-

gration: The concentration of the state associated to the purple disk in (**a**) decays exponentially to a concentration of $1.5 \cdot 10^{-7}$ before recovering transiently. Numerical integration cannot rule out permanent recovery at later times. Supplementary Fig. 2 demonstrates such a case.

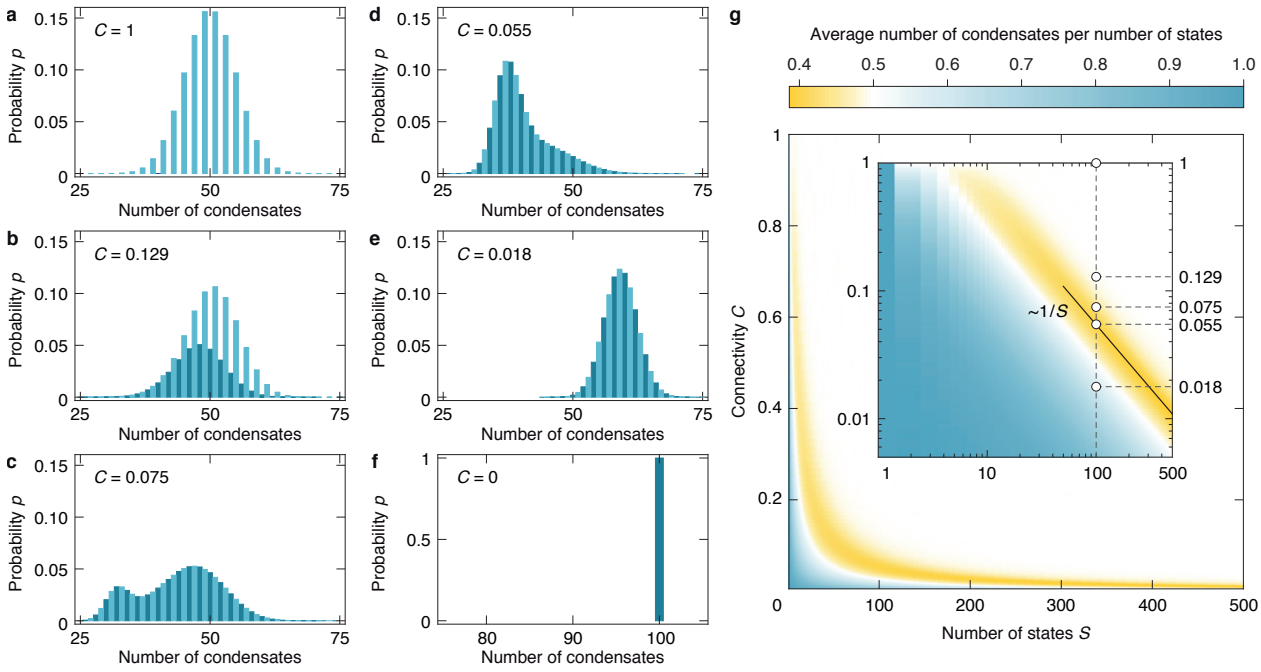


Figure 3 Dependence of the number of condensates on the connectivity of states in random networks. (**a-f**) Measured probability p of finding a particular number of condensates for a system with $S = 100$ states and with connectivity C ($5 \cdot 10^6$ systems analysed per histogram). The connectivity specifies the percentage of states between which transitions of particles occur with a non-zero effective rate constant $a_{ij} = r_{ij} - r_{ji}$. Effective rate constants a_{ij} were sampled from a Gaussian distribution (zero mean, unit variance). (**a**) At full connectivity, the distribution is pseudo-binomial with only odd numbers of condensates ($C = 1$; light blue bars)[19,60,66]. (**b**) As the connectivity is reduced, even numbers of condensates become possible when systems decouple into even numbers of subsystems ($C = 0.129$; dark blue bars). (**c**) The distribution exhibits bimodality ($C = 0.075$) and (**d**) approaches a minimal average number of 40.2 condensates ($C = 0.055$). (**e**)

This average subsequently increases ($C = 0.018$) because isolated states are trivially selected as condensates ($C = 0$) as shown in (**f**). (**g**) Average number of condensates per number of states (colour-coded) plotted against the number of states S and the connectivity C (log-log graph in inset; $\geq 10^4$ systems per data point, see Supplementary Fig. 4 for the reliability of the linear programming algorithm). White circles correspond to distributions shown in (**a-e**). The minimal relative number of condensates conforms to the power law $C \sim 1/S^{\gamma}$ with $\gamma = 0.998 \pm 0.008$ (s.e.m.) and can be related to the criticality of random networks[20].

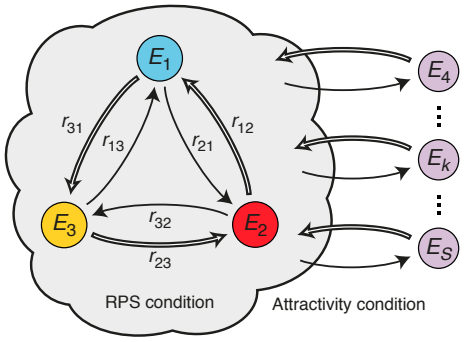
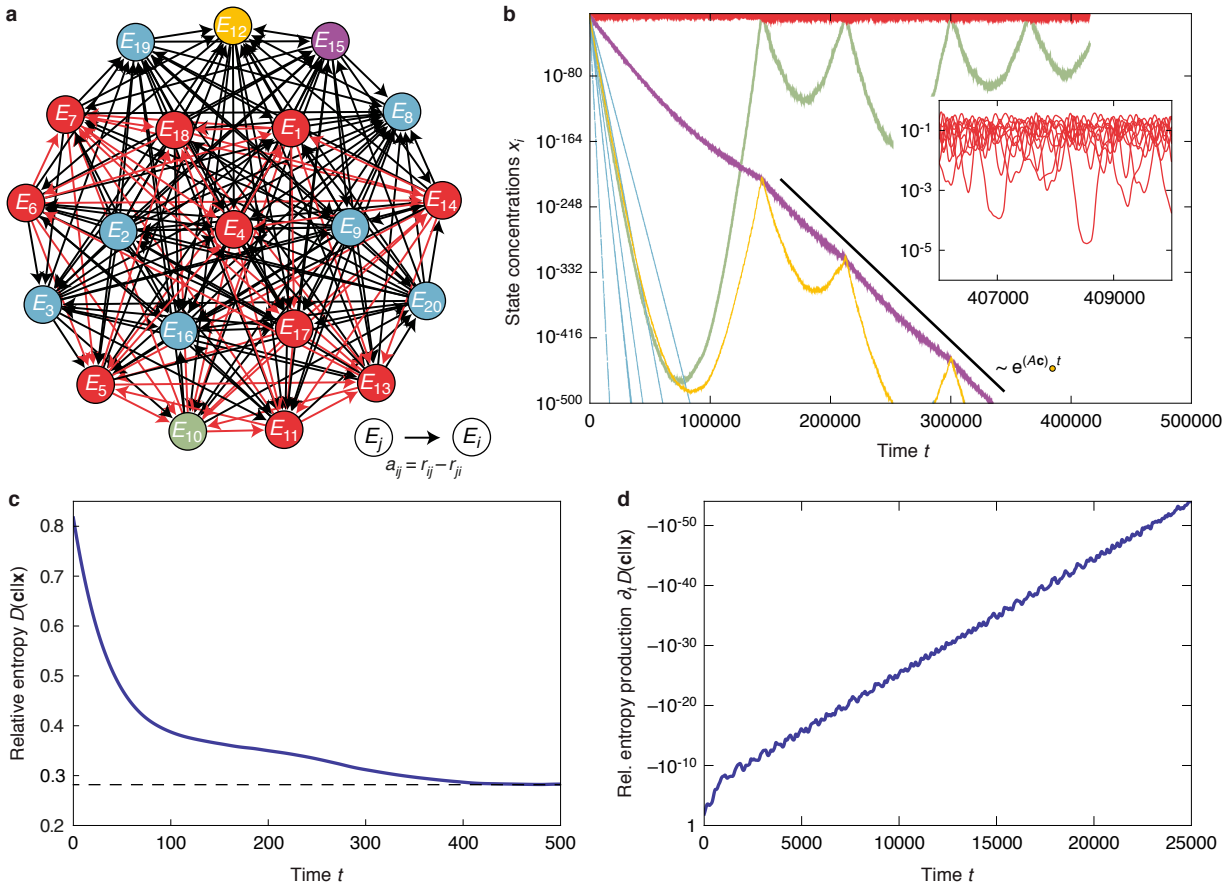


Figure 4 Conditions for the emergence of a rock-paper-scissors cycle of condensates. Three particular states E_1, E_2, and E_3 (blue, red, and yellow disks) of a network condense into a rock-paper-scissors (RPS) cycle if, and only if, two conditions are fulfilled: First, the "RPS condition" requires that the rate constants r_{ij} between the three states form a RPS network: $r_{i-1,i+1} > r_{i+1,i-1}$ (indices are counted modulo 3, for example, $r_{42} = r_{12}$ (framed arrows denote rate constants that are larger than rate constants for the respective reverse direction). Differences between these rate constants define the entries $c_i = r_{i-1,i+1} - r_{i+1,i-1}$ of an admissible condensate vector $\mathbf{c}$. Second, the "attractivity condition" requires that the weighted sum of rates from any exterior state E_k (purple disks) into the RPS cycle, $\sum_{j=1}^{3} c_j r_{jk}$ (framed arrows), is larger than the weighted sum of outbound rates, $\sum_{j=1}^{3} c_j r_{kj}$ (black arrows). In other words, the inflow of particles into the RPS cycle from any exterior state needs to be greater than the outflow to that state.

Field of research	Entity	State	Process	Dynamics
Quantum physics[19]	Boson	Quantum state	Incoherent transition	Condensation/depletion
Evolutionary game theory[59, 60]	Agent	Strategy	Game (+mutation)	Win/loss
Population dynamics[51-53, 47]	Individual	Species	Competition (+mutation)	Survival/extinction
Chemical kinetics[56-58]	Molecule	Chemical species	Reaction (+conversion)	Production/consumption
Plasma physics[54, 55]	Plasmon	Jet	Scattering	Increase/decrease

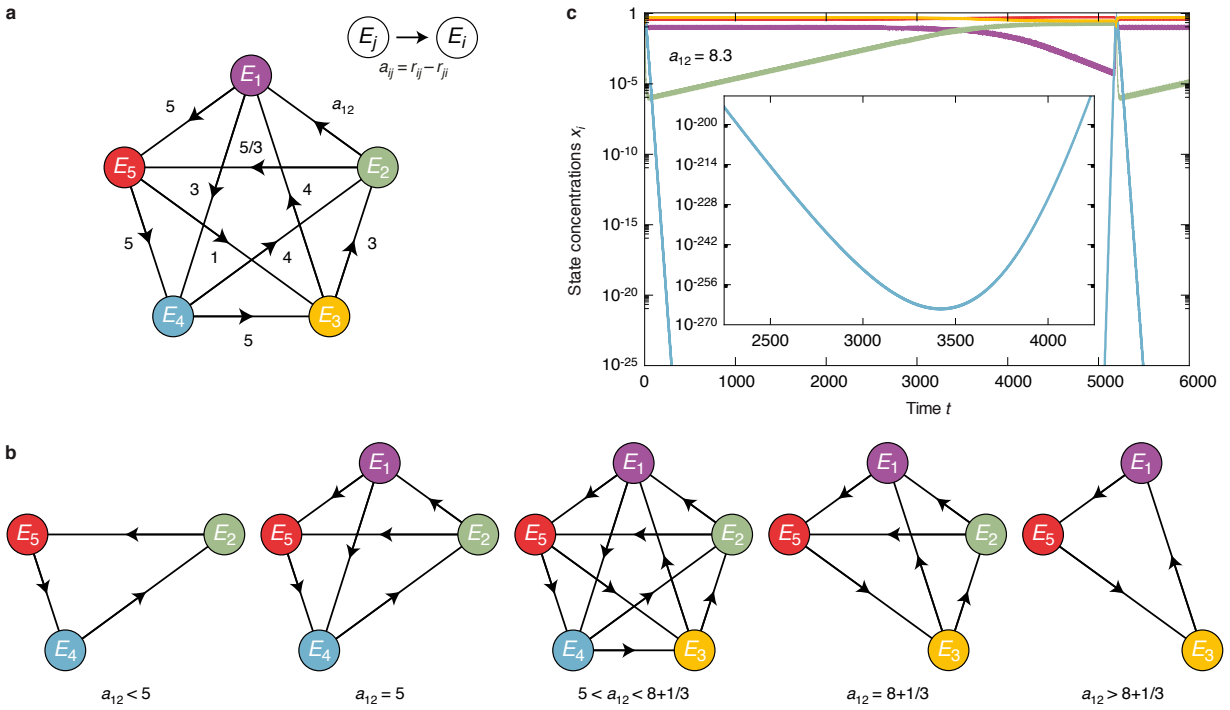
Table 1. Condensation processes described by the ALVE in different fields of research. The ALVE (3) governs condensation processes in diverse fields of research. For example, for incoherently driven-dissipative bosonic systems, the ALVE describes condensation and depletion of states by incoherent transitions of non-interacting bosons. In EGT, the ALVE occurs in the context of winning and losing strategies played by agents.

SUPPLEMENTARY FIGURE 1



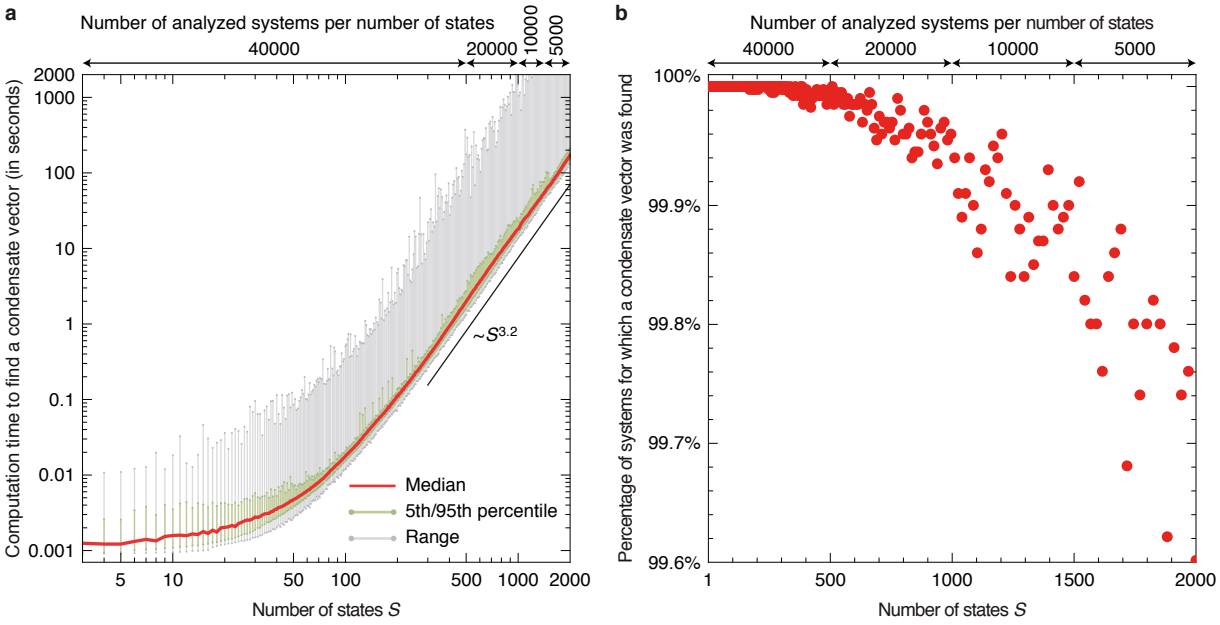
Supplementary Figure 1. Advantage of the algebraic algorithm over numerical integration of the antisymmetric Lotka-Volterra equation (ALVE). (**a**) A fully connected network with 20 states (colored disks). Effective transition rates a_{ij} (arrows) are listed in Supplementary Note 3 and were sampled from a Gaussian distribution (zero mean, unit variance). Computation of the unique normalized condensate vector $\mathbf{c}$ identifies 11 states as condensates (red and green disks) and depletion of 9 states (blue, yellow, and purple disks). (**b**) Trajectories show the temporal evolution of state concentrations x_i (colors in accordance with (**a**)). Numerical integration of the ALVE is highly unstable. Only the routine NDSolve with method "StiffnessSwitching" of Mathematica from Wolfram Research was able to track the concentrations for a sufficiently long time. Routines offered by the GNU Scientific Library failed (Dormand-Prince and Runge-Kutta-Fehlberg). The concentration corresponding to the green state transiently decreases to a value of $4{\cdot}10^{-473}$ before recovering. Its asymptotic temporal average is given by the corresponding entry of the condensate vector $\mathbf{c}$ as $1.6 \cdot 10^{-4}$. This average could, however, not be verified by integration due to numerical failure at $t = 414960$. Negative entries of $A\mathbf{c}$ determine the rates of exponentially fast depletion as illustrated by the black line for the yellow state (note the logarithmic scaling). (**c**) Temporal evolution of the relative entropy $D(\mathbf{c}||\mathbf{x})(t)$ (blue line). The relative entropy decreases towards a non-zero asymptotic value. (**d**) The production of relative entropy $\partial_t D(\mathbf{c}||\mathbf{x})(t)$ (blue line) is, therefore, negative and vanishes for large times.

SUPPLEMENTARY FIGURE 2



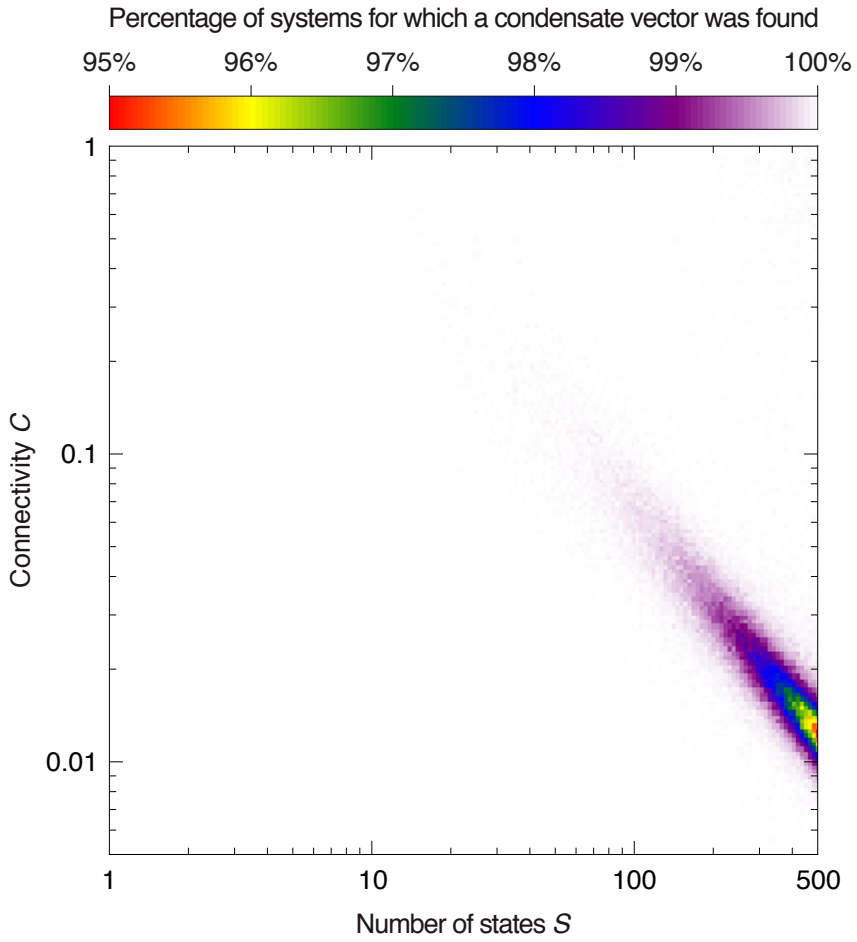
Supplementary Figure 2. Identification of condensates for a system with five states. (**a**) Colored disks represent states E_i. An arrow from E_j to E_i denotes an effective rate constant $a_{ij} = r_{ij} - r_{ji}$. (**b**) Computation of condensate vectors for different values of a_{12} yields depletion of states E_1 and E_3 for $a_{12} < 5$, depletion of state E_3 for $a_{12} = 5$, condensation of all states for $5 < a_{12} < 8 + 1/3$, depletion of state E_4 for $a_{12} = 8 + 1/3$, and depletion of states E_2 and E_4 for $a_{12} > 8 + 1/3$. These results can be verified by using the Mathematica code supplied in Supplementary Note 3. (**c**) Numerical integration of the ALVE confirms the selection of condensates, but becomes error-prone in the vicinity of values of a_{12} at which the set of condensates changes. The trajectories were obtained for $a_{12} = 8.3$, which is slightly smaller than the value at which E_4 becomes depleted. Identification of condensates from trajectories requires the introduction of a threshold for concentrations below which states are considered as depleted. If such a threshold is set to values larger than $\sim 2.5 \cdot 10^{-265}$, state E_4 is erroneously considered to be depleted, despite its periodic recovery.

SUPPLEMENTARY FIGURE 3



Supplementary Figure 3. Run-time analysis of the linear programming algorithm for fully connected random networks of states. The computation time to find a condensate vector was measured for systems with up to $S = 2000$ states. The systems were constructed by sampling effective rate constants $a_{ij} = r_{ij} - r_{ji}$ from a Gaussian distribution (zero mean, unit variance). The numbers of systems (ensemble size) that were analyzed for different numbers of states S are indicated at the top. (**a**) The median of computation times is shown in red, the regime between the 5th and 95th percentiles in green, and the range of computation times in grey (log-log graph). Computation time increases at most polynomially as indicated by the black line (linear fit of the median with exponent 3.221 ± 0.008 (s.e.m.)). (**b**) Percentage of networks for which a condensate vector was found (red dots). The lowest percentage was 99.6%. All computations were performed on machines with 10 Intel Xeon E5-2670v2 cores (2.50 GHz) and 128 GB RAM. Parallelization was not used in CPLEX.

SUPPLEMENTARY FIGURE 4



Supplementary Figure 4. Reliability of the linear programming algorithm for large random networks. Color-coded representation of the percentage of random networks with S states and connectivity C for which the linear programming algorithm found a suitable condensate vector ($\geq 10^4$ analyzed systems per data point; the figure accompanies Fig. 3b). The percentage was close to 100% for most parameters but decreased in the vicinity of the power law shown in Fig. 3b (lowest percentage: 94.9% for networks with $S = 500$ states and connectivity $C = 0.013$).

SUPPLEMENTARY NOTE 1: APPROXIMATION OF THE CLASSICAL MASTER EQUATION BY STOCHASTIC DIFFERENTIAL EQUATIONS AND OCCURRENCE OF THE ANTISYMMETRIC LOTKA-VOLTERRA EQUATION

In the following, we supplement the discussion of the classical master equation that governs condensation of non-interacting bosons in an incoherently driven-dissipative system. When the total number of particles in the system is large ($N \gg 1$), the classical master equation can be approximated by a Fokker-Planck equation. This Fokker-Planck equation is rewritten as an Itō stochastic differential equation, which is equivalent to a Langevin equation. From the analysis of the stochastic differential equation, we find that the leading order dynamics of the condensation process is governed by the antisymmetric Lotka-Volterra equation (equation (3) in the main text).

Classical master equation for the coupled birth-death processes with conservation of total particle number

We consider a system of S non-degenerate states $E_i, i = 1, \ldots, S$, each of which is occupied by N_i indistinguishable particles (Fig. 1a). The configuration of the system at time t is fully characterized by the vector of occupation numbers $\mathbf{N} = (N_1, N_2, \ldots, N_S) \in \mathbb{Z}^S$ with $N_i \geq 0$. It changes only due to coupled creation (birth) and annihilation (death) processes between two connected states E_i and E_j:

$$(N_1, \ldots, N_i - 1, \ldots, N_j + 1, \ldots, N_S) \leftrightarrows (N_1, \ldots, N_i, \ldots, N_j, \ldots, N_S) \leftrightarrows (N_1, \ldots, N_i + 1, \ldots, N_j - 1, \ldots, N_S)\,. \tag{1}$$

We introduce the following short-hand notation for these processes:

$$\mathbf{N} - \mathbf{e}_i + \mathbf{e}_j \quad \leftrightarrows \quad \mathbf{N} \quad \leftrightarrows \quad \mathbf{N} + \mathbf{e}_i - \mathbf{e}_j\,. \tag{2}$$

Here, the vector $\mathbf{e}_i \in \mathbb{Z}^S$ denotes the unit vector in direction i (equal to one at index i, otherwise zero). The above creation and annihilation processes conserve the total number of particles $N = \sum_i N_i$. We are interested in the probability $P(\mathbf{N}, t)$ of finding the system in configuration $\mathbf{N}$ at time t, given that it was initially in configuration $\mathbf{N}_0$ at time t_0. The temporal evolution of the probability distribution $P(\mathbf{N}, t)$ is governed by the classical master equation (equation (1) in the main text):

$$\partial_t P(\mathbf{N}, t) = \sum_{\substack{i,j=1 \\ j \neq i}}^{S} \Big(\Gamma_{i \leftarrow j}(N_i - 1, N_j + 1) P(\mathbf{N} - \mathbf{e}_i + \mathbf{e}_j, t) - \Gamma_{i \leftarrow j}(N_i, N_j) P(\mathbf{N}, t) \Big)\,. \tag{3}$$

In our work, we consider the following transition rate from configuration $\mathbf{N}$ to configuration $\mathbf{N} + \mathbf{e}_i - \mathbf{e}_j$:

$$\Gamma_{i \leftarrow j}(N_i, N_j) = r_{ij}(N_i + s_{ij}) N_j, \text{ with } s_{ij} \geq 0 \text{ and } r_{ij} \geq 0\,. \tag{4}$$

This transition rate encompasses the two model classes of condensate selection in bosonic systems and in evolutionary game theory (EGT) as described in the main text. In the context of bosonic condensation, the parameters s_{ij} are equal to 1 for all i and j, whereas it may assume any non-negative value in the context of EGT. In EGT, s_{ij} contributes to the mutation or switching rate [1].

For the study of incoherently driven-dissipative systems of non-interacting bosons, the above description in terms of a classical master equation requires that the system under consideration is weakly coupled to a reservoir [2–5]. The reservoir has to be such that correlations in it decay rapidly. In particular, the Born-Markov and the rotating wave approximation are assumed in the derivation of the classical master equation. These assumptions imply that off-diagonal entries in the reduced density operator of the system decay fast enough such that coherence in the quantum system is negligible. In addition, the initial states of the system and the reservoir should not be correlated after their preparation.

These conditions are, for example, fulfilled in systems of non-interacting bosons that are both periodically driven in time (Floquet systems) and weakly coupled to a thermal bath (see Vorberg et al. [2] for a detailed discussion).

Derivation of the Fokker-Planck equation

In the following, we approximate the classical master equation (3) in the limit of a large number of particles ($N \gg 1$) [6]. For that purpose, we introduce state concentrations $\mathbf{x} = (x_1, \ldots, x_S)$ with $x_i = N_i/N$. The concentrations are intensive variables and elements of the $(S-1)$-dimensional open simplex $\Delta_{S-1} = \{\mathbf{x} \in \mathbb{R}^S \mid x_i > 0 \text{ for all } i,\ \sum_{i=1}^S x_i = 1\}$. In the limit $N \to \infty$, they become continuous variables. We denote their corresponding probability distribution by $p(\mathbf{x}, t)$. Furthermore, we rescale time by $t \to t/N$. It can be straightforwardly seen that the following formulation of the classical master equation for $p(\mathbf{x}, t)$ is equivalent to the form given in equation (3):

$$\partial_t p(\mathbf{x},t) = \sum_{\substack{i,j=1\\ j\neq i}}^{S} \int_{\mathbb{R}} d\Delta x_i \int_{\mathbb{R}} d\Delta x_j \sum_{k\in\{0,1\}} N^{1-k} \Big(\gamma^{(k)}(\Delta x_i, \Delta x_j; x_i - \Delta x_i, x_j - \Delta x_j)\, p(\mathbf{x} - \Delta x_i \mathbf{e}_i - \Delta x_j \mathbf{e}_j, t) - \gamma^{(k)}(\Delta x_i, \Delta x_j; x_i, x_j)\, p(\mathbf{x}, t) \Big) , \tag{5}$$

$$\text{with } \gamma^{(0)}(\Delta x_i, \Delta x_j; x_i, x_j) = r_{ij} x_i x_j\, \delta(\Delta x_i - 1/N)\delta(\Delta x_j + 1/N) , \tag{6}$$

$$\text{and } \gamma^{(1)}(\Delta x_i, \Delta x_j; x_i, x_j) = r_{ij}\, s_{ij}\, x_j\, \delta(\Delta x_i - 1/N)\delta(\Delta x_j + 1/N) . \tag{7}$$

This classical master equation can be approximated by performing a Kramers-Moyal expansion. Truncation of the expansion at second order, leads to the Fokker-Planck equation:

$$\partial_t p(\mathbf{x},t) = -\sum_{i=1}^{S} \partial_i \big(\alpha_i(\mathbf{x}) p(\mathbf{x},t)\big) + \frac{1}{2N} \sum_{i,j=1}^{S} \partial_i \partial_j \big(\beta_{ij}(\mathbf{x}) p(\mathbf{x},t)\big) . \tag{8}$$

Here, $\alpha(\mathbf{x})$ is the drift vector and $\beta(\mathbf{x})$ the diffusion matrix.

The drift vector is given by:

$$\alpha_i(\mathbf{x}) = \sum_{\substack{j=1\\ j\neq i}}^{S} \int_{\mathbb{R}} d\Delta x_i \int_{\mathbb{R}} d\Delta x_j \sum_{k\in\{0,1\}} \frac{1}{N^k} \Big(\gamma^{(k)}(\Delta x_i, \Delta x_j; x_i, x_j) + \gamma^{(k)}(\Delta x_j, \Delta x_i; x_j, x_i) \Big) N \Delta x_i , \tag{9}$$

$$= x_i \sum_{j=1}^{S} (r_{ij} - r_{ji}) x_j + \frac{1}{N} \sum_{j=1}^{S} (r_{ij} s_{ij} x_j - r_{ji} s_{ji} x_i) =: \alpha_{i,0}(\mathbf{x}) + \frac{1}{N} \alpha_{i,1}(\mathbf{x}) . \tag{10}$$

At leading order in $1/N$ (that is, at order $\mathcal{O}(1)$), the drift vector is determined by the antisymmetric part $A = R - R^T$ of the rate constant matrix $R = \{r_{ij}\}_{i,j}$. In other words, the matrix A is antisymmetric with entries $a_{ij} = r_{ij} - r_{ji}$.

For the diffusion matrix, we have to distinguish between its diagonal entries,

$$\beta_{ii}(\mathbf{x}) = \sum_{\substack{j=1\\j\neq i}}^{S} \int_{\mathbb{R}} d\Delta x_i \int_{\mathbb{R}} d\Delta x_j \sum_{k\in\{0,1\}} \frac{1}{N^k}\Big(\gamma^{(k)}(\Delta x_i, \Delta x_j; x_i, x_j) + \gamma^{(k)}(\Delta x_j, \Delta x_i; x_j, x_i)\Big) N^2 \Delta x_i^2 \;, \tag{11}$$

$$= \sum_{\substack{j=1\\j\neq i}}^{S} (r_{ij} + r_{ji}) x_i x_j + \frac{1}{N} \sum_{\substack{j=1\\j\neq i}}^{S} (r_{ij} s_{ij} x_j + r_{ji} s_{ji} x_i) =: \beta_{ii,0}(\mathbf{x}) + \frac{1}{N}\beta_{ii,1}(\mathbf{x}) \;, \tag{12}$$

and its off-diagonal entries ($i \neq j$):

$$\beta_{ij}(\mathbf{x}) = \int_{\mathbb{R}} d\Delta x_i \int_{\mathbb{R}} d\Delta x_j \sum_{k\in\{0,1\}} \frac{1}{N^k}\Big(\gamma^{(k)}(\Delta x_i, \Delta x_j; x_i, x_j) + \gamma^{(k)}(\Delta x_j, \Delta x_i; x_j, x_i)\Big) N^2 \Delta x_i \Delta x_j \;, \tag{13}$$

$$= -(r_{ij} + r_{ji}) x_i x_j - \frac{1}{N}\big(r_{ij} s_{ij} x_j + r_{ji} s_{ji} x_i\big) =: \beta_{ij,0}(\mathbf{x}) + \frac{1}{N}\beta_{ij,1}(\mathbf{x}) \;. \tag{14}$$

Both diagonal and off-diagonal entries of the diffusion matrix are determined by the symmetric part of the rate constant matrix R at leading order in $1/N$.

Derivation of the stochastic differential equations

The Fokker-Planck equation (8) can be transformed into a system of Itō stochastic differential equations (SDEs) [6]:

$$\mathrm{d}x_i = \alpha_i(\mathbf{x})\,\mathrm{d}t + \frac{1}{\sqrt{N}} \sum_{j=1}^{S} \zeta_{ij}(\mathbf{x})\,\mathrm{d}W_j \;. \tag{15}$$

Here, $\mathrm{d}W_j$ represents a Wiener increment of zero mean and unit variance. The matrix $\zeta(\mathbf{x})$ is a square root of the diffusion matrix $\beta(\mathbf{x})$ in the sense that $\zeta\zeta^T = \beta$ (the diffusion matrix β is positive semi-definite). Although ζ is not unique, its choice does not change the stochastic nature of the process (an orthogonal transformation $\zeta \to \zeta\mathcal{T}$ with $\mathcal{T}\mathcal{T}^T = \mathbb{I}_S$ does not change the corresponding Fokker-Planck equation). The decomposition $\beta_{ij} = \beta_{ij,0} + \frac{1}{N}\beta_{ij,1}$ in equations (11) and (13) implies that ζ can be written as $\zeta_{ij} = \zeta_{ij,0} + \mathcal{O}\big(\frac{1}{N}\big)$.

The Itō SDEs (15) can also be written in Langevin form, which are often used in the physics literature [6, 7]:

$$\frac{\mathrm{d}}{\mathrm{d}t}x_i = \alpha_i(\mathbf{x}) \; + \frac{1}{\sqrt{N}} \sum_{j=1}^{S} \zeta_{ij}(\mathbf{x})\,\eta_j \;. \tag{16}$$

Here, η_j represents uncorrelated Gaussian white noise.

The antisymmetric Lotka-Volterra equation

We identify a leading (fast) and a subleading (slow) timescale of the SDE (15). On the leading timescale ($t \sim \mathcal{O}(1)$), only the drift term $\alpha_{i,0}$ is relevant, whereas on the subleading timescale ($t \sim \mathcal{O}(N)$), the terms $\alpha_{i,1}$ and $\zeta_{ij,0}$ compete. The latter terms cause only slow changes on the leading $\mathcal{O}(1)$-timescale. In other words, the dynamics on the subleading timescale cause only slow changes of the $\mathcal{O}(1)$-trajectory.

More specifically, we find that at order $t \sim \mathcal{O}(1)$, only $\alpha_{i,0}$ determines the change in concentrations such that:

$$\frac{\mathrm{d}}{\mathrm{d}t}x_i = x_i \sum_j (r_{ij} - r_{ji})x_j = x_i(A\mathbf{x})_i \, . \tag{17}$$

As stated in the main text, we refer to this equation as the antisymmetric Lotka-Volterra equation (ALVE). The initial concentrations are assumed to lie in the open simplex Δ_{S-1}, that is $\mathbf{x}(t=0) =: \mathbf{x}_0 \in \Delta_{S-1}$. We note that the dynamics defined by equation (17) cannot leave the simplex, that is $\mathbf{x}(t) \in \Delta_{S-1}$ for all times [8].

We note that the van Kampen system size expansion [9] of the master equation (3) yields the same deterministic equation (17) as our derivation via Fokker-Planck and Langevin equation at the leading order timescale.

Stability in the linear approximation around fixed points of the ALVE

We discuss the fixed points $\mathbf{x}^* \in \overline{\Delta}_{S-1}$ ($x_i^* \geq 0$ and $\sum_i x_1^* = 1$) of the ALVE (17), that is the points for which the dynamics is stationary ($\partial_t x_i^* = x_i^*(A\mathbf{x}^*)_i = 0$). In the following, we show that a linear stability analysis of these fixed points does not yield insight into the global dynamics of the ALVE.

First, every condensate vector $\mathbf{c}$ (normalized such that $\sum_i c_i = 1$) of the antisymmetric matrix A yields a fixed point of the ALVE. This can be seen from the properties of a condensate vector $\mathbf{c}$ (see Methods section of the main text) [10]:

$$c_i > 0 \;\text{ and } (A\mathbf{c})_i = 0 \text{ for } i \in I, \text{ and} \tag{18}$$

$$c_i = 0 \;\text{ and } (A\mathbf{c})_i < 0 \text{ for } i \in \bar{I} \, . \tag{19}$$

Notably, the index set I is unique although more than one condensate vector may exist. Furthermore, there exist fixed points $\mathbf{x}^* \in \overline{\Delta}_{S-1}$ and a different index set $J \neq I$ for which $x_j^* > 0$ and $(A\mathbf{x}^*)_j = 0$ for $j \in J$, and $x_j^* = 0$ for $j \in \bar{J}$ but $(A\mathbf{x}^*)_j < 0$ does not hold for all $j \in \bar{J}$ (in other words, condition (18) is fulfilled, but condition (19) is not).

We first study the stability in the linear approximation around the fixed points that are given by condensate vectors. Upon introducing the distance $\Delta\mathbf{x}$ of a normalized condensate vector $\mathbf{x}^* = \mathbf{c}$ from the concentrations $\mathbf{x}$ as a new variable, $\Delta\mathbf{x} \coloneqq \mathbf{x} - \mathbf{c}$, one obtains from the ALVE (17) the temporal behavior of that translated variable as follows:

$$\frac{\mathrm{d}}{\mathrm{d}t}\Delta x_i = \Delta x_i (A\mathbf{c})_i + c_i (A\Delta\mathbf{x})_i + \Delta x_i (A\Delta\mathbf{x})_i \;\text{ , that is,} \tag{20}$$

$$\text{for } i \in I: \;\; \frac{\mathrm{d}}{\mathrm{d}t}\Delta x_i = \sum_{j=1}^{S} c_i a_{ij} \Delta x_j + R(\Delta\mathbf{x}) \, , \tag{21}$$

$$\text{for } i \in \bar{I}: \;\; \frac{\mathrm{d}}{\mathrm{d}t}\Delta x_i = (A\mathbf{c})_i \Delta x_i + R(\Delta\mathbf{x}) \, , \tag{22}$$

with $R(\Delta\mathbf{x}) = \Delta x_i (A\Delta\mathbf{x})_i = \mathcal{O}(\|\Delta\mathbf{x}\|^2)$.

Next, we discuss the stability of the condensate vectors in the linear approximation (linear stability analysis of fixed points). The cardinality of the set I is referred to as $|I|$. After relabeling of the indices, one obtains up to linear order in $\|\Delta\mathbf{x}\|$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\Delta\mathbf{x}_I\\ \Delta\mathbf{x}_{\bar{I}}\end{pmatrix} = \begin{pmatrix}\tilde{A}_c & B\\ 0 & \tilde{A}_s\end{pmatrix}\begin{pmatrix}\Delta\mathbf{x}_I\\ \Delta\mathbf{x}_{\bar{I}}\end{pmatrix} =: \tilde{A}\Delta\mathbf{x}\,, \tag{23}$$

with $\tilde{A}_c$ denoting the $(|I| \times |I|)$-dimensional matrix with elements $(\tilde{A}_c)_{i,j} = c_i a_{ij}$ for $i, j \in I$. $\tilde{A}_s$ denotes the diagonal, $((S-|I|)\times(S-|I|))$-dimensional matrix with entries $(\tilde{A}_s)_{i,j} = (A\mathbf{c})_i\delta_{ij}$ for $i, j \in \bar{I}$. The matrix B is of dimension $|I| \times (S-|I|)$ with elements $B_{ij} = c_i a_{ij}$ for $i \in I$ and $j \in \bar{I}$.

The eigenvalues of the matrix $\tilde{A}$ determine the linear stability of the fixed points. Because of the block upper triangular structure of the matrix $\tilde{A}$, its eigenvalues are given by the eigenvalues of the matrices $\tilde{A}_s$ and $\tilde{A}_c$. All eigenvalues of the diagonal matrix $\tilde{A}_s$ are negative because $(A\mathbf{c})_i < 0$ for $i \in \bar{I}$. All eigenvalues of the matrix $\tilde{A}_c$ have vanishing real part. The latter can be seen from defining the nonsingular, $(|I| \times |I|)$-dimensional matrix V with elements $(V)_{i,j} = c_i\delta_{ij}$ for $i, j \in I$. Since the matrix A^I is an antisymmetric matrix, all eigenvalues of $\tilde{A}_c = VA^I$ are purely imaginary as well (the $(|I| \times |I|)$-dimensional submatrix A^I corresponds to the system of condensates with indices in I; see Methods section of the main text). However, the matrix $\tilde{A}_c$ is not antisymmetric in general. This argument can be seen as follows (see for example the Appendix in [11]): Consider the diagonal, nonsingular matrix $V^{1/2}$, whose square is the matrix V, and whose inverse is the matrix $V^{-1/2}$. The matrix $V^{-1/2}(VA^I)V^{1/2}$ has the same eigenvalues as the matrix VA^I. Since the matrix $V^{-1/2}(VA^I)V^{1/2} = V^{1/2}A^IV^{1/2}$ is antisymmetric and, thus, has purely imaginary eigenvalues, also the matrix VA^I has purely imaginary eigenvalues.

Consequently, the fixed points of the ALVE (17) that are given by the condensate vectors possess a $(S-|I|)$-dimensional local, invariant stable manifold $\mathcal{M}_s$ and a $|I|$-dimensional local, invariant center manifold $\mathcal{M}_c$ (see for example Theorem 3.2.1 in [12]). Concentrations with initial conditions chosen in $\mathcal{M}_s$ decay to zero exponentially fast. However, for initial concentrations that do not lie in $\mathcal{M}_s$, the temporal behavior cannot be inferred ("linearly stable solutions may be nonlinearly unstable" [12]).

We note that the above linear stability analysis applies to any fixed point of the ALVE. Any fixed point of the ALVE possesses an at least $|J|$-dimensional local, invariant center manifold $\mathcal{M}_c$ (see definition of J above). Therefore, a linear stability analysis of the fixed points of the ALVE (17) does not yield insight into the global dynamics of the ALVE, at least not in the straightforward fashion.

SUPPLEMENTARY NOTE 2: DYNAMICS OF CASES FOR NON-GENERIC ANTISYMMETRIC MATRICES

Some of the results presented in the main text for generic antisymmetric matrices ($A \in \Omega$) can be extended to matrices with higher dimensional kernels. Here, Ω is defined as in equation (18) of the main text as the set of antisymmetric matrices whose submatrices have a kernel with dimension less than one or equal to one:

$$\Omega = \left\{ A \in \mathbb{R}^{S\times S} \mid A \text{ is antisymmetric and } \dim\ker A^J \leq 1 \text{ for all } J \subseteq \{1,\ldots,S\} \right\} . \tag{24}$$

When submatrices of A have a kernel of dimension greater than or equal to two ($A \notin \Omega$) the statements generalize as follows: The temporal average of the projection to the surviving concentrations converges to the positive kernel of the attractive subsystem. This convergence takes place on a time scale which is not slower than $1/t$. More precisely, the distance between the time average of the concentrations of the selected states and the kernel of the surviving subsystem tends to zero for large times:

$$\text{dist}\left(\langle \mathbf{x}_I \rangle_t , \ker A^I\right) \leq \frac{Const(A, \mathbf{x}_0)}{t} . \tag{25}$$

Typically, one still finds exponentially fast depletion of states. The effective bounds on the depletion rates may depend on initial conditions.

SUPPLEMENTARY NOTE 3: LINEAR PROGRAMMING ALGORITHMS

In the following, we supplement the description and discussion of the linear programming algorithm from the Methods section in the main text. We detail on the CPEX algorithm and discuss its calibration, and provide a simplified Mathematica code.

CPLEX algorithm

The IBM ILOG CPLEX Optimization Studio 12.5 was used to numerically search for a condensate vector $\mathbf{c}$ for a given antisymmetric matrix A. Direct solution of the inequalities $A\mathbf{c} \leq 0$ and $\mathbf{c} - A\mathbf{c} \geq 1$ turned out to be numerically infeasible for systems with a large number of states S. Therefore, condensate vectors were primarily determined by solving the following linear programming problem: minimize $\epsilon_1 + \ldots + \epsilon_S$, subject to $-A\mathbf{c} + \epsilon \geq 0$ and $\mathbf{c} - A\mathbf{c} \geq 1$, with non-negative auxiliary variables $\epsilon \coloneqq (\epsilon_1, \ldots, \epsilon_S)$. The resulting vector $\mathbf{c}$ was used to define the set $I \coloneqq \{i \,|\, c_i > 5 \cdot 10^{-8}\}$ and its complement $\bar{I} = \{1, \ldots, S\} - I = \{i \,|\, c_i \leq 5 \cdot 10^{-8}\}$. The vector $\mathbf{c}$ was accepted as condensate vector, the set I as set of condensates, and the set $\bar{I}$ as set of depleted states if: $|(A\mathbf{c})_i| < 10^{-6}$ for all $i \in I$, and $-(A\mathbf{c})_i > 10^{-6}$ for all $i \in \bar{I}$. Minor extensions were added to the CPLEX algorithm to handle matrices for which an appropriate condensate vector could not be found. The Mathematica code in the following section exemplifies one of these extensions.

The above numerical thresholds were optimized by comparing inferred sets of condensates and of depleted states to sets that were derived using an alternative method. This alternative method is based on an analytical expression for kernel vectors (see equation (19) in the Methods section of the main text and [13, 14]). It is reliable but restricted to systems in which the network of states has connectivity $C = 1$ and in which the number of states is small (computational complexity grows exponentially with S). The correct identification of condensates and depleted states by the CPLEX algorithm was validated for 10^6 randomly sampled networks of states with connectivity $C = 1$ and $S = 18$ states. A detailed evaluation of the reliability of the CPLEX linear programming algorithm is provided in Supplementary Figs. 3 and 4.

Mathematica code

The following code for Mathematica 9.0 from Wolfram Research determines a condensate vector $\mathbf{c}$ by minimizing $\epsilon_1 + \ldots + \epsilon_S$, subject to $-A\mathbf{c} \geq 0$ and $\mathbf{c} - A\mathbf{c} + \epsilon \geq 1$. The resulting vector is used to infer the set of condensates I and the set of depleted states $\bar{I}$. The code can be used to verify the selection of states for the systems shown in Supplementary Figs. 1 and 2.

```
lpAlgorithm[noOfStates_, matrix_] :=
  Block[{condensateVector, condensates, depletedStates},
    Block[
    {
    solution
    (* solution: first half: condensate vector b, second half: aux. vector ε *),
    vector = Join[ConstantArray[0., noOfStates], ConstantArray[1., noOfStates]]
    (* vector: first half: w.r.t condensate vector b, second half: w.r.t. aux. vector ε *),
    lhs = ConstantArray[0., {2*noOfStates, 2*noOfStates}]
    (* lhs: upper left: -Ab, lower half: b-Ab+ε *),
    rhs = ConstantArray[0., 2*noOfStates]
    (* rhs: upper left: 0, lower half: 1 *)
    },

    Do[
      (* -Ab >= 0 *)
      Do[lhs[[i, j]] = -matrix[[i, j]], {j, 1, noOfStates}];
      rhs[[i]] = 0.;

      (* b-Ab+ε >= 1 *)
      lhs[[noOfStates + i, i]] = 1.;
      Do[lhs[[noOfStates + i, j]] -= matrix[[i, j]], {j, 1, noOfStates}];
      lhs[[noOfStates + i, noOfStates + i]] = 1.;
      rhs[[noOfStates + i]] = 1.;
    , {i, 1, noOfStates}];

    solution = Check[LinearProgramming[vector, lhs, rhs], {}];
    condensateVector = solution[[1;;noOfStates]];
    ];

  condensates = Flatten[Position[condensateVector, _?(# > 0.1&)]];
  depletedStates = Complement[Range[noOfStates], condensates];
  Return[{condensates, depletedStates, condensateVector/Total[condensateVector]}];
  ];
```

The following code can be used to verify the selection of states for the system with five states shown in Supplementary Fig. 2:

```
noOfStates = 5;
a12 = 8.3;
matrix = {{0,a12,4,-3,-5},{-a12,0,3,4,-5/3},{-4,-3,0,5,1},{3,-4,-5,0,5},{5,5/3,-1,-5,0}};

{condensates, depletedStates, condensateVector} = lpAlgorithm[noOfStates, matrix];

Print["Condensates: ", Length[condensates]];
Print["   Indices: ", condensates];
Print["   b: ", condensateVector[[condensates]]];
Print["   A.b: ", (matrix.condensateVector)[[condensates]]];
Print[];

Print["Depleted states: ", Length[depletedStates]];
Print["   Indices: ", depletedStates];
Print["   b: ", condensateVector[[depletedStates]]];
Print["   A.b: ", (matrix.condensateVector)[[depletedStates]]];
Print[];
```

To verify the selection of states for the system with 20 states that is shown in Supplementary Fig. 1, the first two lines of the above code have to be changed to:

```
noOfStates = 20;
matrix=
{{0.0000000000,0.1012965582,0.0960864501,0.1257833702,-0.0595764929,0.0924501179,-0.1016301739,
-0.0795618281,0.0487881199,-0.1122071087,-0.0286971727,-0.1000496191,-0.0103185636,0.0130714180,
-0.0572150115,-0.0129797149,-0.0706576396,-0.0389344063,0.0663966936,0.1467041892},{-0.1012965582,
0.0000000000,-0.0712011514,0.0415422138,-0.0019111200,-0.0816168047,-0.0477211059,-0.1006350936,
0.1522277516,0.0720296603,-0.2058646002,-0.1253355555,0.0717377587,-0.1063033227,-0.0942343547,
-0.0391532296,0.0049849208,-0.0929232005,-0.0840063075,0.0559612167},{-0.0960864501,0.0712011514,
0.0000000000,-0.0304803123,-0.1741821153,0.0076812872,0.0297497936,-0.0075143775,-0.0120359985,
-0.0110515825,-0.0808005342,0.2738320344,0.1170451029,-0.0357777508,-0.0896471841,-0.1577036387,
-0.3757310353,0.0296054422,-0.2259424452,0.0989073494},{-0.1257833702,-0.0415422138,0.0304803123,
0.0000000000,0.0510755453,0.0115416481,0.0985481106,0.0272202852,0.0213164598,-0.0535439544,
-0.0456534129,0.1362532620,0.1074372756,0.0375029473,-0.0533083982,0.1648089686,-0.0526653130,
0.0389051602,0.0438611333,0.0539352012},{0.0595764929,0.0019111200,0.1741821153,-0.0510755453,
0.0000000000,0.0109896214,0.1353304974,0.0912349277,0.0296142827,-0.1538255140,0.0431260101,
-0.0502994390,-0.0163964089,-0.1759377604,0.0481992186,0.0664062093,0.1736770927,-0.0448630119,
0.0535605474,-0.0257754999},{-0.0924501179,0.0816168047,-0.0076812872,-0.0115416481,-0.0109896214,
0.0000000000,-0.1090515383,-0.1188694334,0.0514000882,-0.0331443533,0.1623742129,0.0492738763,
0.0176787814,-0.1341072593,-0.0009543378,-0.0789222831,-0.0579314512,0.0892386351,-0.0686414708,
0.0492364011},{0.1016301739,0.0477211059,-0.0297497936,-0.0985481106,-0.1353304974,0.1090515383,
0.0000000000,-0.0303370160,0.0506580949,0.0225254369,0.1119589132,-0.2732763845,0.0903284019,
0.0780506743,0.1487517615,-0.0050831919,-0.0357202770,-0.1006725919,-0.0014275275,0.0744309213},
{0.0795618281,0.1006350936,0.0075143775,-0.0272202852,-0.0912349277,0.1188694334,0.0303370160,
0.0000000000,0.0881041025,0.0210453129,-0.0131581374,-0.0515644614,0.0300418276,0.0765770257,
0.1482013668,0.0876706565,-0.1600022303,-0.1501954314,0.0381309952,0.1069018065},{-0.0487881199,
-0.1522277516,0.0120359985,-0.0213164598,-0.0296142827,-0.0514000882,-0.0506580949,-0.0881041025,
0.0000000000,-0.0937867616,-0.1253362044,0.1051027292,-0.0160557361,0.0120747605,0.0410327424,
-0.1178120937,-0.0104974723,0.1001865178,0.0915443356,-0.0590317396},{0.1122071087,-0.0720296603,
0.0110515825,0.0535439544,0.1538255140,0.0331443533,-0.0225254369,-0.0210453129,0.0937867616,
0.0000000000,-0.1718419085,0.0842948432,0.1084407671,-0.1297238294,0.0768833880,-0.0866723482,
0.0062786219,-0.0986826408,-0.1352434973,-0.1425316892},{0.0286971727,0.2058646002,0.0808005342,
0.0456534129,-0.0431260101,-0.1623742129,-0.1119589132,0.0131581374,0.1253362044,0.1718419085,
0.0000000000,0.1936649126,-0.0428450371,0.0782428143,0.0592244942,-0.0554995964,0.0105651402,
-0.0184545141,0.1317087624,-0.1175718037},{0.1000496191,0.1253355555,-0.2738320344,-0.1362532620,
0.0502994390,-0.0492738763,0.2732763845,0.0515644614,-0.1051027292,-0.0842948432,-0.1936649126,
0.0000000000,0.0668839360,0.0174093226,0.0573713882,-0.0610992396,-0.0280347431,0.1058623608,
0.1781275929,0.0152443634},{0.0103185636,-0.0717377587,-0.1170451029,-0.1074372756,0.0163964089,
-0.0176787814,-0.0903284019,-0.0300418276,0.0160557361,-0.1084407671,0.0428450371,-0.0668839360,
0.0000000000,0.0571450290,0.1871354534,0.0147123474,0.1010276308,0.0366350187,0.0630761367,
-0.0728212225},{-0.0130714180,0.1063033227,0.0357777508,-0.0375029473,0.1759377604,0.1341072593,
-0.0780506743,-0.0765770257,-0.0120747605,0.1297238294,-0.0782428143,-0.0174093226,-0.0571450290,
0.0000000000,0.0158510770,-0.0301637492,0.0379895572,0.0353221008,-0.0410300505,0.0399902646},
{0.0572150115,0.0942343547,0.0896471841,0.0533083982,-0.0481992186,0.0009543378,-0.1487517615,
-0.1482013668,-0.0410327424,-0.0768833880,-0.0592244942,-0.0573713882,-0.1871354534,-0.0158510770,
0.0000000000,-0.1410043405,0.0473724443,0.1164556594,0.0120263929,0.0383652365},{0.0129797149,
0.0391532296,0.1577036387,-0.1648089686,-0.0664062093,0.0789222831,0.0050831919,-0.0876706565,
0.1178120937,0.0866723482,0.0554995964,0.0610992396,-0.0147123474,0.0301637492,0.1410043405,
0.0000000000,-0.0051486466,0.0770482371,-0.0619160030,0.1041763643},{0.0706576396,-0.0049849208,
0.3757310353,0.0526653130,-0.1736770927,0.0579314512,0.0357202770,0.1600022303,0.0104974723,
-0.0062786219,-0.0105651402,0.0280347431,-0.1010276308,-0.0379895572,-0.0473724443,0.0051486466,
0.0000000000,0.0771683333,0.0494219197,0.3405228485},{0.0389344063,0.0929232005,-0.0296054422,
-0.0389051602,0.0448630119,-0.0892386351,0.1006725919,0.1501954314,-0.1001865178,0.0986826408,
0.0184545141,-0.1058623608,-0.0366350187,-0.0353221008,-0.1164556594,-0.0770482371,-0.0771683333,
0.0000000000,-0.1856676553,0.0357968997},{-0.0663966936,0.0840063075,0.2259424452,-0.0438611333,
-0.0535605474,0.0686414708,0.0014275275,-0.0381309952,-0.0915443356,0.1352434973,-0.1317087624,
-0.1781275929,-0.0630761367,0.0410300505,-0.0120263929,0.0619160030,-0.0494219197,0.1856676553,
0.0000000000,0.1384008887},{-0.1467041892,-0.0559612167,-0.0989073494,-0.0539352012,0.0257754999,
-0.0492364011,-0.0744309213,-0.1069018065,0.0590317396,0.1425316892,0.1175718037,-0.0152443634,
0.0728212225,-0.0399902646,-0.0383652365,-0.1041763643,-0.3405228485,-0.0357968997,-0.1384008887,
0.0000000000}};
```

SUPPLEMENTARY REFERENCES

[1] M. A. Nowak and K. Sigmund, Science **303**, 793 (2004).

[2] D. Vorberg, W. Wustmann, R. Ketzmerick, and A. Eckardt, Phys. Rev. Lett. **111**, 240405 (2013).

[3] M. Grifoni and P. Hänggi, Phys. Rep. **304**, 229 (1998).

[4] C. W. Gardiner and P. Zoller, *Quantum Noise* (Springer, Berlin Heidelberg, 2004).

[5] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, 2002).

[6] C. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences* (Springer, Berlin, 2009).

[7] U. C. Täuber, *Critical Dynamics* (Cambridge University Press, Cambridge, MA, 2014).

[8] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge University Press, Cambridge, 1998).

[9] N. G. Van Kampen, *Stochastic Process in Physics and Chemistry* (Elsevier, Amsterdam, 2007).

[10] H. Kuhn and A. Tucker, *Linear Inequalities and Related Systems* (Princeton University Press, Princeton, NJ, 1956).

[11] R. M. May, *Stability and Complexity in Model Ecosystems* (Princeton University Press, Princeton, NJ, 1973).

[12] S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos* (Springer-Verlag, New York, 2003).

[13] C. E. Cullis, *Matrices and Determinoids*, Vol. I and II (Cambridge University Press, Cambridge, 1913).

[14] J. Knebel, T. Krüger, M. F. Weber, and E. Frey, Phys. Rev. Lett. **110**, 168106 (2013).

A Linear algebra of antisymmetric matrices

A.1 Antisymmetric matrices and weighted networks

We here briefly describe the formal link between antisymmetric matrices and weighted networks. A network (or directed graph) $\mathcal{N}$ consists of a set of labeled nodes (or vertices) $V(\mathcal{N}) = \{1, ..., S\}$ and a set of links (or directed edges) each of which connects two nodes $E(\mathcal{N}) = \{(1 \to 2), \ldots, (i \to j), \ldots\}$. In a *weighted network* (or weighted directed graph, or interaction network), every link is assigned the positive weight $w(i \to j)$; see figure 2 for an illustration.

On the one hand, every antisymmetric matrix $A \in \mathbb{R}^{S\times S}$ gives rise to a weighted network $\mathcal{N}(A)$. The set of vertices of $\mathcal{N}(A)$ is given by $V(\mathcal{N}(A)) = \{1, ..., S\}$ and positive matrix entries $a_{ji} > 0$ give rise to the links $(i \to j)$ with weight $w(i \to j) = a_{ji}$. If a matrix entry vanishes, $a_{ji} = 0$, nodes i and j are not connected in the network. Naturally, two nodes are connected by at most by one link; self-loops do not appear such that we deal with simple networks here.

On the other hand, the antisymmetric adjacency matrix $A(\mathcal{N}) \in \mathbb{R}^{S\times S}$ is obtained from a simple, weighted network $\mathcal{N}$ being composed of S nodes by defining for all links of the network the matrix entry $A(\mathcal{N})_{ij} = -A(\mathcal{N})_{ji}$ through the weight $a_{ij} > 0$ that corresponds to the edge $j \to i$. Therefore, the sign of an entry in the adjacency matrix corresponds to the direction of the edge (positive weight for incoming link, negative weight for outgoing link) and the absolute value denotes the magnitude of the weight.

Note that we use A to refer to both the antisymmetric matrix occurring in the ALVE and its corresponding weighted network. Note also that in our publication [36] (reprinted on pages 22–29) the direction of links and, thus, the sign of the matrix entries was defined conversely.

A.2 The Pfaffian of an antisymmetric matrix

The Pfaffian is a determinant-like function that is tailored to antisymmetric matrices. Here, the combinatorial definition of the Pfaffian of an antisymmetric matrix is presented. Let Π denote the set of all partitions of the set $\{1, 2, \ldots, 2n\}$ into ordered pairs. In other words, every partition $\alpha \in \Pi$ is pairwisely ordered in the form $\alpha = \big((i_1, j_1), (i_2, j_2), \ldots, (i_n, j_n)\big)$ with $i_k < j_k$ for all k and $i_k < i_l$ for all $k < l$. Note that there are $|\Pi| = (2n-1)\cdot(2n-3)\cdot\ldots\cdot 3\cdot 1 = (2n-1)!!$ different pairwisely ordered partitions of the set $\{1, 2, \ldots, 2n\}$. We define the permutation σ_α of such a pairwisely ordered partition $\alpha \in \Pi$ as:

$$\sigma_\alpha := \begin{pmatrix} 1 & 2 & 3 & 4 & \ldots & 2n-1 & 2n \\ (i_1 & j_1) & (i_2 & j_2) & \ldots & (i_n & j_n) \end{pmatrix} \equiv \big(i_1\ j_1\ i_2\ j_2\ \ldots\ i_n\ j_n\big)\,. \tag{26}$$

With these notions, the Pfaffian of an antisymmetric matrix $A \in \mathbb{R}^{S\times S}$ of even size $S = 2n$ with $n = 1, 2, 3, \ldots$ is defined as:

$$\mathrm{Pf}(A) \coloneqq \sum_{\alpha\in\Pi} \left(\mathrm{sign}(\sigma_\alpha) \prod_{e=1}^{n} a_{\alpha_e} \right). \tag{27}$$

Note that, because the elements of every partition α are ordered pairs ($i < j$ for every element $(i, j) \in \alpha$), every summand $\prod_{e=1}^{n} a_{\alpha_e}$ is a product of above-diagonal matrix entries of A. For an odd-sized antisymmetric matrix, the Pfaffian is 0. With this definition, one can show that $\mathrm{Pf}(A)^2 = \mathrm{Det}(A)$ [55]. Note that the Pfaffian of an antisymmetric matrix carries a sign as opposed to its determinant, which is always non-negative.

As an example, consider an arbitrary antisymmetric 2×2 matrix (with $a_{12} > 0$),

$$A_2 = \begin{pmatrix} 0 & a_{12} \\ -a_{12} & 0 \end{pmatrix} . \tag{28}$$

The set of all pairwisely ordered partitions of $\{1, 2\}$ is simply $\Pi = \{\big((1, 2)\big)\}$. Therefore, the Pfaffian of A_2 is given by:

$$\mathrm{Pf}(A_2) = a_{12} . \tag{29}$$

Consider now an arbitrary antisymmetric 4×4 matrix,

$$A_4 = \begin{pmatrix} 0 & a_{12} & a_{13} & a_{14} \\ -a_{12} & 0 & a_{23} & a_{24} \\ -a_{13} & -a_{23} & 0 & a_{34} \\ -a_{14} & -a_{24} & -a_{34} & 0 \end{pmatrix} . \tag{30}$$

The set of all pairwisely ordered partitions of the set $\{1, 2, 3, 4\}$ is given by $\Pi = \{\big((1, 2), (3, 4)\big),\ \big((1, 3), (2, 4)\big),\ \big((1, 4), (2, 3)\big)\}$. The Pfaffian of A_4 is, thus, obtained as:

$$\mathrm{Pf}(A_4) = (+1)a_{12}a_{34} + (-1)a_{13}a_{24} + (+1)a_{14}a_{23} . \tag{31}$$

The sign of the permutation was highlighted in front of the corresponding factor (for example, $\mathrm{sign}(1\ 3\ 2\ 4) \coloneqq \mathrm{sign}(\big((1, 3), (2, 4)\big)) = -1$).

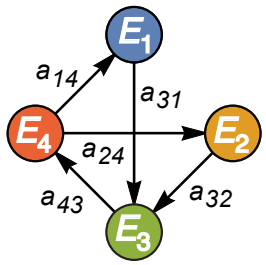


Figure 4. Computation of the Pfaffian of an antisymmetric matrix. The Pfaffian of the pretzel-like interaction network is computed according to the combinatorial definition (27).

For illustration, we compute the Pfaffian of the pretzel-like interaction network depicted in figure 4. The corresponding antisymmetric adjacency matrix is given by:

$$A_{\text{pretzel}} = \begin{pmatrix} 0 & 0 & -a_{31} & a_{14} \\ 0 & 0 & -a_{32} & a_{24} \\ a_{31} & a_{32} & 0 & -a_{43} \\ -a_{14} & -a_{24} & a_{43} & 0 \end{pmatrix} . \tag{32}$$

Its Pfaffian is computed via equation (27) as:

$$\begin{aligned} \mathrm{Pf}(A_{\text{pretzel}}) &= \mathrm{sign}(1\,2\,3\,4) \cdot 0 \cdot (-a_{43}) \\ &\quad + \mathrm{sign}(1\,3\,2\,4)(-a_{31})a_{24} + \mathrm{sign}(1\,4\,2\,3)a_{14}(-a_{32}) , && (33) \\ &= (-1)(-a_{31})a_{24} + (+1)a_{14}(-a_{32}) , && (34) \\ &= a_{31}a_{24} - a_{14}a_{32} . && (35) \end{aligned}$$

Therefore, the kernel of A_{pretzel} is only nontrivial ($\dim \mathrm{Ker}(A_{\text{pretzel}}) \geq 2$ in this case because S is even) if the matrix entries of A_{pretzel} fulfil $\mathrm{Pf}(A_{\text{pretzel}}) = a_{31}a_{24} - a_{14}a_{32} = 0$, that is, if they are fine-tuned.

For a general antisymmetric 6×6 matrix A_6,

$$A_6 = \begin{pmatrix} 0 & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ -a_{12} & 0 & a_{23} & a_{24} & a_{25} & a_{26} \\ -a_{13} & -a_{23} & 0 & a_{34} & a_{35} & a_{36} \\ -a_{14} & -a_{24} & -a_{34} & 0 & a_{45} & a_{46} \\ -a_{15} & -a_{25} & -a_{35} & -a_{45} & 0 & a_{56} \\ -a_{16} & -a_{26} & -a_{36} & -a_{46} & -a_{56} & 0 \end{pmatrix} , \tag{36}$$

the Pfaffian is obtained as:

$$\begin{aligned}\mathrm{Pf}(A_6) = &(+1)a_{12}a_{34}a_{56} + (-1)a_{12}a_{35}a_{46} + (+1)a_{12}a_{36}a_{45} \\ &+(-1)a_{13}a_{24}a_{56} + (+1)a_{13}a_{25}a_{46} + (-1)a_{13}a_{26}a_{45} \\ &+(+1)a_{14}a_{23}a_{56} + (-1)a_{14}a_{25}a_{36} + (+1)a_{14}a_{26}a_{35} \\ &+(-1)a_{15}a_{23}a_{46} + (+1)a_{15}a_{24}a_{36} + (-1)a_{15}a_{26}a_{34} \\ &+(+1)a_{16}a_{23}a_{45} + (-1)a_{16}a_{24}a_{35} + (+1)a_{16}a_{25}a_{34}\,.\end{aligned} \tag{37}$$

For further illustration of how the Pfaffian is computed, let us consider a directed cycle of even length: $S \to 1 \to 2 \to 3 \to \ldots \to S-1 \to S$ (such that $a_{1,S}, a_{21}, a_{32}, \ldots, a_{S,S-1} > 0$ with $S = 2n$ even, $n = 2, 3, \ldots$) with according antisymmetric adjacency matrix $A_{\text{even-cycle}}$:

$$A_{\text{even-cycle}} = \begin{pmatrix} 0 & -a_{21} & 0 & 0 & \ldots & a_{1,S} \\ a_{21} & 0 & -a_{32} & 0 & \ldots & 0 \\ 0 & a_{32} & 0 & -a_{43} & \ldots & 0 \\ \vdots & 0 & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & a_{S-1,S-2} & 0 & -a_{S,S-1} \\ -a_{1,S} & 0 & \ldots & 0 & a_{S,S-1} & 0 \end{pmatrix}. \tag{38}$$

The Pfaffian according to definition (27) is obtained as:

$$\mathrm{Pf}(A_{\text{even-cycle}}) = (-1)^{n \bmod 2} a_{21}a_{43}\cdots a_{S,S-1} - (-1)^{n \bmod 2} a_{32}a_{54}\cdots a_{1,S}\,. \tag{39}$$

As can be seen from above, the computation of the Pfaffian of an antisymmetric matrix proceeds in a similar manner as the computation of the determinant of an arbitrary matrix, tailored to the antisymmetry of the matrix. In this line of thought, it is not surprising that a recursive definition of the Pfaffian for an antisymmetric matrix of even size is obtained as [55, 56]:

$$\mathrm{Pf}(A) = \sum_{i=2}^{S} (-1)^i a_{1i} \mathrm{Pf}(A_{\hat{1}\hat{i}})\,, \tag{40}$$

which is mentioned here for completeness because we employed this definition of the Pfaffian in our publication [36].

A.3 The kernel of an antisymmetric matrix

The eigenvalues of an antisymmetric matrix $A \in \mathbb{R}^{S\times S}$ are either conjugated pairs of purely imaginary values or zero. To see this, consider an eigenvector $\mathbf{w}$ of A (that is,

$A\mathbf{w} = \lambda\mathbf{w}$) and compute:

$$\lambda||\mathbf{w}||^2 = \lambda\overline{\mathbf{w}}^T\mathbf{w} = \overline{\mathbf{w}}^T(A\mathbf{w}) = \left(A^T\overline{\mathbf{w}}\right)^T\mathbf{w}\,, \tag{41}$$

$$= (-A\overline{\mathbf{w}})^T\mathbf{w} = (-\overline{A\mathbf{w}})^T\mathbf{w} = -\overline{\lambda}\overline{\mathbf{w}}^T\mathbf{w}\,, \tag{42}$$

$$= -\overline{\lambda}||\mathbf{w}||^2\,. \tag{43}$$

Thus, an eigenvalue λ of A is purely imaginary (and $-\overline{\lambda}$ is an eigenvalue as well), or it is zero. Consequently, the determinant of an antisymmetric matrix is either 0 or positive. Furthermore, the rank (that is, the dimension of the image) of an antisymmetric matrix is always even, $\mathrm{rk}(A) = 0, 2, 4, \ldots, S$.

From the rank-nullity theorem it follows that, for odd $S = 2n + 1$ (with $n = 0, 1, \ldots$), the dimension of the kernel can only attain the values $\dim\mathrm{Ker}(A) = 1, 3, \ldots, S-2$, whereas it can only attain dimension $\dim\mathrm{Ker}(A) = 0, 2, \ldots, S-2, S$ for an antisymmetric matrix of even size $S = 2n$ with $n = 1, 2, \ldots$. In other words, the kernel of an odd-sized antisymmetric matrix is at least one-dimensional, that is, always nontrivial, and its determinant vanishes, $\mathrm{Det}(A) = 0$. The kernel of an even-sized antisymmetric matrix, however, is only nontrivial if a polynomial relation is fulfilled by the matrix entries. If the determinant of an even-sized antisymmetric matrix is non-zero, it is positive because all eigenvalues are pairs of purely imaginary values. Note that, because the rank of an antisymmetric matrix is even, the dimension of the kernel has the same parity as the size S of the matrix. By using the notion of the Pfaffian of an antisymmetric matrix A, it is possible to obtain an analytical expression for the kernel of A if its dimension is 1 (in case $S = 2n + 1$ is odd) or 2 (in case $S = 2n$ is even) via the adjugate vector or the adjugate matrix of A, respectively.

If S is odd, the kernel is odd-dimensional as well, and further characterized by the *adjugate vector* $\mathbf{r} \in \mathbb{R}^S$, which is defined as:

$$r_i = (-1)^{i+1}\mathrm{Pf}(A_{\hat{i}}) \quad,\ i = 1, \ldots, S\,. \tag{44}$$

Here, $A_{\hat{i}}$ denotes the matrix that is created by deleting the i-th row and column from A. The adjugate vector is a kernel vector of A if $\dim\mathrm{Ker}(A) = 1$; if $\dim\mathrm{Ker}(A) = 3, 5, \cdots, S-2, S$, the adjugate vector is the zero-vector $\mathbf{r} = 0$ [56].

If S is even, the kernel is even-dimensional as well, and is further characterized by the *adjugate matrix* $B \in \mathbb{R}^{S\times S}$, whose entries are defined as follows:

$$B_{ij} = \mathrm{sign}(\sigma_{ij})\mathrm{Pf}(A_{\hat{i}\hat{j}}) \quad,\ i, j = 1, \ldots, S\,. \tag{45}$$

Here, $A_{\hat{i}\hat{j}}$ denotes the matrix created by deleting both the i-th and j-th row and column from A. In case $i = j$, one defines $A_{\hat{i}\hat{i}} \coloneqq A_{\hat{i}}$ with vanishing Pfaffian because

$A_{\hat{i}}$ is of odd size. Furthermore, $\mathrm{sign}(\sigma_{ij})$ denotes the signum of the permutation,

$$\sigma_{ij} \coloneqq \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & i & i+1 & \dots & j & j+1 & \dots & 2n-1 & 2n \\ (i & j) & (1 & 2) & \dots & \hat{i} & i+1 & \dots & \hat{j} & j+1 & \dots & (2n-1 & 2n) \end{pmatrix}, \quad (46)$$

in which i and j are taken out of and put in front of the sequence $(1, 2, \dots, 2n)$. In general, the adjugate matrix can be thought of as the generalized inverse of the antisymmetric matrix A having the property $AB = -\mathrm{Pf}(A)\mathbb{1}_S$ with $\mathbb{1}_S$ denoting the unit matrix of size S. The adjugate matrix comprises two independent kernel vectors if $\dim \mathrm{Ker}(A) = 2$; if $\dim \mathrm{Ker}(A) = 4, 6, \dots, S$, the adjugate matrix is the zero-matrix [56]. If the kernel is trivial ($\dim \mathrm{Ker}(A) = 0$, that is, $\mathrm{Det}(A) \neq 0$), the adjugate matrix is proportional to the inverse matrix of A.

II Ecological feedback in quorum-sensing microbial populations

Abstract of the chapter

Autoinducers are small signaling molecules that mediate intercellular communication in microbial populations and trigger coordinated gene expression via "quorum sensing". Elucidating the mechanisms that control autoinducer production is, thus, pertinent to understanding collective microbial behavior, such as virulence and bioluminescence. Recent experiments have shown a heterogeneous promoter activity of autoinducer synthase genes, suggesting that some of the isogenic cells in a population might produce autoinducers, whereas others might not. However, the mechanism underlying this phenotypic heterogeneity in quorum-sensing microbial populations has remained elusive. In our theoretical model that is presented in this chapter, cells synthesize and secrete autoinducers into the environment, up-regulate their production in this self-shaped environment, and non-producers replicate faster than producers. We show that the coupling between ecological and population dynamics through quorum sensing can induce phenotypic heterogeneity in microbial populations, suggesting an alternative mechanism to stochastic gene expression in bistable gene regulatory circuits.

Acknowledgments related to this project

The results presented in sections 1 and 3 were developed over the course of many discussions with experimental microbiologists. I acknowledge fruitful discussions on phenotypic heterogeneity and quorum sensing with Paul Rainey, Kirsten Jung, Kai Papenfort, Wolfgang Streit, Jessica Grote, Vera Bettenworth, Friedrich Simmel, and Madeleine Opitz. The results presented in sections 2 and 4, and in appendices A, B, and D were obtained in close collaboration with Matthias Bauer. The numerical simulations presented in this chapter were carried out by Matthias Bauer. I thank Mauro Mobilia, Meike Wittmann, Florian Gartner, Markus F. Weber, and Karl Wienand for discussions on the quorum-sensing model. Our results presented in sections 1–4 and in appendices A, B, and D are currently under review for publication. Matthias Bauer and I share co-first authorship for this publication. The results presented in appendix C were obtained in close collaboration with Peter Pickl. The conceptual ideas for the proof go back to Peter's suggestions. Our results are currently in preparation for publication in a journal on mathematical physics.

1 Introduction

1.1 Microbiological background: Heterogeneous production of autoinducers in quorum-sensing microbial populations

Autoinducers are small molecules that are produced by microbes, secreted into the environment, and sensed by the cells in the population [9, 57]. Autoinducers can trigger a collective behavior of all cells in a population, which is called quorum sensing. For example, quorum sensing regulates the transcription of virulence genes in the Gram-positive bacterium *Listeria monocytogenes* [10, 58, 59] and the transcription of bioluminescence genes in the Gram-negative bacterium *Vibrio harveyi* [60, 61], and it may also autoregulate the transcription of autoinducer synthase genes [8, 62]. When the concentration of autoinducers reaches a threshold value, a coordinated and homogeneous expression of target genes may be initiated in all cells of the population [8, 57, 63], or a heterogeneous gene expression in the population may be triggered at low concentrations [10–12, 61, 63–68]. To implement all of these functions and behaviors, a microbial population needs to dynamically self-regulate the average autoinducer production.

Within a given population, the promoter activity of autoinducer synthase genes may vary between genetically identical cells [10, 11, 67, 69–71]. For example, during the growth of *L. monocytogenes* under well-mixed conditions two subpopulations were observed, one of which expressed autoinducer synthase genes, while the other did not [10]. Such a phenotypic heterogeneity was associated with biofilm formation [10, 57, 59, 71]. The stable coexistence of different phenotypes in one population may serve the division of labor or act as a bet-hedging strategy and, thus, may be beneficial for the survival and resilience of a microbial species on long time scales [12].

The mechanism by which a heterogeneous expression of autoinducer synthase genes is established when their expression is autoregulated by quorum sensing has remained elusive. For example, expression of the above mentioned autoinducer synthase genes in *L. monocytogenes* is up-regulated through quorum-sensing in single cells [8, 10, 72]. From an experimental point of view it is often not known, however, whether autoinducer synthesis is up-regulated for all autoinducer levels or only above a threshold level. To explain phenotypic heterogeneity of autoinducer production, currently favored threshold models of quorum sensing typically assume a bistable gene regulation function [14, 73–75]. For bistable regulation, cellular autoinducer synthesis is up-regulated above a threshold value of the autoinducer concentration in the population, whereas it is down-regulated below the threshold ("all-or-none" expression); see figure 1(B). Stochastic gene expression at the cellular level then explains the coexistence of different phenotypes in one population. If, however, cellular autoinducer synthesis is up-regulated for all autoinducer concentrations

(monostable up-regulation), the mechanism by which phenotypic heterogeneity can arise and is controlled has not been explained.

1.2 Summary of our quorum-sensing model

In our work, which is presented in this chapter, we demonstrated that the coupling between ecological and population dynamics through quorum sensing can control a heterogeneous production of autoinducers in quorum-sensing microbial populations. At the same time, the overall autoinducer level in the environment is robustly self-regulated, so that further quorum-sensing functions such as virulence or bioluminescence can be triggered. We studied the collective behavior of a stochastic many-particle model of quorum sensing, in which cells produce autoinducers to different degrees and secrete them into the well-mixed environment. Production of large autoinducer molecules (for example oligopeptides) and accompanied gene expression are assumed to reduce fitness such that non-producers reproduce faster than producing cells. Moreover, it is assumed that quorum sensing enables up-regulation of autoinducer production, that is, individuals can increase their production in response to the sensed average production level in the population (figure 1). As a central result, we found that the population may split into two subpopulations: one with a low, and a second with a high production rate of autoinducers. This phenotypic heterogeneity in the autoinducer production is stable for many generations and the autoinducer concentration in the population is tightly controlled by how production is up-regulated. If cellular response to the environment is absent or too frequent, phase transitions occur from heterogeneous to homogeneous populations in which all individuals produce autoinducers to the same degree. To capture these emergent dynamics, we derived the macroscopic mean-field equation (1) from the microscopic stochastic many-particle process in the spirit of the kinetic theory in statistical physics, which we refer to as the *autoinducer equation*. The analysis of the autoinducer equation explains both phenotypic heterogeneity through quorum sensing and the phase transitions to homogeneity.

The key aspect of our work is how the composition of a population changes in time when its constituents respond to an environment that is being shaped by their own activities. This ecological feedback is mediated by quorum sensing and creates an effective global coupling between the individuals in the population. Such a global coupling is reminiscent of long-range interactions in models of statistical mechanics, such as in the classical XY spin model with infinite range interactions [15, 76–81]. Our analysis suggests that quorum sensing in microbial populations can induce and control phenotypic heterogeneity as a collective behavior through such a global coupling and, notably, does not rely on a bistable gene regulatory circuit (see section 3).

2 Phenomenology and analysis of the quorum-sensing model

2.1 Definition of the stochastic many-particle process

We now introduce the quorum-sensing model for a well-mixed population of N individuals (figure 1). The phenotype of each individual $i = 1, \ldots, N$ is characterized by its production degree $p_i \in [0, 1]$, that is, the extent to which it produces and secretes autoinducers. In an experiment with microbes, the promoter activity of autoinducer synthase genes or their enzymatic activity could be a proxy for the production degree. The limiting case $p_i = 0$ denotes a non-producer, and $p_i = 1$ denotes a full producer.

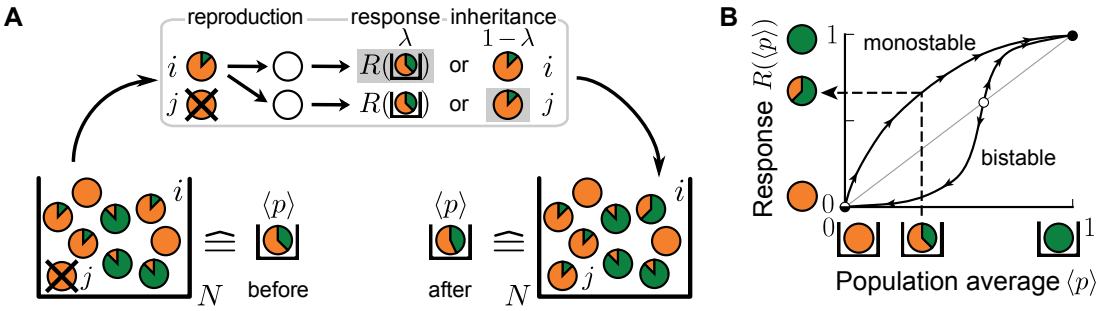


Figure 1. The quorum-sensing model for the production of autoinducers in microbial populations. (A) Sketch of a typical update step. Individuals are depicted as disks and the degree of autoinducer production ($p_i \in [0, 1]$) is indicated by the size of the green fraction. Non-producers (orange disks) reproduce fastest, full producers (green disks) slowest. Individual i with $p_i = 1/6$ divides into two offspring individuals, one of which replaces another individual j. Both offspring individuals sense the average production level in the population ($\langle p \rangle = 1/3$), and may either respond to this environment, with probability λ, by adopting the value $R(\langle p \rangle)$ of the response function (= 2/3 here, see (B)) or, with probability $1 - \lambda$, retain the production degree from the ancestor (= 1/6). Here, offspring individual i responds to the environment while j does not (denoted by gray shading). (B) Quorum sensing is characterized by the response function. Perception of the average production level in the population ($\langle p \rangle$) enables individuals to change their production degree to the value $R(\langle p \rangle) \in [0, 1]$. Sketched are a monostable response function (stable fixed point at 1, unstable fixed point at 0), and a bistable response function (stable fixed points at 0 and 1, unstable fixed point at an intermediate threshold value). Stable fixed points of the response function are depicted as black circles while unstable fixed points are colored in white. For the sketched bistable response function, autoinducer production is down-regulated with respect to the sensed production level in the population below the threshold value, and up-regulated above this threshold. For the monostable response function, autoinducer production is up-regulated at all sensed production levels.

The state of the population $\mathbf{p} = (p_1, \ldots, p_N)$ changes stochastically (figure 1(A)) through a continuous-time Markov process: An individual i reproduces randomly after a time that is exponentially distributed with rate ϕ_i, which we refer to as the individual's fitness. We assume that fitness decreases with incurring metabolic costs of induction and synthesis of autoinducers, and with other metabolic burdens in the cell's phenotypic state [82–84]. For simplicity, we choose $\phi_i = \phi(p_i) = 1 - sp_i$. The selection strength $0 \leq s < 1$ scales the fitness difference with respect to the non-producing phenotype ($\phi(0) = 1$). Thus, the larger an individual's production, the smaller its reproduction rate. This assumption is discussed in detail further below (see section 3).

Whenever an individual divides into two offspring individuals in the stochastic process, another individual from the population is selected at random to die such that the population size N remains constant. Qualitative results of our model remain valid if only the average population size is constant, which may be assumed, for example, for the stationary phase of microbial growth in batch culture. One recovers the mathematical set-up of frequency-dependent Moran models for Darwinian selection [26, 85–87] if one restricts the production degrees to a discrete set, for example, to full producers or non-producers only, $p_i \in \{0, 1\}$. The well-known Prisoner's dilemma in evolutionary game theory is recovered if, in addition, the secreted molecules would confer a fitness benefit on the population [26, 88–90]. Since we are interested in the mechanism by which heterogeneous production of autoinducers might be induced and do not study the context under which it might have evolved, we do not include any fitness benefits through signaling, for example at the population level, into the modeling here; see section 3.

A central feature of our model is the fact that individuals may adjust their production degree via a sense-and-response mechanism through quorum sensing, which is implemented as follows. After reproduction, both offspring individuals sense the average production level of autoinducers $\langle p \rangle = 1/N \sum_i p_i$ in the well-mixed population. With probability λ, they independently adopt the value $R(\langle p \rangle) \in [0, 1]$ as their production degree in response to the sensed environmental cue $\langle p \rangle$, whereas they retain the ancestor's production degree with probability $1 - \lambda$ through non-genetic inheritance. In an experimental setting, the response probability λ relates to the rate with which cells respond to the environment [91–93] and regulate their production through quorum sensing. We refer to the function $R(\langle p \rangle)$ as the response function, which is the same for all individuals. The response function encapsulates all biochemical steps involved in the autoinducer production between perception of the average production level $\langle p \rangle$ and adjustment of the individual production degree to $R(\langle p \rangle)$ in response [57, 64, 84, 94, 95]; see figure 1(B). For example, it may be a bistable step or bistable Hill function, which is often effectively assumed in threshold models of phenotypic heterogeneity [14, 73–75]. For a bistable response function, cellular production is up-regulated above a threshold value of $\langle p \rangle$, whereas

it is down-regulated below the threshold. For the bistable response function sketched in figure 1(B), both values $\langle p \rangle = 0$ and $\langle p \rangle = 1$ are stable fixed points. In this work, however, we particularly focus on monostable response functions $R(\langle p \rangle)$ to model microbial quorum-sensing systems in which autoinducer synthesis is up-regulated at all autoinducer production levels in the population [8, 72]. In other words, cellular production always increases with respect to the sensed production level in the population (stable fixed point at $\langle p \rangle = 1$ and unstable fixed point at $\langle p \rangle = 0$). The sense-and-response mechanism is further discussed in section 3.

From a mathematical point of view, the introduced sense-and-response mechanism through quorum sensing constitutes a source of innovation in the space of production degrees because an individual may adopt a production degree that was not previously present in the population. Thus, a continuous production space with $p_i \in [0, 1]$ as opposed to a discrete production space is a technical necessity for the implementation of the quorum-sensing model. The coupling of ecological dynamics (given by the average production level of autoinducers $\langle p \rangle$) with population dynamics (determined by growth rate differences between the phenotypes) through quorum sensing results in interesting collective behavior, as we show next. We emphasize that, as long as this coupling is present, the effects of the quorum-sensing model that we found and report next are qualitatively robust against noise at all steps; see below.

2.2 Heterogeneous production of autoinducers

The quorum-sensing model was numerically simulated by employing Gillespie's stochastic simulation algorithm [96, 97] for a population size of $N = 10^4$ individuals and an exemplary selection strength $s = 0.2$, such that $sN \gg 1$. In this regime, demographic fluctuations are subordinate [26, 87, 98]. Within the scope of our quorum-sensing model, the precise value of the selection strength s that scales the fitness differences is not important for the reported mechanism by which phenotypic heterogeneity can be induced, see below. We tracked the state of the population $\mathbf{p}$ over time, and depict the histogram of production degrees and the population average in figure 2.

First, we studied the stochastic many-particle process without sense-and-response ($\lambda = 0$); see figure 2(A, D). In this case, non-producers always proliferate because they reproduce at the highest rate in the population, which is well-studied in evolutionary game theory [21, 24, 32]. Thus, the initially uniform distribution in the population shifts to a peaked distribution at low production degrees. Ultimately, a *homogeneous* (unimodal) stationary state is reached in which all individuals produce autoinducers to the same low degree $p_{\text{low}} \simeq 0$. Such a stationary state is absorbing [99], that is, the stochastic process offers no possibility of escape from this state of the population.

With quorum sensing ($\lambda > 0$), absorbing states are reached if, again, all individuals produce to the same degree p^* and, in addition, the value of this production degree

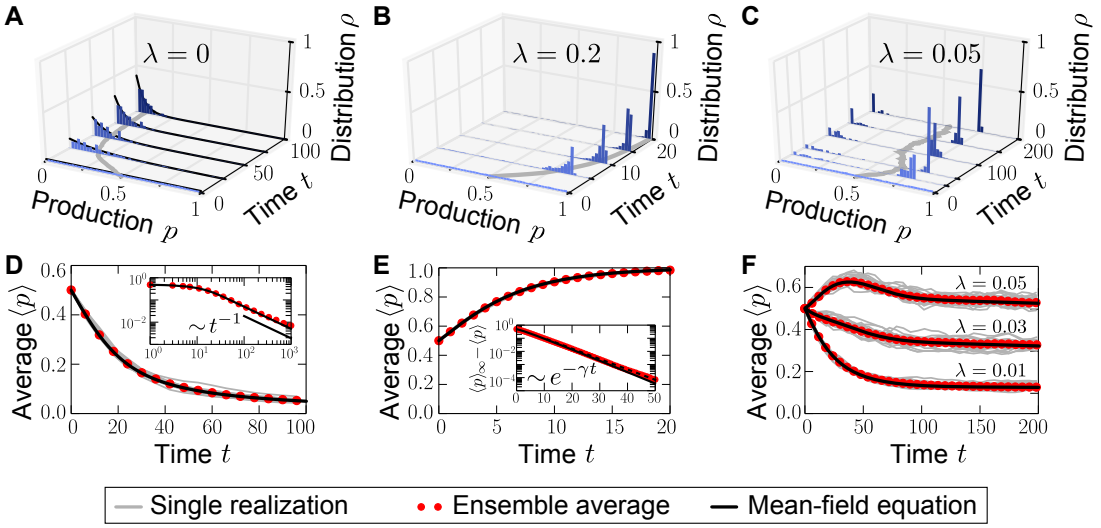


Figure 2. Homogeneous and heterogeneous production of autoinducers in the quorum-sensing model. Temporal evolution of autoinducer production in the quorum-sensing model depicted as histograms of production degrees (normalized values, A-C), and average production level of autoinducers in the population (D-F). (A) In the absence of sense-and-response ($\lambda = 0$), only non-producers proliferate. The approach to stationarity is asymptotically algebraically slow for a quasi-continuous initial distribution of production degrees (D). The black line $\langle p \rangle \sim t^{-1}$ serves as a guide for the eye. (B) Sense-and-response through quorum sensing ($\lambda = 0.2$ here) promotes autoinducer production, and the population becomes homogeneous (ultimately, fixation at a single production degree, data not shown). The response function used here, $R(\langle p \rangle) = \langle p \rangle + 0.2 \cdot \sin(\pi \langle p \rangle)$, was chosen such that an individual's production degree is always up-regulated through quorum sensing (see figure 1(B)). Approach to stationarity is exponentially fast (E), but timescales may diverge at bifurcations of the response function (see supplementary figure A.3). The dashed line in (E) shows fit to an exponential decay. (C) When λ is small ($\lambda = 0.05$ here), the population becomes heterogeneous: quasi-stationary states arise in which the population splits into two subpopulations, one of which does not produce autoinducers, while the other does. The same monostable response function was chosen as in (B). Therefore, heterogeneity may arise without bistable response. For very long times, one of the two absorbing states (A, B) is reached, data not shown (see figure 3(A)). Heterogeneous, quasi-stationary states arise for a broad class of initial distributions (see supplementary figure A.1 and our mathematical analysis). At the same time, the average production level of autoinducers in the population is adjusted by the response probability λ if s is fixed (F) or vice versa (data not shown). Bimodal, quasi-stationary states also arise when noisy inheritance, noisy perception, and noisy response are included in the model set-up (see supplementary figure A.2). Mean-field theory (1) agrees with all observations. The time unit $\Delta t = 1$ means that in a population consisting solely of non-producers, each individual will have reproduced once on average. Ensemble size $M = 100$, $s = 0.2$, $N = 10^4$. Numerical simulations were carried out by Matthias Bauer.

is a fixed point of the response function ($R(p^*) = p^*$); see figure 2(B, E). In such a homogeneous absorbing state with $\langle p \rangle_\infty = p^*$, an offspring individual can no longer alter its production degree. It either takes over the production degree p^* from its ancestor or it adopts that same degree $R(\langle p \rangle_\infty) = \langle p \rangle_\infty = p^*$ through sense-and-response. Thus, all individuals continue to produce with degree p^* and the state of the population remains *homogeneous* (unimodal).

Surprisingly, for small response probabilities λ, we found that the population may get trapped in *heterogeneous* (bimodal) states for long times before a homogeneous absorbing state is reached. The temporal evolution of such a heterogeneous state is shown in figure 2(C, F) for $\lambda = 0.05$. A monostable response function was chosen with $R(\langle p \rangle) > \langle p \rangle$ for all $\langle p \rangle \in (0, 1)$ (unstable fixed point at 0, and stable fixed point at 1) such that the production degree is always up-regulated through quorum sensing; see sketch in figure 1(B). After some time has elapsed, the population is composed of two subpopulations: one in which individuals produce autoinducers to a low degree p_{low}, and a second in which individuals produce to a higher degree p_{high} that is separated from p_{low} by a gap in the space of production degrees. Only through strong demographic fluctuations can the population reach one of the homogeneous absorbing states ($\langle p \rangle_\infty = 0$ or 1 for the response function chosen above). The time taken to reach a homogeneous absorbing state grows exponentially with N (figure 3(A)). Therefore, states of phenotypic heterogeneity are quasi-stationary and long-lived. These heterogeneous states arise for a broad class of response functions and initial distributions (supplementary figure A.1), and they are robust against demographic noise that is always present in populations of finite size (figure 3(A)); see our mathematical analysis below. We demonstrated that states of phenotypic heterogeneity are also robust against changes of the model set-up, which might account for more biological details (see, for example, [63] and references therein). Upon including, for example, noisy inheritance of the production degree, noisy perception of the environment, and noisy response to the environment into the quorum-sensing model, heterogeneous states still arise; see supplementary figure A.2. Furthermore, the average production in the heterogeneous state is finely adjusted by the interplay between the response probability λ and the selection strength s (figure 2(F)).

The establishment of long-lived, heterogeneous states induced by quorum sensing is one central finding of our study. We interpret this phenotypic heterogeneity as the result of the robust balance between population and ecological dynamics coupled through quorum sensing (see figure 4). On the one hand, fitness differences due to costly production favor non-producers. On the other hand, sensing the population average and accordingly up-regulating individual production enables producers to persist. Remarkably, fitness differences and sense-and-response balance such that *separated* production degrees may stably coexist in one population; the population does not become homogeneous at an intermediate production degree as one might

naively expect. Heterogeneity of the autoinducer production is a robust outcome of the dynamics (and not a fine-tuned effect), and the average production level in the population is adjusted by the interplay of the response probability λ and the selection strength s. Phenotypic heterogeneity does not rely on a bistable response function, but arises due to the global intercellular coupling of ecological and population dynamics through quorum sensing, as we show next. The relevance of quorum sensing for phenotypic heterogeneity in microbial populations is further explored below (see section 3).

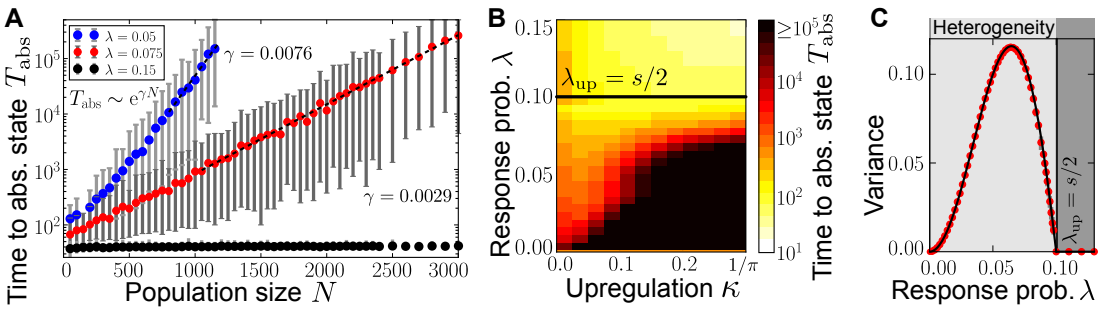


Figure 3. Characterization of phenotypic heterogeneity in the quorum-sensing model. (A) For small response probability λ, populations get stuck in heterogeneous quasi-stationary states. The time taken to reach a homogeneous absorbing state, T_{abs}, increases exponentially with the population size N (filled circles denote the mean, gray bars denote the range within which 95% of the data points lie closest to the mean; dashed lines show fit to $T_{abs} \sim e^{\gamma N}$). (B) Heterogeneous states are long-lived only if the response probability λ is small and the response function is nonlinear (in particular, up-regulation is required for some average production level such that $R(\langle p \rangle) > \langle p \rangle$). Here, the monostable response function $R(\langle p \rangle) = \langle p \rangle + \kappa \sin(\pi \langle p \rangle)$ was chosen such that $\kappa \in [0, 1/\pi]$ scales the magnitude of up-regulation. As κ increases, the gap between the low-productive and high-productive peaks of the heterogeneous state becomes larger such that it takes longer to reach the absorbing state. Mean-field theory (1) predicts the existence and local stability of heterogeneous stationary distributions for $0 < \lambda < \lambda_{up} = s/2$ (regime below the black line). Deviations between the stochastic process and mean-field theory are due to demographic fluctuations that vanish as $N \to \infty$. (C) The variance of production degrees in the population reveals whether the population is in a homogeneous ($\mathrm{Var}(p) = 0$) or heterogeneous state ($\mathrm{Var}(p) > 0$). The variance was averaged over long times in the quasi-stationary state. Mean-field theory (1) (black line) agrees with our numerical observations (red filled circles); see Methods and materials section 4. Ensemble size $M = 100$, $s = 0.2$, in (B) $N = 10^3$ and in (C) $N = 10^4$ and $N = 5 \cdot 10^4$ close to λ_{up}, in (A, C) $\kappa = 0.2$. Numerical simulations were carried out by Matthias Bauer.

2.3 Derivation and analysis of the mean-field equation (1)

In the following, the observed long-lived states of phenotypic heterogeneity in the quorum-sensing model are explained. First, we derived the macroscopic mean-field equation (the *autoinducer equation* (1)) from the microscopic dynamics of the quorum-sensing model. Second, we analyzed this mean-field equation and characterized phenotypic heterogeneity of autoinducer production.

The microscopic dynamics of the quorum-sensing model are captured by a memoryless stochastic birth-death process as sketched in figure 1. Starting from the microscopic many-particle stochastic process, we first heuristically derived (by neglecting correlations) and then proved a mean-field equation for the probability distribution of finding *any* individual at a specified production degree p at time t in the spirit of the kinetic theory in statistical physics [100]. We call this one-particle probability distribution the production distribution ρ; figure 2 shows the corresponding histogram numerically obtained from the stochastic many-particle process. The mean-field equation for ρ, which we also refer to as the *autoinducer equation*, is obtained as:

$$\partial_t \rho(p,t) = 2\lambda \overline{\phi}_t \big(\delta(p - R(\overline{p}_t)) - \rho(p,t)\big) + (1-2\lambda)\big(\phi(p) - \overline{\phi}_t\big)\rho(p,t)\,, \qquad (1)$$

where $\overline{\cdot}_t$ denotes averaging with respect to ρ at time t. The details of the derivation of the autoinducer equation from the microscopic dynamics are given in the Methods and materials section 4 and in appendix B and C.

The autoinducer equation (1) involves two contributions: the *sense-and-response* term with prefactor 2λ, and the *replicator* term with prefactor $1-2\lambda$. Through the replicator term, probability weight at production degree p changes if the fitness $\phi(p)$ is different from the mean fitness in the population $\overline{\phi}_t$ (here $\phi(p) - \overline{\phi}_t = -s(p - \overline{p}_t)$). Without quorum sensing ($\lambda = 0$), equation (1) reduces to the well-known replicator equation of the continuous Prisoner's dilemma [101–105]. The sense-and-response term, on the other hand, encodes the global feedback by which individuals adopt the production degree $R(\overline{p}_t)$ upon sensing the average $\overline{p}_t$ through quorum sensing at rate 2λ. The difference between the current state ρ and the state in which all individuals have this production degree $R(\overline{p}_t)$ determines the change in ρ at every production degree. Through the replicator term and the sense-and-response term, the ecological dynamics (average production level $\overline{p}_t$) are coupled with the dynamics of ρ.

We now present our results for the long-time behavior of the autoinducer equation (1). First, the autoinducer equation (1) admits homogeneous stationary distributions. Without quorum sensing ($\lambda = 0$), the initially lowest production degree in the population, p_{low}, constitutes the *homogeneous* stationary distribution $\rho_\infty(p) = \delta(p - p_{\mathrm{low}})$, which is attractive for generic initial conditions. With quorum sensing ($\lambda > 0$), fixed points of the response function $p^* = R(p^*)$ yield *homogeneous* stationary distributions as $\rho_\infty(p) = \delta(p - p^*)$, which are attractors of the quorum-

sensing dynamics (1) for all initial distributions if $\lambda > s/2$; see analysis below. These homogeneous stationary distributions confirm our observations of homogeneous absorbing states in the quorum-sensing model, in which all individuals produce to the same degree; see figure 2(A, B). Time scales at which stationarity is approached are discussed in the Methods and materials section 4.

Second, to analytically characterize long-lived heterogeneous states of the population, we decomposed ρ into a distribution at low production degrees and a remainder distribution at higher degrees. We found that such a decomposition yields the bimodal, *heterogeneous*, stationary distribution of the autoinducer equation (1):

$$\begin{aligned} &\rho_\infty(p) = y\delta(p) + (1-y)\delta(p-p_{\text{high}})\,, \\ &\text{with } \; p_{\text{high}} = R(\beta) \;\text{ and }\; y = 1-\beta/R(\beta)\,, \end{aligned} \tag{2}$$

if the conditions $0 < p_{\text{high}} \leq 1$ and $0 < y < 1$ are fulfilled; see figure 4 for an illustration and appendix D for the derivation. The parameter $\beta = 2\lambda/s$ quantifies the balance between fitness differences and sense-and-response mechanism through quorum sensing. Heterogeneous stationary distributions (2) are constituted of a probability mass y at the low-producing degree $p_{\text{low}} = 0$ and a coexisting δ-peak with stationary value $1-y$ at a high-producing degree p_{high} separated from p_{low} by a gap. Such heterogeneous stationary distributions have mean $\overline{p}_\infty = \beta$ and variance $\text{Var}(p)_\infty = \beta(R(\beta)-\beta)$. Therefore, the interplay between selection strength s and response probability λ adjusts the average production of autoinducers in the population (figure 2(F)). For simplicity, we assumed in equation (2) that the initially lowest production degree in the population is $p_{\text{low}} = 0$; generalized bimodal distributions for arbitrary initial distributions ρ_0 are given in appendix D.

From the conditions on p_{high} and y below equation (2), one can derive the following conditions on the response function and the value of the response probability λ (for given selection strength s) for the existence of heterogeneous stationary distributions: (i) The response function needs to be nonlinear with $R(\overline{p}_\infty) = p_{\text{high}} > \overline{p}_\infty$; that is, quorum sensing needs to up-regulate the cellular production in some regime of the average production level. Therefore, both monostable and bistable response functions depicted in figure 1(B) may induce heterogeneous stationary distributions through the ecological feedback. (ii) The response probability needs to be small with $\lambda < \lambda_{\text{up}} = s/2$; that is, to induce phenotypic heterogeneity, cells must respond only rarely to the environmental cue $\overline{p}$. This estimate of an upper bound on λ is confirmed by our numerical results of the stochastic process (figure 3(A-C)). Vice versa, for a given response probability, the selection strength needs to be big enough to induce heterogeneous stationary distributions. As we show in the Methods and materials section 4, phase transitions in the space of stationary probability distributions govern the long-time dynamics of the autoinducer equation (1) from heterogeneity

to homogeneity as the response probability changes ($\lambda \to 0$ and $\lambda \to \lambda_{up}$); see figure 3(C).

For small λ, the coexistence of the low-producing and the high-producing peaks in solution (2) is stable due to the balance of fitness differences and sense-and-response through quorum sensing. In appendix D we show that the heterogeneous stationary distributions (2) are stable up to linear order in perturbations around stationarity. As our numerical simulations show, these bimodal distributions are the attractor of the mean-field dynamics (1) for a broad range of initial distributions when λ is small; see supplementary figure A.1 for some examples. They are also robust against noisy inheritance, noisy perception, and noisy response as demonstrated in supplementary figure A.2. We interpret the stability of the bimodal stationary distributions (2) as follows (see also figure 4). Fitness differences quantified by the selection strength s increase probability mass at production degree p_{low}, whereas nonlinear response to the environment with probability λ pushes probability mass towards the up-regulated production degree $p_{high} = R(\overline{p}_\infty)$. The gap $p_{high} - p_{low} > 0$ ensures that the exponential time scales of selection and sense-and-response stably balance the coexistence of both peaks; see Methods and materials section 4. Because heterogeneous stationary distributions (2) are attractive and stable, heterogeneous states of the stochastic many-particle process arise and are quasi-stationary. Consequently, the time to reach a homogeneous absorbing state in the stochastic process through demographic fluctuations scales exponentially with the population size N [22, 106–109]; see figure 3(A). Thus, phenotypic heterogeneity is long-lived.

In summary, our mathematical analysis explains how phenotypic heterogeneity in the autoinducer production arises when quorum sensing up-regulates the autoinducer production in microbial populations (figure 4). As an emergent phenomenon, the population may split into two subpopulations: one in which cells do not produce autoinducers ('off' state, $p_{low} = 0$) and a second in which cells produce autoinducers ('on' state, $p_{high} = R(2\lambda/s) > 0$), but grow slower. The fraction of individuals in the 'off' state is given by the value of y in equation (2). If quorum sensing is absent ($\lambda = 0$), the whole population is in the 'off' state ($y = 1$), whereas all individuals are in the 'on' state ($y = 0$) if quorum sensing is frequent ($\lambda \geq \lambda_{up}$). Only when the response to the environment is rare ($0 < \lambda < \lambda_{up}$) can the two phenotypic states, p_{low} and p_{high}, coexist in the population ($0 < y < 1$). The transitions from heterogeneous to homogeneous populations are governed by nonequilibrium phase transitions when the response probability changes ($\lambda \to 0$ and $\lambda \to \lambda_{up}$). Our mathematical analysis shows that phenotypic heterogeneity arises dynamically, is robust against perturbations of the autoinducer production in the population, and is robust against noise at the level of inheritance, sense, and response.

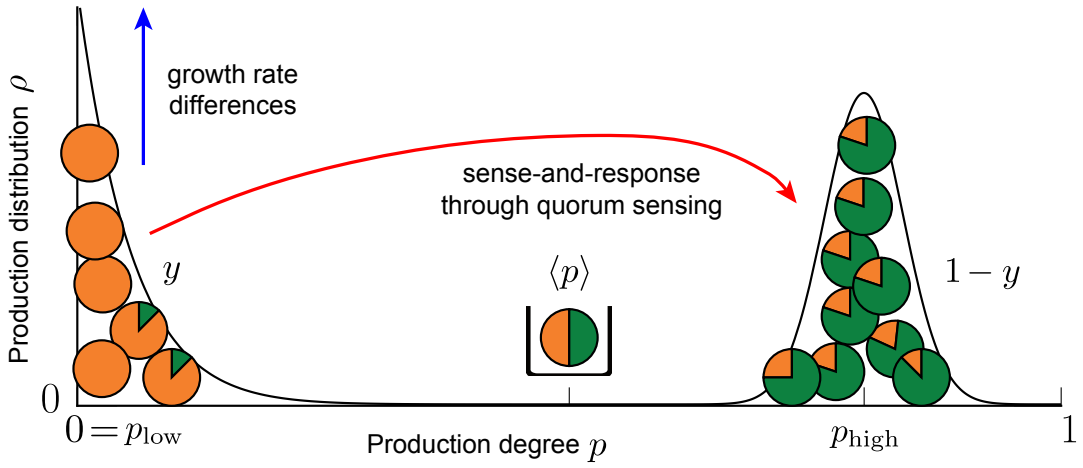


Figure 4. Dynamic stability of phenotypic heterogeneity induced by an ecological feedback in quorum-sensing microbial populations. The sketch illustrates the effective picture of how heterogeneous production of autoinducers is controlled in the quorum-sensing model. The coupling of growth rate differences between non-producers and producers (selection strength s) and sense-and-response to the self-shaped environment through quorum sensing (response probability λ and up-regulation of production with response function $R(\langle p \rangle)$) ensures the stable coexistence of the two subpopulations. The value $\beta = 2\lambda/s$ quantifies this balance. In one subpopulation (fraction $y = 1 - \beta/R(\beta)$ of the total population), individuals do not produce ($p_{\mathrm{low}} = 0$), while in the other (fraction $1 - y$) individuals produce autoinducers to the degree $p_{\mathrm{high}} = R(\beta)$. The average production level in the population is robustly adjusted to the value $\langle p \rangle = \beta$. States of phenotypic heterogeneity arise for a broad range of initial distributions and are robust against noisy inheritance, noisy perception, and noisy response (see supplementary figures A.1 and A.2).

3 Discussion of the quorum-sensing model

3.1 Summary: Phenotypic heterogeneity in the quorum-sensing model as a collective phenomenon through an ecological feedback

In this work, we studied a conceptual model for the heterogeneous production of autoinducers in quorum-sensing microbial populations. The two key assumptions of our quorum-sensing model are as follows. First, production of large autoinducer molecules and accompanied gene expression in the cell's phenotypic state are negatively correlated with fitness such that non-producers reproduce faster than producers. Second, cells sense the average production level of autoinducers in the population and may accordingly up-regulate their production through quorum sensing. As a result, not only does the interplay between fitness differences and sense-and-response give

rise to homogeneously producing populations, but it can also induce a heterogeneous production of autoinducers in the population as a stable collective phenomenon. In these heterogeneous states, the average production level of autoinducers in the population is adjusted within narrow limits by the balance between fitness differences (selection strength s in the model), and the rate with which cells respond to the environment and up-regulate their production through quorum sensing (response probability λ and response function $R(\langle p \rangle)$ in the model). Due to this robust adjustment of the production level in the population, the expression of other genes (for example, bioluminescence and virulence genes) can be regulated by quorum sensing even when the production of autoinducers is heterogeneous in the population.

In the following, we discuss the assumptions of our model in the light of the empirical reality for both quorum sensing and phenotypic heterogeneity. Furthermore, we indicate possible directions to experimentally test the ecological feedback that is suggested by the results of our theoretical work.

3.2 Does autoinducer production reduce individual growth rate?

In our quorum-sensing model, it is assumed that the individual's production degree of autoinducers is negatively correlated with its growth rate ($\phi_i = 1 - sp_i$). Is this assumption of growth impairment for producing phenotypes justified [110]? This would be the case if cellular production of autoinducers directly causes a reduction of the cell's growth rate. For example, in *L. monocytogenes* populations, heterogeneous production was observed for an autoinducer oligopeptide that is synthesized via the *agr* operon [10, 72]. This signaling oligopeptide incurs high metabolic costs through the generation of a larger pre-protein. For the oligopeptide signal synthesized via the *agr* operon in *Staphylococcus aureus*, the metabolic costs were conservatively estimated by Keller and Surette to be 184 ATP per molecule (metabolic costs for precursors were disregarded in this estimate); see [9] for details. In contrast, basically no costs (0–1 ATP) incur for the different signaling molecule Autoinducer-2 (AI-2) that is considered as a metabolic by-product. As to what extent the production of oligopeptides for signaling reduces an individual's growth rate has, to our knowledge, not been studied quantitatively.

For quorum-sensing systems that involve N-acyl homoserine lactones (AHLs) as signaling molecules, however, a reduced fitness of producers has been reported for microbial growth in batch culture [82–84]. Even though metabolic costs for the synthesis of C_4-HSL (one of the simplest AHL signaling molecules that is synthesized via the *rhl* operon) were conservatively estimated with only 8 ATP per molecule [9], a growth impairment was experimentally reported only recently for a C_4-HSL-producing strain [82]. Furthermore, a strain producing a long-chain AHL (OC_{12}-HSL, synthesized via the *las* operon) showed a reduced fitness in both mono and mixed culture compared with a non-producing strain. The reduced fitness of

AHL-producers was attributed to (i) metabolic costs of autoinducer production, in particular also to metabolic costs of precursors that were disregarded in the estimates by Keller and Surette [9], and (ii) accumulation of toxic side products accompanying the synthesis of autoinducers [82]. As another example, the strain *Sinorhizobium fredii* NGR234 synthesizes AHLs via both the *ngr* and the *tra* operon [111], and it was shown that gene expression related to autoinducer production reduces the strain's growth rate in mono culture [84]. On the other hand, a heterogeneous expression of the corresponding autoinducer synthase genes was observed during growth of NGR234 only recently [67]. As to what extent the production of AHLs reduces fitness of NGR234 in mixed culture and, thus, whether the phenotypic heterogeneity observed in [67] could be explained through the ecological feedback proposed by our quorum-sensing model, remains to be explored experimentally.

In the quorum-sensing model, even small growth rate differences between producer and non-producer, which are quantified by the ratio (growth rate of producer) / (growth rate of non-producer) = $1 - s$, may give rise to a bimodal production of autoinducers in the population. Furthermore, it would be interesting to track the expression level of autoinducer synthase genes of a microbial strain during growth for which growth differences between the producing and the non-producing phenotype are known such as in the study of [82]. We emphasize that it would be desirable to report the full distribution of expression levels in the population in order to detect whether a population splits into several subpopulations; note that variance or percentiles are not suitable measures to characterize and compare the bimodality of distributions. A bimodal expression of autoinducer synthase genes in the population together with a tightly controlled average expression level could be a signature of the feedback between ecological and population dynamics underlying the observation of phenotypic heterogeneity as suggested by our results.

3.3 A question of spatio-temporal scales: How stable and how dispersed are autoinducers in the environment?

Autoinducers are secreted into the environment where they get dispersed and are degraded. For simplicity and to facilitate our mathematical analysis, we assumed in the quorum-sensing model that individuals respond to the current average production level of autoinducers in the whole population. Temporal availability and spatial dispersal of autoinducers determine whether this assumption is valid or not. On the one hand, temporal availability of autoinducers in the environment for signaling depends on many factors. For example, pH and temperature influence the stability of autoinducers [112–116]. Biochemical mechanisms that inhibit or disrupt the functioning of signaling molecules (commonly referred to as "quorum quenching") further determine the time scales at which autoinducers are degraded in the environment [115–117]. On the other hand, spatial dispersal of autoinducers in the

population depends, for example, upon cellular mechanisms that import and export autoinducers into the cell from the environment and vice versa, and upon the spatial structure of the microbial population [57, 118]. The degree of dispersal determines whether autoinducers remain spatially privatized to a single cell, diffuse to neighboring cells, or are spread evenly between all cells of the population. Consequently, the spatio-temporal organization of the microbial population determines as to what extent microbes sense rather the current average production level or a time-integrated production of autoinducers, and to what extent they sense rather the global or a local average production level. Our quorum-sensing model assumes that autoinducers are uniformly degraded in a well-mixed environment. These assumptions do not hold true for a spatially structured microbial biofilm, but should be fulfilled during the stationary phase of microbial growth in a well-mixed batch culture [112, 113].

3.4 How is production of autoinducers up-regulated at the single-cell level?

3.4.1 Monostable or bistable up-regulation of autoinducer synthesis at the single-cell level

Our theoretical results also relate to the question of how cells regulate the production of autoinducers upon sensing the level of autoinducers in the environment. In this work, we showed that positive feedback loops and, thus, up-regulation of cellular autoinducer production may give rise to phenotypic heterogeneity. Positive feedback loops are mathematically introduced in our model as a stable fixed point at the producing phenotype of the response function (up-regulation to the stable 'on' state at $p = 1$; see figure 1(B)). Such a positive feedback is not present in all autoinducer synthase systems, but was reported for the strains *L. monocytogenes* and *S. fredii* NGR234 [8, 72, 84, 119] that showed a heterogeneous synthesis of autoinducers at the population level [10, 67]. From an experimental point of view it is often not known, however, whether autoinducer synthesis is up-regulated for all autoinducer levels or only above a threshold level. Up-regulation at all production levels in the population corresponds to a monostable response function with an unstable fixed point at the 'off' state at $p = 0$, whereas up-regulation only above a threshold level corresponds to a bistable response function with a stable fixed point at the 'off' state at $p = 0$ and an additional unstable fixed point at the threshold value; see figure 1(B). Most models of quorum-sensing microbial populations explicitly or implicitly assume a bistable gene regulation for positive feedback loops without experimental verification; see [57] for further discussion. Why might it be relevant to distinguish between bistable (for example, a Hill function with Hill coefficient > 1) and monostable (for example, a Hill function with Hill coefficient ≤ 1) regulation of autoinducer synthesis – apart from the insight on how regulation proceeds at the

molecular level? As the results of our quorum-sensing model show, the qualitative form of the regulation could discriminate between different mechanisms that control phenotypic heterogeneity of the autoinducer production at the population level as we describe in the following.

3.4.2 Heterogeneity through stochastic gene expression only for bistable gene regulation

In recent years, a deeper mechanistic understanding of phenotypic heterogeneity has been achieved by exploring how the presence of different phenotypes in a population of genetically identical cells depends upon molecular mechanisms and stochasticity at the cellular level [12]. For example, a bistable gene regulation function enables cells to switch between an 'on' and an 'off' state with respect to the expression of a certain gene or operon. Depending on environmental cues, cells are either in the stable 'on' or in the stable 'off' state. A noisy expression at intermediate concentrations of an environmental cue may then cause some cells to be in the 'on' state while others are still in the 'off' state. Thus, stochastic gene expression explains the coexistence of different phenotypic states in one population in many experimental situations [13, 120–125]. In the context of quorum sensing, the level of autoinducers in the population is the environmental cue that triggers the stochastic switch between 'on' and 'off' state explaining heterogeneous autoinducer production when the response function is bistable [14, 73–75]. In other words, bistable regulation together with stochastic gene expression can explain a bimodal autoinducer synthesis in the population. If, however, regulation of autoinducer synthesis is monostable, an explanation of phenotypic heterogeneity in the autoinducer production in terms of stochastic gene expression appears questionable to us.

3.4.3 Heterogeneity through an ecological feedback for monostable and for bistable gene regulation

The analysis of our quorum-sensing model suggests that an alternative mechanism could explain a heterogeneous production of autoinducers in quorum-sensing microbial populations. Our results show that phenotypic heterogeneity may also arise dynamically as a collective phenomenon for monostable regulation of autoinducer production when quorum sensing creates an ecological feedback by coupling ecological with population dynamics. Cells need to up-regulate their expression with respect to the sensed production level in the population. A threshold-like, bistable response function does not need to be assumed in the quorum-sensing model, but would work as well, to establish a bimodal production of autoinducers in the population.

Therefore, if phenotypic heterogeneity of autoinducer synthesis is observed in a microbial population and if cellular growth rate is correlated with the cell's production degree of autoinducers, then it would be worth testing experimentally

whether regulation of autoinducer synthesis is monostable or bistable. Monostable regulation would be an indicator that heterogeneity on the population level is not caused by stochastic gene expression, but actually is caused by a different mechanism such as the ecological feedback proposed here.

3.4.4 On which timescales do microbes respond to autoinducers in the environment?

Furthermore, in our implementation of the quorum-sensing model, individuals respond to the environment with response probability λ upon reproduction. The rule that offspring individuals can only respond at reproduction events represents a coarse-grained view in time to facilitate the mathematical analysis and to identify the ecological feedback. The response probability can actually be interpreted as the rate with which individuals respond to autoinducers in the environment. This cellular response rate is then effectively measured in units of the cell's reproduction rate (ϕ_i) in the quorum-sensing model. Phenotypic heterogeneity of autoinducer production arises in the quorum-sensing model if the time scale at which cells respond to autoinducers in the environment is of similar order as or larger than the time scale at which growth rate differences affect the population dynamics. This can be inferred from the prefactors of the sense-and-response term and the replicator term in the autoinducer equation (1): *Effective changes* of the distribution of autoinducer production in the population occur (i) through cellular response to autoinducers in the environment at rate $\sim 2\lambda$ and (ii) through growth rate differences at rate $\sim s$. Both contributions need to balance each other such that a bimodal production in the population is established (quantified in our model by the ratio $\beta = 2\lambda/s$; see also figure 4 for an illustration). This balance is robust against several kinds of perturbations and noise as discussed above; see supplementary figures A.1 and A.2. To understand how bacteria respond to changes of autoinducer levels in the environment and to quantify response rates, experiments at the single-cell level seem most promising to us at present.

3.5 Single-cell experiments

Some of the questions raised above may be addressed most effectively with single-cell experiments. For example, it would be desirable to simultaneously monitor, at the single-cell level, the correlations between autoinducer levels in the environment, the expression of autoinducer synthase genes, and the transcriptional regulators that mediate response to quorum sensing. Upon adjusting the level of autoinducers in a controlled manner, for example in a microfluidic device, one could characterize how cells respond to autoinducers in the environment. This way, it might be possible to answer questions of (i) how the cellular production of autoinducers is regulated (monostable or bistable regulation, or a different form of regulation), (ii) whether response times to environmental changes are stochastic and whether response rates

can be identified, (iii) as to what extent cellular response in the production of autoinducers depends on both the level of autoinducers in the environment and on the cell's present production degree, and (iv) how production of autoinducers is correlated with single-cell growth rate. In the context of the quorum-sensing model, the results of such single-cell experiments would help to identify the form of the fitness function ϕ and the response function R, to quantify the selection strength s and response probability λ, and to refine the model set-up.

Different mechanisms at the cellular (microscopic) level may yield the same behavior at the population (macroscopic) level. Therefore, observations at the population level might not discriminate between different mechanisms at the cellular level. Is phenotypic heterogeneity in the production of autoinducers an example of such a case? In this work, we discussed that phenotypic heterogeneity in the autoinducer production could be the result of stochastic gene expression in bistable gene regulation or, as suggested by our model, the result of the feedback between ecological and population dynamics. We believe that the above-mentioned single-cell experiments could elucidate the mechanisms that allow for phenotypic heterogeneity in quorum-sensing microbial populations, and help to understand how population dynamics and ecological dynamics influence each other.

3.6 What is the function of phenotypic heterogeneity in autoinducer production?

The purpose of the quorum-sensing model presented here is to explain how phenotypic heterogeneity in the autoinducer production arises and how it is controlled in quorum-sensing microbial populations. With the current model set-up, however, we did not address its function. Why might this phenotypic heterogeneity in the autoinducer production be beneficial for a microbial species on long times? From an experimental point of view, the evolutionary contexts and ecological scenarios under which this phenotypic heterogeneity may have arisen are still under investigation [10, 11, 67]. From a modeling perspective, one could extend, for example, our chosen fitness function with a term that explicitly accounts for the benefit of signaling either at the cellular or population level, and study suitable evolutionary contexts and possible ecological scenarios [57, 126–129]. Such theoretical models together with further experiments might help to clarify whether heterogeneous production of autoinducers can be regarded as a bet-hedging strategy of the population or rather serves the division of labor in the population [12].

3.7 Conclusion

Overall, our analyses suggest that feedbacks between ecological and population dynamics through signaling might generate phenotypic heterogeneity in the production

of signaling molecules itself, providing an alternative mechanism to stochastic gene expression in bistable gene-regulatory circuits. Spatio-temporal scales are important for the identified ecological feedback to be of relevance for microbial population dynamics: growth rate differences between producers and non-producers need to balance the rate at which cells respond to the environment, degradation of signaling molecules should be faster than time scales at which growth rate differences affect the population composition significantly, and signaling molecules should get dispersed in the whole population faster than they are degraded. In total, if microbes sense and respond to their self-shaped environment under these conditions, the population may not only respond as a homogeneous collective as is typically associated with quorum sensing, but may also become a robustly controlled heterogeneous collective. Further experimental and theoretical studies are needed to clarify the relevance of the different mechanisms that might control phenotypic heterogeneity, in particular for quorum-sensing microbial populations.

4 Methods and materials

4.1 Derivation of the mean-field equation (1) from section 2

The microscopic dynamics are captured by a memoryless stochastic birth-death process (a continuous-time Markov process) as sketched in figure 1. The state of the population $\mathbf{p}$ is updated by nongenetic inheritance and sense-and-response through quorum sensing such that at most two individuals i and $j \neq i$ change their production degree at one time. The temporal evolution of the corresponding joint N-particle probability distribution $P(\mathbf{p}, t)$ is governed by a master equation for the stochastic many-particle process [16–18], whose explicit form is derived from figure 1 and given in appendix B. This master equation tracks the correlated microscopic dynamics of the production degrees of all N individuals. To make analytical progress, we focused on the reduced one-particle probability distribution $\rho^{(1)}(p, t) = 1/N \langle \sum_i \delta(p - p_i) \rangle_P$ in the spirit of a kinetic theory [100] starting from the microscopic stochastic dynamics. $\rho^{(1)}$ denotes the probability distribution of finding *any* individual at a specified production degree p at time t; the numerically obtained histogram of $\rho^{(1)}$ was plotted in figure 2. The temporal evolution of $\rho^{(1)}$ is derived from the master equation, and couples to the reduced two-particle probability distribution and to the full probability distribution P through quorum sensing. By assuming that correlations are negligible, one may approximate $\rho^{(1)}$ by the mean-field distribution ρ, which we refer to as the production distribution. The mean-field equation (1) for ρ is derived in appendix B and referred to as the autoinducer equation. Note that equation (1) conserves normalization of ρ, that is, $\int \mathrm{d}p \, \partial_t \rho(p, t) = 0$.

We also proved that $\rho^{(1)}$ converges in probability to ρ as $N \to \infty$ for any finite time if initial correlations are not too strong. In other words, the autoinducer

equation (1) captures exactly the collective dynamics of the stochastic many-particle process for large N. To show this convergence, we introduced the bounded Lipschitz distance d between ρ and $\rho^{(1)}$, applied Grönwall's inequality to the temporal evolution of d, and used the law of large numbers. Similar distance measures and estimates have been used, for example, to prove that the Vlasov equation governs the macroscopic dynamics of the above-mentioned classical XY spin model with infinite range interactions [77, 130–132].

4.2 Analysis of homogeneous stationary distributions of the mean-field equation (1)

Without quorum sensing ($\lambda = 0$), one finds the analytical solution for ρ by applying the method of characteristics to equation (1) in the space of moment and cumulant generating functions as: $\rho(p,t) = \rho_0(p)e^{-stp}/\int_0^1 \mathrm{d}p\, e^{-stp}\rho_0(p)$; see appendix D for details. Thus, the initially lowest production degree in the population, p_{low}, constitutes the homogeneous stationary distribution $\rho_\infty(p) = \delta(p - p_{\mathrm{low}})$, which is attractive for generic initial conditions. Only δ-peaks at production degrees greater than p_{low} are stationary as well, but they are neither attractive nor stable. The temporal approach to the homogeneous stationary distribution is algebraically slow for continuous initial distributions ρ_0, and exponentially fast if p_{low} is separated from all greater degrees by a gap in production space; see appendix D and figure 2(D).

With quorum sensing ($\lambda > 0$), fixed points of the response function $p^* = R(p^*)$ yield homogeneous stationary distributions of the mean-field equation (1) as $\rho_\infty(p) = \delta(p - p^*)$. In particular, stable fixed points of the response function ($R'(p^*) < 1$) constitute homogeneous stationary distributions that are stable up to linear order in perturbations around stationarity. For $\lambda > s/2$, these distributions are also attractors of the mean-field dynamics (1) for all initial distributions; see appendix D. The temporal approach towards homogeneous stationary distributions with quorum sensing is generically exponentially fast (figure 2(E)). This exponentially fast approach is illustrated for the special case of a linear response function and $\lambda = 1/2$, for which one finds the analytical solution as: $\rho(p,t) = y(t)\rho_0(p) + (1 - y(t))\delta(p - \overline{p}_0)$ with $y(t) = e^{-\overline{\phi}_0 t}$. However, time scales at which stationarity is approached may diverge at bifurcations of the response function. Such can be seen, for example, if one chooses a supercritical pitchfork bifurcation of a polynomial response function and $\lambda = 1/2$; see supplementary figure A.3 and appendix D.

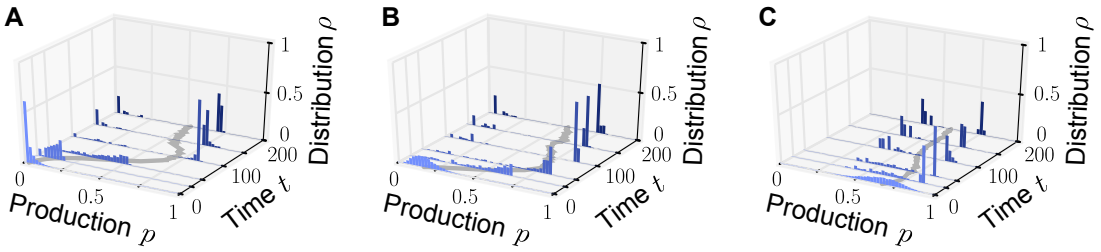
4.3 Phase transitions from heterogeneity to homogeneity in the mean-field equation (1)

Here we discuss how the long-time behavior of the quorum-sensing model changes from heterogeneous to homogeneous populations as the response probability λ

vanishes or reaches the upper threshold λ_{up} while the selection strength s is kept fixed. For small response probabilities, $0 < \lambda < \lambda_{\mathrm{up}}$, the heterogeneous stationary distributions of the autoinducer equation (1) explain the long-lived, heterogeneous states of the stochastic quorum-sensing process. The coexisting δ-peaks at the low-producing and high-producing degree in the heterogeneous stationary distribution are separated by a gap in production space, which gives rise to the non-vanishing variance $\mathrm{Var}(p)_\infty$ in the phase of heterogeneity (figure 3(C)). As $\lambda \to \lambda_{\mathrm{up}}$, the gap closes, $p_{\mathrm{high}} \to R(p_{\mathrm{high}})$, and $y \to 0$, such that a homogeneous stationary distribution with $\mathrm{Var}(p)_\infty = 0$ is recovered in a continuous transition. This nonequilibrium phase transition from heterogeneity to homogeneity proceeds without any critical behavior. As $\lambda \to 0$, and under the assumption that 0 is an unstable fixed point of the response function ($R(0) = 0$ and $1 < R'(0)$; we further assume $R'(0) < \infty$), the gap between the low-producing and the high-producing peak closes as well because $p_{\mathrm{high}} \to 0$. However, y does not approach 1, but the value $1 - 1/R'(0) < 1$. The probability weight at the low-producing mode jumps by the value $1/R'(0)$ and the homogeneous stationary distribution with $\mathrm{Var}(p)_\infty = 0$ is recovered in a discontinuous transition. Therefore, a discontinuous phase transition in the space of stationary probability distributions governs the long-time dynamics of the autoinducer equation (1) from heterogeneity to homogeneity as the response probability λ vanishes (for fixed selection strength s).

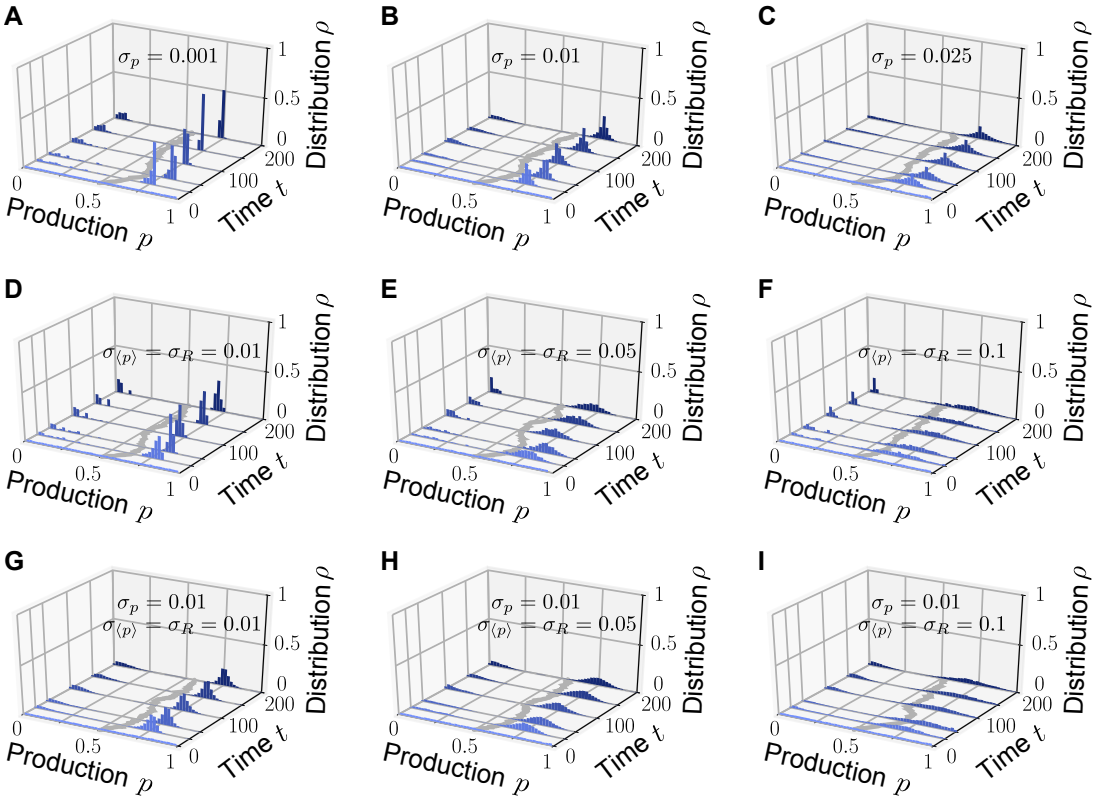
A Supplementary Figures

A.1 Supplementary figure: Phenotypic heterogeneity in the quorum-sensing model arises for diverse initial distributions



Supplementary figure A.1. Phenotypic heterogeneity in the quorum-sensing model arises for diverse initial distributions. Bimodal quasi-stationary states arise for a broad class of initial distributions if the value of the response probability λ is small and an individual's production degree is upregulated by the sense-and-response mechanism through quorum sensing ($R(p) > p$ for some $p \in [0, 1]$). Depicted is the temporal evolution of the histograms of production degrees (normalized values) as in figure 2. The monostable response function $R(p) = p + 0.2 \cdot \sin(\pi p)$ was chosen (see figure 1(B)). (A, B) $\lambda = 0.05$. Initially, the population consists of mainly non-producers (in (A) initial distribution $p_i \sim \text{Beta}(0.5, 20)$ i.i.d. and in (B) initial distribution $p_i \sim \text{Beta}(4, 20)$ i.i.d.). Due to the balance of fitness differences and sense-and-response through quorum sensing, the population splits into a heterogeneous population with producers and non-producers coexisting for long times. (C) $\lambda = 0.02$. If the initial distribution of production degrees is centered around high production degrees (initial distribution $p_i \sim \text{Beta}(10, 5)$ i.i.d.), the population may still evolve in time into a heterogeneous quasi-stationary state. However, the peak at the low-producing degree is typically located away from 0, that is, $p_{\text{low}} > 0$. These exemplary numerical results (A-C) are confirmed by the results of our mean-field theory: heterogeneous stationary distributions are the attractor of the mean-field dynamics (autoinducer equation (1)) for a broad range of initial distributions if conditions (i) $R(\overline{p}_\infty) = p_{\text{high}} > \overline{p}_\infty$ and (ii) $\lambda < \lambda_{\text{up}} = s/2$ are fulfilled. Note that i.i.d. abbreviates "independent and identically distributed". *Parameters:* selection strength $s = 0.2$ and population size $N = 10^4$. Numerical simulations were carried out by Matthias Bauer.

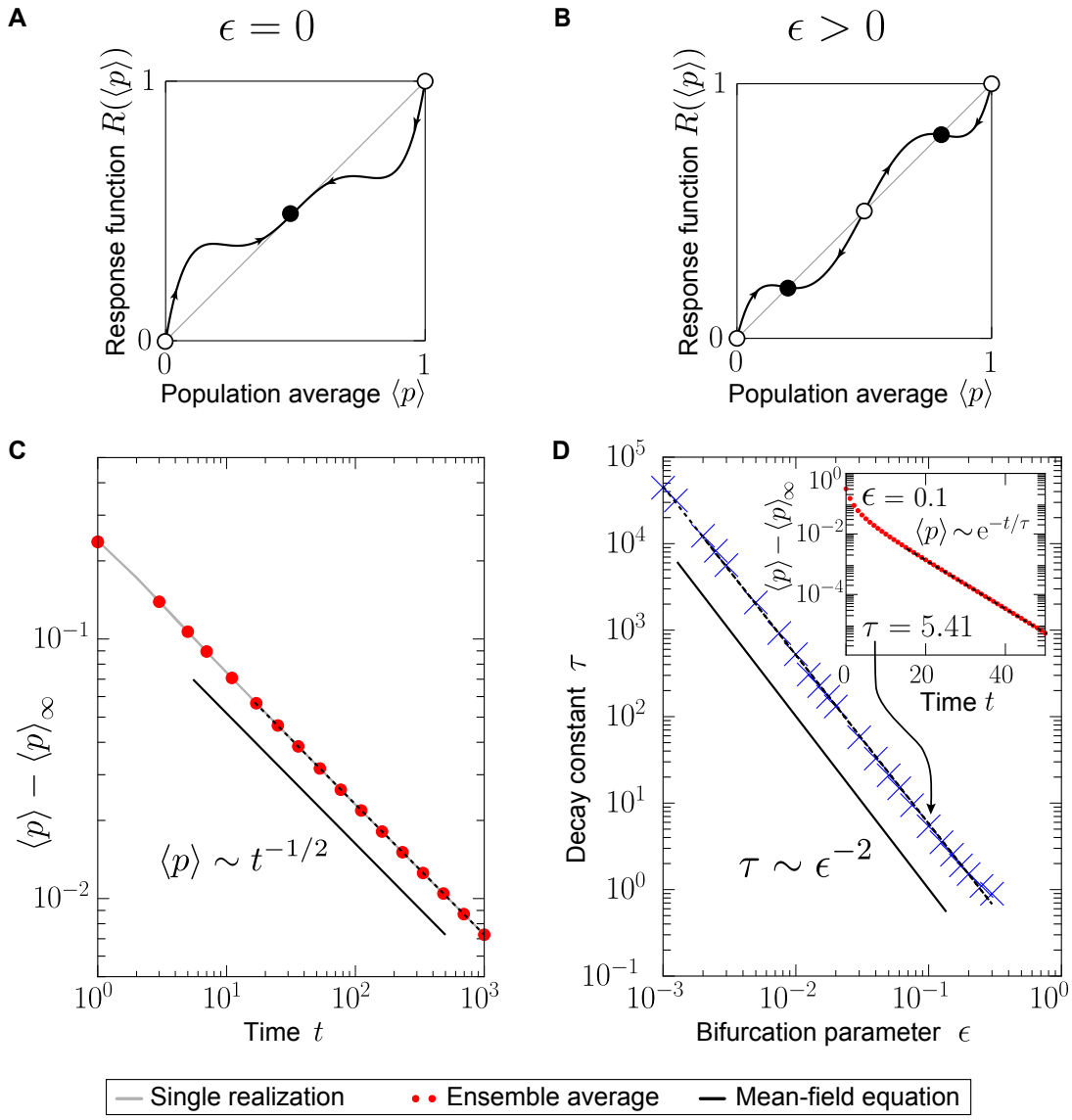
A.2 Supplementary figure: Phenotypic heterogeneity in the quorum-sensing model is robust against noisy inheritance, noisy perception, and noisy response



Supplementary figure A.2. Phenotypic heterogeneity in the quorum-sensing model is robust against noisy inheritance, noisy perception, and noisy response. Upon including either noisy inheritance of the production degree (A-C), or noisy perception of the average production level and noisy response to it (D-F), or both (G-I) into the model set-up, bimodal quasi-stationary states still arise in the relevant parameter regimes (see figure 2(C)). Depicted are representative single realizations of the modified stochastic process (histogram over normalized values of production degrees to make the comparison with figure 2 possible). (A-C) Noisy inheritance is implemented at reproduction events. Production degree p_i is passed on to an offspring as $p_i \mapsto p_i + \eta_p$ with noise $\eta_p \sim \mathcal{N}(0, \sigma_p)$ sampled from a Normal distribution (and are cut off such that $p_i + \eta_p \in [0, 1]$), emulating noisy inheritance of the phenotype. $\sigma_p \geq 0$ characterizes the strength of the noise ($\sigma_p = 0$ recovers noiseless inheritance). As σ_p increases, bimodal quasi-stationary states still arise, but the two peaks become broader than in the noiseless case. (D-F) Noise in the sensing apparatus is implemented as noisy perception of the average production level $\langle p \rangle \mapsto \langle p \rangle + \eta_{\langle p \rangle}$ with Gaussian noise $\eta_{\langle p \rangle} \sim \mathcal{N}(0, \sigma_{\langle p \rangle})$, and noise in the response is implemented at the level of the response function as $R(\langle p \rangle) \mapsto R(\langle p \rangle) + \eta_R$

with Gaussian noise $\eta_R \sim \mathcal{N}(0, \sigma_R)$. Therefore, the production degree of an individual is updated through sense-and-response to the environment as $p_i = R(\langle p \rangle) \mapsto R(\langle p \rangle + \eta_{\langle p \rangle}) + \eta_R$ in the quorum-sensing model. Again, as the strength of both sense and response noise increase, bimodal quasi-stationary states still arise, but the two peaks become broaden compared with the noiseless case. We emphasize that $\sigma_{\langle p \rangle} = \sigma_R = 0.1$ corresponds to very strong noise on the interval $[0, 1]$. (G-I) Combined effect of noisy inheritance and noisy sense-and-response. Representative trajectories demonstrate that bimodal quasi-stationary states also arise in the presence of noise at all update steps. Thus, phenotypic heterogeneity in the quorum-sensing model is qualitatively robust against noise at all steps. *Initial distribution:* $p_i \sim \text{Uniform}(0, 1)$, independent and identically distributed; *Parameters:* selection strength $s = 0.2$, response probability $\lambda = 0.05$, response function $R(\langle p \rangle) = \langle p \rangle + 0.2 \cdot \sin(\pi \langle p \rangle)$, and population size $N = 10^4$. Numerical simulations were carried out by Matthias Bauer.

A.3 Supplementary figure: Time scales at which stationarity is approached may diverge



Supplementary figure A.3. Time scales at which stationarity is approached may diverge. The response probability was set to $\lambda = 1/2$, and a nonlinear response function with bifurcation parameter ϵ was chosen as $R(\langle p \rangle) = \langle p \rangle + 40 \cdot \langle p \rangle(\langle p \rangle - (0.5 - \epsilon))(\langle p \rangle - 0.5)(\langle p \rangle - (0.5 + \epsilon))(\langle p \rangle - 1)$, see equation (166); ϵ controls a supercritical pitchfork bifurcation of the response function at the fixed point $p_{cr} = 0.5$ ($R(p_{cr}) = p_{cr}$): For $\epsilon > 0$, the fixed point at p_{cr} is unstable and non-degenerate (sketch in (B)), and becomes stable and threefold

degenerate ($z = 3$) as $\epsilon = 0$ (sketch in (A)). (D) Away from the bifurcation of the response function ($\epsilon > 0$), the approach of an absorbing state in the stochastic many-particle system is exponentially fast (see inset of (D) for an exemplary measurement of $\langle p \rangle(t) - \langle p \rangle_\infty$ for $\epsilon = 0.1$, dashed line denotes fit to exponential decay). The exponentially fast approach of stationarity is confirmed by mean-field theory ($\overline{p}_t - \overline{p}_\infty \sim e^{-t/\tau}$), see sections 2 and 4, and equation (169). Mean-field theory also predicts that the time scale of this exponentially fast relaxation diverges as $\tau \sim \epsilon^{-2}$ as the bifurcation is approached ($\epsilon \to 0$), indicated by the black line in (D). This prediction agrees with the numerical simulations of the stochastic quorum-sensing model, see (D) (blue crosses denote values of the decay constants obtained from the exponential fits and black dashed line indicates fit to $\tau \sim 1/\epsilon^\gamma$ with $\gamma = 1.95$). The divergence of time scales reflects critical slowing down as $\epsilon \to 0$. (C) At the bifurcation of the response function ($\epsilon = 0$), the approach of an absorbing state is algebraically slow, $\overline{p}_t - \overline{p}_\infty \sim t^{-1/\nu}$ with critical exponent $\nu = z - 1 = 2$ obtained from mean-field theory (black line), see equation (171). This prediction agrees with our numerical simulations of the stochastic quorum-sensing model (black dashed line in (C) indicates fit to $\langle p \rangle(t) - \langle p \rangle_\infty \sim t^\alpha$ with $\alpha = -0.50$). *Initial distribution:* unimodal $p_i \sim \text{Beta}(1, 10)$, independent and identically distributed; *Parameters:* Ensemble size $M = 100$, selection strength $s = 0.1$, population size $N = 10^4$. Numerical simulations were carried out by Matthias Bauer.

B Heuristic derivation of the mean-field equation (1)

B.1 Description of the microscopic dynamics: Master equation of the stochastic many-particle process

To describe the temporal evolution of the population, we introduced the joint N-particle probability distribution $P(\mathbf{p}, t)$. The value $P(\mathbf{p}, t)\mathrm{d}p_1 \ldots \mathrm{d}p_N$ denotes the joint probability of finding the first individual with a production degree in the interval $[p_1, p_1 + \mathrm{d}p_1]$, the second individual with a production degree in the interval $[p_2, p_2 + \mathrm{d}p_2]$, and so on at time t. The stochastic dynamics are captured by a coupled birth-death process (continuous-time Markov process) as described in section 2 and in figure 1. An individual i reproduces randomly after a time that is exponentially distributed with rate ϕ_i, which we referred to as the individual's fitness in section 2. One update step involves reproduction, sense-and-response through quorum sensing, and nongenetic inheritance such that at most two individuals i and $j \neq i$ change their production degree at one time. We denote the state of the population before the update step as $\widetilde{\mathbf{p}}^{i,j} = (p_1, \ldots, p_{i-1}, \widetilde{p}_i, p_{i+1} \ldots, \widetilde{p}_j, \ldots, p_N)$; the production degrees of individual i and j, which might change during the update step, are labeled with a tilde. For the sake of readability, we do not distinguish notationally between a random variable and the value that this random variable attains; both are labeled with the same symbol. The master equation for the joint N-particle probability distribution P for the individuals' production degrees $\mathbf{p} = (p_1, \ldots, p_N)$ at time t can be written as [16–18]:

$$\begin{aligned} \partial_t P(\mathbf{p}, t) = & \sum_{i=1}^{N} \sum_{j \neq i} \int_{[0,1]^2} \mathrm{d}\widetilde{p}_i \mathrm{d}\widetilde{p}_j \; P(\widetilde{\mathbf{p}}^{i,j}, t) \phi_i(\widetilde{\mathbf{p}}^{i,j}) \psi_j(\widetilde{\mathbf{p}}^{i,j}) A_i(\widetilde{\mathbf{p}}^{i,j}; i) A_j(\widetilde{\mathbf{p}}^{i,j}; i) \\ & - P(\mathbf{p}, t) \sum_{i=1}^{N} \sum_{j \neq i} \phi_i(\mathbf{p}) \psi_j(\mathbf{p}) \,, \\ = & \text{ “gain” } - \text{ “loss” } , \end{aligned} \tag{3}$$

with reproduction rate of individual i (fitness) given by (and selection strength $0 \leq s < 1$):

$$\phi_i(\mathbf{p}) = \phi(p_i) = 1 - s p_i \,, \tag{4}$$

and death rate of individual j given by (random death):

$$\psi_j(\mathbf{p}) = \frac{1}{N-1} \,. \tag{5}$$

The transition probabilities A_i and A_j account for the ensuing changes of at most two production degrees in the population due to nongenetic inheritance and sense-and-response (see detailed description below equations (6) and (7)). The initial condition to the master equation (3) is given as $P(\mathbf{p}, t = 0) = P_0(\mathbf{p})$.

The master equation (3) involves two contributions: *gain terms* yielding an increase and *loss terms* yielding a decrease of the probability weight in state $\mathbf{p}$ at time t. *Loss terms* occur

when the population is in state $\mathbf{p}$ and an individual reproduces. The probability of finding the population in this state is given by $P(\mathbf{p}, t)$. Individual i is selected for reproduction at rate $\phi_i(\mathbf{p})$ and splits into two offspring individuals, and a different individual $j \neq i$ is removed with probability $1/(N-1)$ at the same time (random death). *Gain terms* involve all events that take the population from an arbitrary state $\widetilde{\mathbf{p}}^{i,j}$ to state $\mathbf{p}$, and involve again reproduction for individual i and neutral death for individual j. The transition probabilities A_i and A_j account for these changes due to nongenetic inheritance and sense-and-response through quorum sensing, and are given as:

$$A_i(\widetilde{\mathbf{p}}^{i,j}; i) = \lambda \cdot \delta\left(p_i - R(\langle \widetilde{p}^{i,j} \rangle)\right) + (1-\lambda) \cdot \delta\left(p_i - \widetilde{p}_i\right) , \tag{6}$$

$$A_j(\widetilde{\mathbf{p}}^{i,j}; i) = \lambda \cdot \delta\left(p_j - R(\langle \widetilde{p}^{i,j} \rangle)\right) + (1-\lambda) \cdot \delta\left(p_j - \widetilde{p}_i\right) . \tag{7}$$

We abbreviate $\langle \widetilde{p}^{i,j} \rangle = 1/N \sum_k (\widetilde{\mathbf{p}}^{i,j})_k$ as the average production degree before the update step. Both transition probabilities A_i and A_j quantify the probability of attaining the production degrees p_i and p_j, respectively, for the two offspring individuals of ancestor i. The first summand in both A_i and A_j captures the response to the perceived average production ($p_{i/j}$ attains the value $R(\langle \widetilde{p}^{i,j} \rangle)$) with probability λ as the updated production degree, and the second summand accounts for the nongenetic inheritance of the production degree from the ancestor i ($p_{i/j}$ attains the value $\widetilde{p}_i$) with probability $1-\lambda$. Note that for the transition probability A_j, also p_j attains the value $\widetilde{p}_i$ due to our convention that individual i is labeled as the reproducing individual and individual j is chosen for the death event, see figure 1.

In our prescription of the master equation (3), the introduced gain and loss terms also involve terms that actually do not change the state of the population. Such is the case, for example, when the two individuals i and j have the same production degree ($\widetilde{p}_i = \widetilde{p}_j$) and both offspring individuals retain the production degree from their ancestor i ($p_i = p_j = \widetilde{p}_i$, that is both offspring individuals do not update their production through sense-and-response). Such events do not change the state of the population ($\widetilde{\mathbf{p}}^{i,j} = \mathbf{p}$), but are included in the master equation (3). However, these terms always occur both in the gain and loss terms. Therefore, they cancel each other and the master equation can be written in form of equation (3).

The master equation (3) conserves normalization of P because $\partial_t \int \mathrm{d}\mathbf{p}\, P(\mathbf{p}, t) = 0$; see analysis below.

B.2 Coarse-grained description: Reduced one-particle probability distribution

The reduced one-particle probability distribution $\rho^{(1)}$ is defined as:

$$\rho^{(1)}(p,t) = \frac{1}{N}\left\langle \sum_{i=1}^{N} \delta(p-p_i) \right\rangle_{P(\mathbf{p},t)}, \tag{8}$$

$$= \int_{[0,1]^{N-1}} \mathrm{d}p_2 \mathrm{d}p_3 \ldots \mathrm{d}p_N \; P(\mathbf{p},t) = P^{(1)}(p,t)\,, \tag{9}$$

and agrees with the marginal probability distribution for the production degree of the first individual $P^{(1)}$. The equality between the normalized reduced one-particle distribution $\rho^{(1)}$ and the one-particle probability distribution $P^{(1)}$ follows from the symmetry of P with respect to permutation of identical (that is indistinguishable) individuals [100].

We also define the more general reduced n-particle probability distribution:

$$\rho^{(n)}(p_1,\ldots,p_n,t) \coloneqq \frac{(N-n)!}{N!}\left\langle \sum_{i_1=1}^{N} \delta(p_1-p_{i_1}) \; \ldots \sum_{\substack{i_n=1 \\ i_n\neq i_2,\ldots,i_{n-1}}}^{N} \delta(p_n-p_{i_n}) \right\rangle_{P(\mathbf{p},t)}, \tag{10}$$

$$= \int_{[0,1]^{N-n}} \mathrm{d}p_{n+1} \mathrm{d}p_{n+2} \ldots \mathrm{d}p_N \; P(\mathbf{p},t) = P^{(n)}(p_1,\cdots,p_n,t)\,, \tag{11}$$

which agrees with the marginal probability distribution for the production degrees of the first n individuals, $P^{(n)}$. In particular, one also has $P(\mathbf{p},t) = P^{(N)}(\mathbf{p},t) = \rho^{(N)}(\mathbf{p},t)$.

B.3 Towards the macroscopic dynamics: Temporal evolution of the reduced one-particle probability distribution

In the following we show that the temporal evolution equation of the reduced one-particle probability distribution is obtained from the master equation (3) as:

$$\begin{aligned}\partial_t \rho^{(1)}(p,t) = \quad & 2\lambda \int\limits_{[0,1]^N} \mathrm{d}p_1 \mathrm{d}p_2 \ldots \mathrm{d}p_N \; \rho^{(N)}(\mathbf{p},t)\,(1-sp)\,\delta(p - R(\langle p\rangle)) \\ & - 2\lambda \left(\rho^{(1)}(p,t) - s\int_0^1 \mathrm{d}p_2 \; \rho^{(2)}(p,p_2,t)\; p_2\right) \\ & + (1-2\lambda)s\left(\int_0^1 \mathrm{d}p_2 \; \rho^{(2)}(p,p_2,t)\; p_2 - p\rho^{(1)}(p,t)\right) .\end{aligned} \tag{12}$$

To derive the temporal evolution equation for $\rho^{(1)}$, we specify the production degree of one particular individual (here p_1), and integrate out the production degrees of the other $N-1$ individuals in the master equation (3):

$$\begin{aligned}&\partial_t \rho^{(1)}(p_1,t) = \\ &= \int\limits_{[0,1]^{N-1}} \mathrm{d}p_2 \ldots \mathrm{d}p_N \; \sum_{i=1}^{N}\sum_{j\neq i}^{N} \int\limits_{[0,1]^2} \mathrm{d}\widetilde{p}_i \mathrm{d}\widetilde{p}_j \; P(\widetilde{\mathbf{p}}^{i,j},t)\frac{\phi(\widetilde{p}_i)}{N-1} A_i(\widetilde{\mathbf{p}}^{i,j};i) A_j(\widetilde{\mathbf{p}}^{i,j};i) \\ &\quad - \int\limits_{[0,1]^{N-1}} \mathrm{d}p_2 \ldots \mathrm{d}p_N \; P(\mathbf{p},t) \sum_{i=1}^{N}\sum_{j\neq i}^{N} \frac{\phi(p_i)}{N-1} , \\ &= \int\limits_{[0,1]^{N-1}} \mathrm{d}p_2 \ldots \mathrm{d}p_N \;\; (\text{“gain”} - \text{“loss”}) =: I_{\text{gain}} - I_{\text{loss}} .\end{aligned} \tag{13}$$

For the loss term I_{loss}, we split the sum $\sum_i *_{(i)}$ into two contributions, $\sum_i *_{(i)} = *_{(i=1)} + \sum_{i>1} *_{(i)}$, and deal with both contributions separately to obtain:

$$I_{\text{loss}} = N P^{(1)}(p,t) - spP^{(1)}(p,t) - s(N-1)\int_0^1 \mathrm{d}p_2 \; P^{(2)}(p,p_2,t)\; p_2 . \tag{14}$$

For the gain term I_{gain}, we split up the sum $\sum_i \sum_{j\neq i} *_{(i,j)}$ that occurs in the master equation (3) into three terms as follows:

$$\sum_{i\geq 1}\sum_{\substack{j\geq 1\\ j\neq i}} *_{(i,j)} = \sum_{j>1} *_{(i=1,j)} + \sum_{i>1} *_{(i,j=1)} + \sum_{i>1}\sum_{\substack{j>1\\ j\neq i}} *_{(i,j)} . \tag{15}$$

We also introduce the notation $\mathrm{d}\widetilde{\mathbf{p}}^{i,j,\hat{k}} := \mathrm{d}p_1 \mathrm{d}p_2 \ldots \mathrm{d}\widetilde{p}_i \ldots \mathrm{d}\widetilde{p}_j \ldots \mathrm{d}\hat{p}_k \ldots \mathrm{d}p_N$ in which variables in the superscript are labeled with a tilde in the product (indices i and j in the example), and variables with a hat in the superscript are missing in the product (that is, they are not

integrated over; index k in the example). This way, the integral measure in the gain term can be decomposed as follows:

$$\begin{aligned}\int\limits_{[0,1]^{N-1}} \mathrm{d}p_2 \ldots \mathrm{d}p_N \sum_{i=1}^{N}\sum_{j\neq i}^{N} \int\limits_{[0,1]^2} \mathrm{d}\widetilde{p}_i \mathrm{d}\widetilde{p}_j \\ = \int\limits_{[0,1]^{N+1}} \sum_{j=2}^{N} \mathrm{d}\widetilde{\mathbf{p}}^{1,j}\mathrm{d}p_j + \int\limits_{[0,1]^{N+1}} \sum_{i=2}^{N} \mathrm{d}\widetilde{\mathbf{p}}^{1,i}\mathrm{d}p_i + \int\limits_{[0,1]^{N+1}} \sum_{i=2}^{N}\sum_{\substack{j=2\\ j\neq i}}^{N} \mathrm{d}\mathbf{p}^{\hat{1},i,j}\mathrm{d}p_i\mathrm{d}p_j \ .\end{aligned} \tag{16}$$

Upon plugging in the specific form of the transition probabilities and decomposing the integral measure into the three contributions, the gain term can be written as follows (note the asymmetry between the first summand ($i = 1$ term) and the second summand ($j = 1$ term); integration over suitable δ-functions of the transition probabilities was carried out as well, for example, $\int \mathrm{d}p_j \, A_j(\widetilde{\mathbf{p}}^{i,j}; i) = 1$):

$$\begin{aligned} I_{\text{gain}} = \quad & \frac{1}{N-1} \int\limits_{[0,1]^N} \sum_{j=2}^{N} \mathrm{d}\widetilde{\mathbf{p}}^{1,j} \; P(\widetilde{\mathbf{p}}^{1,j}, t)\phi(\widetilde{p}_1) \left(\lambda\delta(p_1 - R(\langle\widetilde{p}^{1,j}\rangle)) + (1-\lambda)\delta(p_1 - \widetilde{p}_1)\right) \\ & + \frac{1}{N-1} \int\limits_{[0,1]^N} \sum_{i=2}^{N} \mathrm{d}\widetilde{\mathbf{p}}^{i,1} \; P(\widetilde{\mathbf{p}}^{i,1}, t)\phi(\widetilde{p}_i) \left(\lambda\delta(p_1 - R(\langle\widetilde{p}^{i,1}\rangle)) + (1-\lambda)\delta(p_1 - \widetilde{p}_i)\right) \\ & + \frac{1}{N-1} \int\limits_{[0,1]^{N-1}} \sum_{i=2}^{N}\sum_{\substack{j=2\\ j\neq i}}^{N} \mathrm{d}\widetilde{\mathbf{p}}^{\hat{1},i,j} \; P(\widetilde{\mathbf{p}}^{i,j}, t)\phi(\widetilde{p}_i) \ . \end{aligned} \tag{17}$$

Making use of the fact that P is symmetric with respect to permutation of individuals (individuals are identical), carrying out possible integrals over δ-functions, plugging in the explicit form of the fitness function (4), and relabeling variables, one obtains for the gain term:

$$\begin{aligned} I_{\text{gain}} = \quad & 2\lambda \int\limits_{[0,1]^N} \mathrm{d}\mathbf{p} \; P(\mathbf{p}, t)\delta(p - R(\langle p\rangle))(1 - sp_1) \\ & + 2(1-\lambda)(1-sp)P^{(1)}(p,t) \\ & + (N-2)P^{(1)}(p,t) - s(N-2)\int_0^1 \mathrm{d}p_2 \; P^{(2)}(p, p_2, t)\, p_2 \ . \end{aligned} \tag{18}$$

Combining loss terms I_{loss} and gain terms I_{gain} leads to the result for the equation of motion of the reduced one-particle probability distribution $\rho^{(1)}$ that is given in equation (12).

B.4 Heuristic derivation of the macroscopic dynamics: Mean-field approximation

Upon assuming that correlations are negligible, one may approximate $\rho^{(1)}$ by its mean-field approximation ρ, which we refer to as the production distribution. The temporal evolution equation for $\rho^{(1)}$ serves as a suitable starting point to guess the mean-field equation for ρ, which is the mean-field approximation of $\rho^{(1)}$. Thus, we naively approximate $\rho^{(1)} \approx \rho$ and $\rho^{(N)} \approx \prod^N \rho$. From the temporal evolution of $\rho^{(1)}$ in equation (12), the mean-field equation for ρ is suggested as:

$$\begin{aligned}\partial_t \rho(p,t) \approx \quad & 2\lambda \int_{[0,1]^N} \mathrm{d}p_1 \mathrm{d}p_2 \ldots \mathrm{d}p_N \prod_{i=1}^{N} \rho(p_i) \ \left(1 - s\overline{p}_t\right) \delta(p - R(\overline{p}_t)) \\ & - 2\lambda \left(\rho(p,t) - s \int_0^1 \mathrm{d}p_2 \ \rho(p,t)\rho(p_2,t)p_2\right) \\ & + (1-2\lambda)s \left(\int_0^1 \mathrm{d}p_2 \ \rho(p,t)\rho(p_2,t)p_2 - p\rho(p,t)\right) ,\end{aligned} \tag{19}$$

where $\overline{\cdot}_t$ denotes averaging with respect to ρ at time t. Further collection of terms yields the mean-field equation (1):

$$\partial_t \rho(p,t) = 2\lambda\overline{\phi}_t\big(\delta(p - R(\overline{p}_t)) - \rho(p,t)\big) + (1-2\lambda)\big(\phi(p) - \overline{\phi}_t\big)\rho(p,t) , \tag{20}$$

with initial condition $\rho(p, t=0) = \rho_0(p)$, $\phi(p) = 1 - sp$, $\overline{\phi}_t = 1 - s\overline{p}_t$, and $\overline{p}_t = \int_0^1 \mathrm{d}p \ p\rho(p,t)$. Alternatively, this mean-field equation can also be written as:

$$\partial_t \rho(p,t) = 2\lambda\big(\overline{\phi}_t \delta(p - R(\overline{p}_t)) - \phi(p)\rho(p,t)\big) + \big(\phi(p) - \overline{\phi}_t\big)\rho(p,t) . \tag{21}$$

We emphasize that the mean-field equation (1) is to be understood in distributional sense, that is, it needs to be integrated over observables (for example, suitable test functions $g : [0,1] \to \mathbb{R}$, g smooth) and ρ is interpreted as a linear functional on the space of these observables. This way, ρ can be a continuous probability density function or a discrete probability mass function, or a probability distribution with both density parts and mass parts. For further mathematical details and a more rigorous notation, see appendix C.

The proof that $\rho^{(1)}$ converges in probability to ρ as $N \to \infty$ for any finite time if initial correlations are not too strong is outlined in appendix C.

C Proof for the convergence towards the mean-field equation (1)

In this appendix, we prove that the temporal evolution of the quorum-sensing model converges to the macroscopic mean-field equation (1). To prove this, we introduce an "auxiliary stochastic mean-field process" that mimics the temporal evolution of the mean-field equation and updates the production degrees of the individuals in the population in an independent manner. This way, the law of large numbers can be separated from the propagation of errors that build up due to correlations of the individuals' production degrees. We expect that our developed method of an auxiliary stochastic mean-field process will be helpful to prove the convergence of mean-field equations in other stochastic many-particle processes as well.

The results presented in this appendix were obtained in close collaboration with Peter Pickl and are in preparation for publication in a mathematical physics journal. The main ideas of the proof go back to Peter's suggestions.

C.1 Main result and outline of the proof

In this appendix, we prove that for any time $t > 0$ the empirical distribution $\rho_N^{(1)}(t)$ of the stochastic many-particle process (the microscopic description, here referred to as "microscopic system") converges in probability to the mean-field distribution ρ (the macroscopic description, here referred to as the "macroscopic system") as the number of individuals becomes large and if initial correlations are not too strong. The notion of the empirical distribution $\rho_N^{(1)}(t)$ as the one-particle distribution function was introduced in appendix B. In other words, the mean-field equation (1) captures exactly the collective dynamics of the stochastic many-particle process of the quorum-sensing model as $N \to \infty$.

First, let us define the notion of closeness between the microscopic and the macroscopic system, and formulate our main result. Since the empirical density $\rho_N^{(1)}$ is a a sum of delta functions (a histogram) while ρ is a continuous function, closeness of the two can only hold in a weak sense. Consequently, we now introduce a weak notion of distance between functionals (for example, $\rho_N^{(1)}$ and ρ) in $(L^\infty)^*$, which denotes the dual space of L^∞ (the space of all essentially bounded measurable functions). We then measure distances between probability distributions with the bounded Lipschitz metric, which is based on the bounded Lipschitz norm defined as follows. First, defining the Lipschitz norm of a function $f \in C[0,1]$ (the space of all continuous functions on $[0,1]$) as:

$$\|f\|_L := \sup_{x,y\in[0,1]} \frac{|f(x) - f(y)|}{|x - y|}, \tag{22}$$

the bounded Lipschitz norm of any functional $g \in (L^\infty)^*$ is given by:

$$\|g\|_{BL} := \sup_{\|f\|_L=1; f(0)=0} \left| \int_0^1 \mathrm{d}p \; f(p) g(p) \right| . \tag{23}$$

Note that $(L^\infty)^*$ can be identified with the space of all finitely additive finite signed measures. Furthermore, any normalized $g \in (L^\infty)^*$ (such that $\int_0^1 \mathrm{d}p\, g(p) = 1$) can naturally be identified with a probability distribution. For such a normalized g one may drop the boundary condition $f(0) = 0$ in the definition (23) of the bounded Lipschitz norm. Note that:

$$\|g\|_{BL} \leq \left| \int_0^1 \mathrm{d}p\; p|g(p)| \right| \leq \|g\|_1 = \int_0^1 \mathrm{d}p\; |g(p)| \,. \tag{24}$$

The bounded Lipschitz metric measures the distance between two functionals g and $h \in (L^\infty)^*$ as:

$$d(g,h) := \|g-h\|_{BL} = \sup_{\|f\|_L=1; f(0)=0} \left| \int \mathrm{d}p\; f(p)g(p) - \int \mathrm{d}p\; f(p)h(p) \right| \,. \tag{25}$$

The convergence of the empirical density of the microscopic system $\rho_N^{(1)}$ against the solution ρ of the macroscopic mean-field equation (1) can at most hold in a probabilistic sense: With very small probability, always the same individual might be reproduced in a single realization of the stochastic process. Such a realization would lead to a big deviation from the solution of the mean-field equation (1). However, the occurrence of such a trajectory is improbable. To capture this intuition in mathematical terms, we define convergence in probability as follows:

Definition 1. *Let $(\nu_N)_N$ be a sequence of probability densities and ν be a probability density. We write:*

$$\nu_N \xrightarrow[N\to\infty]{\text{in prob}} \nu \,, \quad \text{if for any } \epsilon > 0 : \lim_{N\to\infty} \mathbb{P}\left(d(\nu_N, \nu) > \epsilon\right) = 0 \,. \tag{26}$$

With this notion, the main result of this appendix is formulated as follows:

Theorem 1. *Let $\rho_N^{(1)}(t)$ be the empirical one-particle probability distribution of the stochastic many-particle process (the microscopic system) and $\rho(t)$ a solution of the mean-field equation (1) (the macroscopic system). We assume for the initial distribution that $\rho_N^{(1)}(0) \xrightarrow[N\to\infty]{\text{in prob}} \rho(0)$. Then:*

$$\rho_N^{(1)}(t) \xrightarrow[N\to\infty]{\text{in prob}} \rho(t) \quad \text{for any } t > 0 \,. \tag{27}$$

It is not surprising, that one of the crucial steps in proving our result exploits the law of large numbers. However, controlling the propagation of errors, which build up upon neglecting correlations of the individuals' production degrees, with mathematical rigor is not trivial. The skeleton of the proof is summarized in figure 5.

The key idea of our proof is to separate the law of large numbers argument from the estimate of the error propagation by introducing an auxiliary stochastic mean-field process. By virtue of this auxiliary process, individuals are created and annihilated in an explicitly independent manner between consecutive update steps such that the temporal evolution of the auxiliary process mimics the mean-field dynamics (1). The auxiliary process is characterized

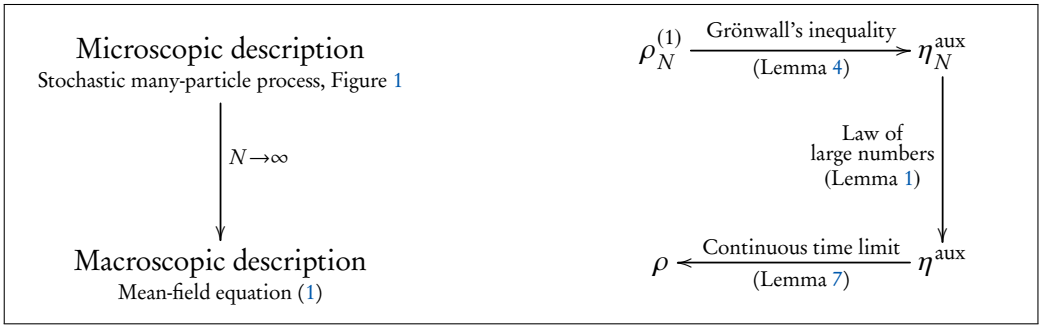


Figure 5. Sketch of the main steps of the proof for the convergence of mean-field in the quorum-sensing model. We prove that the microscopic description of the stochastic many-particle process (see figure 1) converges to the macroscopic description of the quorum-sensing model (given by the mean-field equation (1)) as $N \to \infty$. More precisely, we establish that the empirical density of the microscopic process $\rho_N^{(1)}$ converges in probability to the macroscopic mean-field density ρ as $N \to \infty$ if initial correlations are not too strong, see theorem 1. The steps of the proof are summarized on the right hand side of the figure. The central idea lies in the introduction of an auxiliary process, which mimics the time evolution of the mean-field equation as a stochastic process and updates the production degrees of the individuals in an independent manner between different update steps ("auxiliary stochastic mean-field process" with probability density η^{aux}). This way, arguments involving the law of large numbers can be separated from controlling the propagation of errors that build up due to correlations of the individuals' production degrees. Along all arrows, we show weak convergence in probability (see definition 1). The central argument and the lemma, in which the respective convergence is proven, are written next to the according arrows. Empirical densities ($\rho_N^{(1)}$ and η_N^{aux}) are denoted by the subscript N while average densities (ρ and η^{aux}) do not carry a subscript.

by the average density η^{aux}, and a single realization of the auxiliary process is denoted as η_N^{aux} (the empirical distribution of the auxiliary process). Note that we denote empirical densities ($\rho_N^{(1)}$ and η_N^{aux}) by the subscript N, whereas average densities (ρ and η^{aux}) do not depend on the system size.

The idea of an auxiliary stochastic mean-field process with independent birth and death between consecutive update steps may seem paradox because the annihilation of an individual in the microscopic, stochastic process always depends upon the actual state of the population: an individual with degree p can only be annihilated if it is existing. Thus, for the realization of the auxiliary process, we count an individual with a positive mass +1 at a birth event and an individual with a negative mass −1 at a death event. In other words, instead of creating/annihilating an individual, an individual with a positive/negative mass is created. The auxiliary process is implemented in such a way that we do not lose independence between consecutive update steps, see definition 8. As a consequence, the convergence of the empirical distribution of the auxiliary process η_N^{aux} to the average density η^{aux} as $N \to \infty$ is controlled with a standard law of large numbers argument (see lemma 1).

The propagation of errors for the convergence of the microscopic process $(\rho_N^{(1)})$ to the empirical auxiliary process (η_N^{aux}) is then controlled by Grönwall's inequality (see lemma 4). Grönwall's inequality was also applied for for the convergence of the average auxiliary process (η^{aux}) to the mean-field density ρ (see lemma 7; the law of large numbers was also implicitly applied).

Most of the following sections do *not* include the "real" time t as a variable: the microscopic system, the empirical auxiliary system, and the average auxiliary system are synchronized in time. That is, we use $k = 0, 1, 2, \ldots$ as a variable to count the ordinal number of creation/annihilation steps of the various processes. In other words, k labels the update steps. Since the time intervals between two creation/annihilation processes are distributed independently, the discrete label k is approximately a function of the continuous time t. Only when we compare the macroscopic system (ρ) and the auxiliary system (η^{aux}) is the time variable recovered, and convergence is controlled by a law of large numbers argument (see lemma 7).

Remark 1. *The basic idea of the proof is to estimate the expectation value of the distance* $d(\rho_N^{(1)}(t), \rho(t))$ *as follows:*

$$\mathbb{E}\left(d(\rho_N^{(1)}(t), \rho(t))\right) \leq Const(t) \cdot \left(\mathbb{E}\left(d(\eta_{N,0}^{aux}, \rho_0)\right) + N^{-1/4}\right), \tag{28}$$

with some constant $0 < Const < \infty$ *for any chosen time* $t > 0$.

Applying Markov's inequality then establishes an error estimate of the convergence for any $t > 0$*:*

$$\mathbb{P}\left(d(\rho_N^{(1)}(t), \rho(t)) > \epsilon_N\right) \leq \frac{\mathbb{E}\left(d(\rho_N^{(1)}(t), \rho(t))\right)}{\epsilon_N} \leq \frac{Const(t)}{\epsilon_N} \cdot \left(\mathbb{E}\left(d(\eta_{N,0}^{aux}, \rho_0)\right) + N^{-1/4}\right). \tag{29}$$

This estimate gives a quantitative control of the propagation of errors with respect to the population size N*. For example, if initial correlations vanish with* N *as* $\mathbb{E}\left(d(\eta_{N,0}^{aux}, \rho_0)\right) < Const \cdot N^{-1/8}$*, the choice* $\epsilon_N = N^{-1/8}$ *yields the estimate:*

$$\mathbb{P}\left(d(\rho_N^{(1)}(t), \rho(t)) > N^{-1/8}\right) \leq Const(t) \cdot N^{-1/8}. \tag{30}$$

This statement quantifies our intuitive reasoning from above. Realizations of the stochastic many-particle process whose one-particle distribution deviate significantly $(d(\rho_N^{(1)}(t), \rho(t)) > N^{-1/8})$ *from the solution of the mean-field equation* (1) *do occur, as is actually seen in numerical simulations of the quorum-sensing model. However, the probability* $\mathbb{P}$ *of such an occurrence is bounded from above by* $\mathbb{P} \leq Const(t) \cdot N^{-1/8}$*. Pictorially speaking, as the population size* N *grows, such occurrences become less and less likely and the magnitude of such deviations becomes smaller and smaller.*

C.2 The auxiliary stochastic mean-field process — the central idea of the proof

C.2.1 The microscopic process revisited

Before the auxiliary stochastic mean-field process is defined in definition 8, let us first revisit the set-up of the microscopic process. In the following, we give an alternative formulation of the stochastic many-particle process that is suitable for our mathematical proof.

As described above, individual i reproduces at rate ϕ_i (equation (4)) that depends both on the individual's production degree p_i and the average production level in the whole population $\langle p \rangle$. The time taken until the next reproduction event of individual i occurs is exponentially distributed with mean ϕ_i. In other words, the waiting time is drawn from the probability density $\phi_i e^{-\phi_i t}$.

For our purposes it is useful to reformulate the stochastic process in the spirit of Gillespie's stochastic kinetic Monte Carlo method [96, 97]. Instead of randomly choosing a time of reproduction for every individual independently, one can choose the time steps at which *some* individual of the population is reproduced randomly and, in a second step, define another random variable that selects *which* of the individuals reproduces. This reformulation does not change the dynamics of the microscopic, stochastic many-particle process.

For this reformulation, let $\tau_0 = 0$ and τ_k with $k = 1, 2, \ldots$ be the time at which for the k^{th} time the configuration of the population is updated (the k^{th} update step), that is, for the k^{th} time an individual is created (and another individual is annihilated at the same time). The total rate of creating any individual is given by the sum of the fitnesses of all individuals: $\sum_{i=1}^{N} \phi_i = N\langle \phi \rangle = N(1 + s(b-c)\langle p \rangle)$. It follows that all time differences $\Delta\tau_{k+1} = \tau_{k+1} - \tau_k$ are exponentially distributed with mean μ_k. Because the configuration is updated at every time τ_k, also the fitness and, thus, the parameter μ_k depend on the update step k. Writing $\langle p \rangle_k$ for the average production degree at time τ_k (that is $\langle p \rangle_k = \frac{1}{N}\sum_{i=1}^{N} p_i(\tau_k)$), it follows that $\mu_k = N(1 + s(b-c)\langle p \rangle_k) = N\langle \phi \rangle_k$.

Definition 2. *Let $\Delta\tau_{k+1}$ be the random variable for the length of the time interval between the k^{th} and $(k+1)^{th}$ update step of the stochastic many-particle process of the quorum-sensing model, that is $\langle \Delta\tau_{k+1} \rangle := 1/\mu_k = N^{-1}\left(1 + s(b-c)\langle p \rangle_k\right)^{-1}$. We define $\tau_k := \sum_{l=1}^{k-1} \Delta\tau_l$ as the update times and $\langle \tau_k \rangle := \sum_{l=1}^{k-1} \langle \Delta\tau_l \rangle$ as their according average. Furthermore, we define $M(t)$ to be the maximal natural number such that $\langle \tau_{M(t)} \rangle \leq t$, and $\kappa(t)$ as the random variable given by the maximal number such that $\tau_{\kappa(t)} \leq t$. Note that, due to the definition of the fitness, the number of update steps up to time t scales linearly with N on average, that is $M(t) \sim \mathcal{O}(N)$ such that $\langle \tau_{M(t)} \rangle \sim \mathcal{O}(N^0)$.*

Having defined the update times τ_k for $k \geq 1$, we next define the random variables that select two new individuals with production degree $p^{\text{micro},+}_{N+2k-1}$ and $p^{\text{micro},+}_{N+2k}$ for creation, and two new individuals with production degree $p^{\text{micro},-}_{2k-1}$ and $p^{\text{micro},-}_{2k}$ for annihilation at time τ_k. The random variables $p^{\text{micro},+}_{N+2k-1}$, $p^{\text{micro},+}_{N+2k}$, $p^{\text{micro},-}_{2k-1}$, and $p^{\text{micro},-}_{2k}$ map from some probability space Ω^k onto the interval $[0, 1]$ at update step k for all k. Before we define this probability space,

we first define the distribution of production degrees that we consider in the microscopic system.

The description of the microscopic system begins with a set of N individuals and their according production degrees $p_1^{\text{micro},+}, \dots, p_N^{\text{micro},+}$. The sequence of production degrees that were initially present and that have been created until time τ_k are denoted as $P_k^{\text{micro},+} := (p_1^{\text{micro},+}, \dots, p_N^{\text{micro},+}, p_{N+1}^{\text{micro},+}, \dots, p_{N+2k}^{\text{micro},+})$; the sequence of production degrees that have been annihilated until τ_k is denoted as $P_k^{\text{micro},-} := (p_1^{\text{micro},-}, \dots, p_{2k}^{\text{micro},-})$.

Definition 3. *For any pair of sequences $P_k^{micro,+}$, $P_k^{micro,-}$, the empirical one-particle distribution of the microscopic process after k update steps is given by:*

$$\rho_{N,k}^{(1)}(p) = \rho_N^{(1)}(p; P_k^{micro,+}, P_k^{micro,-}) := \frac{1}{N}\left(\sum_{j=1}^{N+2k} \delta(p - p_j^{micro,+}) - \sum_{j=1}^{2k} \delta(p - p_j^{micro,-})\right). \quad (31)$$

Note that $\rho_{N,k}^{(1)}$ is positive and fulfils $\int \mathrm{d}p\ \rho_{N,k}^{(1)} = 1$ for all update steps k. Thus, $\rho_{N,k}^{(1)}$ is a probability distribution for all k. Note also that only existing particles may be annihilated in the microscopic process. On the other hand, in the auxiliary process (see definition 8 below), individuals may be created with a negative mass at any production degree according to the present distribution of particles. Therefore, the creation and annihilation of individuals at a certain production degree is independent of the previous existence of individuals at that production degree. This way, positivity of the empirical density (η_N^{aux}) is lost for the auxiliary process, but the normalization is still valid.

For easier comparison of the random variables $p^{\text{micro},+}$ and $p^{\text{micro},-}$ of the microscopic model with the yet to be defined random variables of the auxiliary model (see definition 8 below), it is convenient to assume a constant probability density on Ω^k and choose the maps $p^{\text{micro},+}$ and $p^{\text{micro},-}$ in such a way that the creation and annihilation probabilities coincide with those of the microscopic process.

To sample a random variable from an arbitrary probability density ν, we use the following definition:

Definition 4. *For any probability density $\nu \in (L^\infty)^\star$, we define the random variable X^ν, with $[0,1] \to X^\nu$, through the so-called quantile function:*

$$X^\nu(\omega) := \inf\left\{x : \int_0^x \mathrm{d}p\ \nu(p) > \omega\right\}. \quad (32)$$

Note that, because distribution functions are continuous from the right, the infimum is in fact a minimum. Thus, the random variable $X^\nu(\omega)$, with ω being uniformly distributed on $[0,1]$, is the inverse function to the cumulative distribution function of ν that is given by $x \mapsto \int_0^x \mathrm{d}p\ \nu(p)$; see also figure 6 (upper part) for an illustration.

Following this definition, the random variable $X^\nu(\omega)$, with ω being uniformly distributed on $[0,1]$, has probability density ν. For a given probability density ν, we also define the

reproduction density that accounts for how production degrees change at an update step of the stochastic many-particle process defined by the quorum-sensing model.

Definition 5. *For any probability density $\nu : [0,1] \to \mathbb{R}_0^+$, let $\Phi(\nu)$ be the reproduction density of the quorum-sensing model, that is the probability density given by:*

$$\Phi(\nu)(p) := 2\lambda\delta\left(p - R(\langle p\rangle_\nu)\right) + (1-2\lambda)\frac{1+s\left(b\langle p\rangle_\nu - cp\right)}{1+s(b-c)\langle p\rangle_\nu}\nu(p)\,, \tag{33}$$

where we abbreviated the mean of ν as $\langle p\rangle_\nu = \int_0^1 \mathrm{d}p\; p\;\nu(p)$.

The reproduction density consists of two parts: (i) response to the average by increasing probability mass at the production degree $R(\langle p\rangle_\nu)$ with prefactor 2λ and (ii) reproduction according to relative fitness differences $\phi(p)/\phi(\overline{p})\cdot\nu(p)$ with prefactor $(1-2\lambda)$. The form of the reproduction density is chosen such that the temporal evolution of the one-particle density at discrete time steps can be recovered as we explain further below.

With these definitions, the microscopic, stochastic many-particle process can be reformulated as follows.

Definition 6. *Let the sample space Ω be given by the sequence $\Omega := (\Omega^1, \Omega^2, \ldots)$ where the individual sample spaces Ω^k at update step k are given by $\Omega^k = (\Omega_1^k, \Omega_2^k, \Omega_3^k, \Omega_4^k)$ with $\Omega_1^k = \Omega_2^k = [0,1]$ and $\Omega_3^k = \Omega_4^k = \{0,1\}$. We assume that all of the Ω_j^k are independent (both in the indeces j and k), that ω_1^k and ω_2^k are uniformly distributed on $[0,1]$, and that $\mathbb{P}(\omega_3^k = 0) = \mathbb{P}(\omega_4^k = 0) = \lambda$.*

We use this sample space to formulate now the microscopic process of the quorum-sensing model. In the k^{th} update step ω_1^k and ω_2^k determine the two individuals that are subsequently annihilated (with production degrees $p_{2k-1}^{\text{micro},-}$ and $p_{2k}^{\text{micro},-}$), and ω_3^k and ω_4^k determine the production degrees of the two created individuals ($p_{N+2k-1}^{\text{micro},+}$ and $p_{N+2k}^{\text{micro},+}$). If $\omega_{3/4}^k = 0$, then the first/second newly created individual attains the production degree given by the value $R(\langle p\rangle_\nu)$; if $\omega_{3/4}^k = 1$ then the first/second newly created individual takes over the production degree of the first annihilated individual, that is, it attains the production degree $p_{2k-1}^{\text{micro},-}$; compare with figure 1. Using this sample space and definition 4, the microscopic system of the quorum-sensing model can be reformulated as follows to reproduce the correct distribution of the individuals' production degrees:

Definition 7. *Let $p_j^{micro,+}$ be the initial production degree of the j^{th} individual for $1 \le j \le N$. The random variables $p_{N+2k}^{micro,+}$ and $p_{N+2k-1}^{micro,+}$ denote the values of the production degrees of the two individuals that are created in the k^{th} update step, and the random variables $p_{2k-1}^{micro,-}$ and $p_{2k}^{micro,-}$ denote the values of the production degrees of the two individuals that are annihilated in the k^{th}*

update step. These random variables are given by:

$$p_{2k-1}^{micro,-}(\omega^k) := X^{\Phi(\rho_{N,k-1}^{(1)})}(\omega_1^k)\,, \tag{34}$$

$$p_{2k}^{micro,-}(\omega^k) := X^{\rho_{N,k-1}^{(1)}-\delta(p-p_{2k-1}^{micro,-}(\omega_1^k))}(\omega_2^k)\,, \tag{35}$$

$$p_{N+2k-1}^{micro,+}(\omega^k) := p_{2k-1}^{micro,-}\omega_3^k + R(\langle p\rangle_{\rho_{N,k-1}^{(1)}})(1-\omega_3^k)\,, \tag{36}$$

$$p_{N+2k}^{micro,+}(\omega^k) := p_{2k-1}^{micro,-}\omega_4^k + R(\langle p\rangle_{\rho_{N,k-1}^{(1)}})(1-\omega_4^k)\,. \tag{37}$$

Let $P_k^{micro,+} := (p_1^{micro,+}, p_2^{micro,+}, \ldots, p_N^{micro,+}, \ldots, p_{N+2k}^{micro,+})$ *and* $P_k^{micro,-} := (p_1^{micro,-}, p_2^{micro,-}, \ldots, p_{2k}^{micro,-})$. *Together with definition 3, the empirical density of the microscopic system after* $k = 1, 2, \ldots$ *update steps follows as:*

$$\rho_{N,k}^{(1)}(p) = \rho_N^{(1)}(p; P_k^{micro,+}, P_k^{micro,-})\,. \tag{38}$$

Note that the values of the random variables at the update step k depend upon the probability distribution of production degrees at the update step $k-1$. The definition above assures that only individuals present at τ_{k-1} can be chosen for annihilation and, thus, to inherit their production degree $p_{2k-1}^{\mathrm{micro},-}$. It follows by induction that $\rho_{N,k}^{(1)}$ is in fact positive for all update steps k, as claimed above.

Note that $\rho_N^{(1)}(t)$ denotes the empirical probability density at time t, and $\rho_{N,k}^{(1)}$ denotes the empirical probability density after k update steps of the coupled creation-annihilation (birth-death) process. Thus, with definition 2, it follows that $\rho_N^{(1)}(t) = \rho_{N,\kappa(t)}^{(1)}$.

C.2.2 Definition of the auxiliary stochastic mean-field process

We now define the auxiliary stochastic mean-field process. Heuristically speaking, the temporal evolution of the production degrees of the population in the auxiliary process mimic the mean-field dynamics evaluated at time τ_k. The central idea of the proof for the convergence of mean-field is to construct the auxiliary process in such a way that the production degrees of the individuals at one update step are created with positive and negative masses *independently* of the realization of the previous update step (in contrast to the microscopic process). At one update step two particles of positive mass and two particles of negative mass will be created. However, their production degrees are correlated in general. Important for our purpose is the independence between consecutive update steps. The respective random variables are given by $p_{N+2k-1}^{\mathrm{aux},+}$ and $p_{N+2k}^{\mathrm{aux},+}$ for the production degrees of the created individuals, and $p_{2k-1}^{\mathrm{aux},-}$ and $p_{2k}^{\mathrm{aux},-}$ denote the production degrees of the annihilated individuals at the k^{th} update step of the auxiliary process.

In mathematical terms, a realization of the auxiliary process proceeds as follows:

Definition 8. *The average density of the auxiliary system at the $(k+1)^{th}$ update step follows from the average density at update step k as:*

$$\eta_{k+1}^{aux} = \eta_k^{aux} - \frac{1}{N}\eta_k^{aux} + \frac{1}{N}\Phi(\eta_k^{aux})\,, \tag{39}$$

and mimics the temporal evolution of the mean-field equation (1) at discretized time steps. The initial distribution of production degrees in the population for the auxiliary process is given by $\eta_0^{aux} = \rho_0$. Importantly, the update of η^{aux} is independent of the realization of the auxiliary process at the previous update step, whereas in the microscopic process the time evolution depends upon the realization of the stochastic process.

In a realization of the auxiliary stochastic mean-field process at the update step k, individuals with negative and positive masses are created independent from the production degrees of the individuals present at update step $k-1$. The respective random variables describing the values of the production degrees created and annihilated at the k^{th} update step are given by (two individuals are created and two individuals are annihilated per update step):

$$p_{2k-1}^{aux,-}(\omega^k) \coloneqq X^{\Phi(\eta_k^{aux})}(\omega_1^k)\,, \tag{40}$$

$$p_{2k}^{aux,-}(\omega^k) \coloneqq X^{\eta_k^{aux}}(\omega_2^k)\,, \tag{41}$$

$$p_{N+2k-1}^{aux,+}(\omega^k) \coloneqq p_{2k-1}^{aux,-}\omega_3^k + R(\langle p\rangle_{\eta_k^{aux}})(1-\omega_3^k)\,, \tag{42}$$

$$p_{N+2k}^{aux,+}(\omega^k) \coloneqq p_{2k-1}^{aux,-}\omega_4^k + R(\langle p\rangle_{\eta_k^{aux}})(1-\omega_4^k)\,. \tag{43}$$

The values of $p_{N+2k-1}^{aux,+}$ and $p_{N+2k}^{aux,+}$ depend on $p_{2k}^{aux,-}$ and indirectly also on each other. This dependence, however, is not problematic for our proof because independence holds still true for the vast majority of the individuals' production degrees.

C.3 Proof of the theorem for the convergence to mean-field

C.3.1 Convergence of the auxiliary process $\left(\eta^{\text{aux}}_{N,K} \xrightarrow[N\to\infty]{\text{in prob}} \eta^{\text{aux}}_{K}\right)$ – Law of large numbers argument

Because individuals are created and annihilated in an independent manner between consecutive update steps in the auxiliary process, one expects that the empirical distribution function $\eta^{\text{aux}}_{N,k}$ converges to the average density η^{aux}_{k} of the auxiliary process as $N \to \infty$. Here we show that indeed $\eta^{\text{aux}}_{N,k} \xrightarrow[N\to\infty]{\text{in prob}} \eta^{\text{aux}}_{k}$. More precisely, we have lemma 1.

Lemma 1. *One finds a constant* $0 < Const < \infty$ *such that for a given update step* K *the expected difference between a single realization of the auxiliary process* $(\eta^{aux}_{N,K})$ *and the average density* (η^{aux}_{K}) *is estimated as:*

$$\mathbb{E}\left(d(\eta^{aux}_{N,K}, \eta^{aux}_{K})\right) \leq Const \cdot \frac{K^{3/4}}{N} + \mathbb{E}\left(d(\eta^{aux}_{N,0}, \eta^{aux}_{0})\right) . \tag{44}$$

Proof. Note, that we are dealing with the auxiliary process and, thus, have independence of the random variables for the created and annihilated production degrees between consecutive update steps. The proof of the lemma is based on a law of large numbers argument. Such an argument is standard, of course. However since the proof of the lemma is short and we deal with a special notion of the bounded Lipschitz distance $d(\cdot,\cdot)$ (in equation (23)) we provide it in the following.

We first split the interval $[0,1]$ into n pieces $I_j := [\frac{j-1}{n}, \frac{j}{n}]$ with $1 \leq j \leq n$. Later on, n is chosen as a function of the total number of update steps K. The definition of $d(\cdot,\cdot)$ involves taking a supremum. Taking the supremum does not commute with taking the expectation value. Therefore, we first estimate the distance $d(\eta^{\text{aux}}_{N,K}, \eta^{\text{aux}}_{K})$ of the empirical density $\eta^{\text{aux}}_{N,K}$ from the average density η^{aux}_{K} and take the expectation value later. This distance is given as:

$$d(\eta^{\text{aux}}_{N,K}, \eta^{\text{aux}}_{K}) = \|\eta^{\text{aux}}_{N,K} - \eta^{\text{aux}}_{K}\|_{BL} , \tag{45}$$

$$= \sup_{\|f\|_L=1} \left| \int_0^1 \mathrm{d}p\ f(p)\eta^{\text{aux}}_{N,K}(p) - \int_0^1 \mathrm{d}p\ f(p)\eta^{\text{aux}}_{K}(p) \right| . \tag{46}$$

We now make use of the Lipschitz continuity of f on every interval I_j, that is:

$$f(p) \leq \left| f\left(\frac{j-1}{n}\right) \right| + \left| \frac{j-1}{n} - p \right| , \tag{47}$$

because $\|f\|_L = 1$ on every interval I_j.

Therefore, one estimates:

$$
\begin{aligned}
& d(\eta^{\mathrm{aux}}_{N,K}, \eta^{\mathrm{aux}}_{K}) \\
& \quad \leq \sum_{j=1}^{n} \sup_{\|f\|_L=1} \left| \int_{I_j} \mathrm{d}p \; f(p) \left(\eta^{\mathrm{aux}}_{N,K}(p) - \eta^{\mathrm{aux}}_{N,0}(p) - \eta^{\mathrm{aux}}_{K}(p) + \eta^{\mathrm{aux}}_{0}(p) \right) \right| \\
& \qquad + d(\eta^{\mathrm{aux}}_{N,0}, \eta^{\mathrm{aux}}_{0}) \,, \qquad (48) \\
& \quad \leq \sum_{j=1}^{n} \int_{I_j} \mathrm{d}p \; \left| f\left(\frac{j-1}{n} \right) \right| \left| \eta^{\mathrm{aux}}_{N,K}(p) - \eta^{\mathrm{aux}}_{N,0}(p) - \eta^{\mathrm{aux}}_{K}(p) + \eta^{\mathrm{aux}}_{0}(p) \right| \\
& \qquad + \sum_{j=1}^{n} \int_{I_j} \mathrm{d}p \; \left| \frac{j-1}{n} - p \right| \left| \eta^{\mathrm{aux}}_{N,K}(p) - \eta^{\mathrm{aux}}_{N,0}(p) - \eta^{\mathrm{aux}}_{K}(p) + \eta^{\mathrm{aux}}_{0}(p) \right| \\
& \qquad + d(\eta^{\mathrm{aux}}_{N,0}, \eta^{\mathrm{aux}}_{0}) \,, \qquad (49) \\
& \quad \leq Const \cdot \left(1 + \frac{1}{n} \right) \sum_{j=1}^{n} \int_{I_j} \mathrm{d}p \; \left| \eta^{\mathrm{aux}}_{N,K}(p) - \eta^{\mathrm{aux}}_{N,0}(p) - \eta^{\mathrm{aux}}_{K}(p) + \eta^{\mathrm{aux}}_{0}(p) \right| \\
& \qquad + d(\eta^{\mathrm{aux}}_{N,0}, \eta^{\mathrm{aux}}_{0}) \,. \qquad (50)
\end{aligned}
$$

It follows that:

$$
\begin{aligned}
\mathbb{E}(d(\eta^{\mathrm{aux}}_{N,K}, \eta^{\mathrm{aux}}_{K})) \leq Const \cdot \left(1 + \frac{1}{n} \right) \sum_{j=1}^{n} \mathbb{E} \left(\int_{I_j} \mathrm{d}p \; \left| (\eta^{\mathrm{aux}}_{N,K}(p) - \eta^{\mathrm{aux}}_{N,0}(p) - \eta^{\mathrm{aux}}_{K}(p) + \eta^{\mathrm{aux}}_{0}(p)) \right| \right) \\
+ \mathbb{E} \left(d(\eta^{\mathrm{aux}}_{N,0}, \eta^{\mathrm{aux}}_{0}) \right) \,. \qquad (51)
\end{aligned}
$$

For each interval I_j we now give a law of large numbers argument. We define the random variable $Y^{j,+}_{k,1}$ that takes value 1 if the individual $N + 2k - 1$ with a positive mass is sampled inside the interval I_j in the k^{th} update step, and that takes value 0 otherwise. Accordingly, the random variable $Y^{j,+}_{k,2}$ takes value 1 if individual $N + 2k$ is sampled inside the interval I_j. Furthermore, the random variables $Y^{j,-}_{k,1/2}$ indicate whether an individual is created with a

negative mass in the interval I_j at the k^{th} update step:

$$Y_{k,1}^{j,-}(\omega^k) := \begin{cases} 1 & \text{if } p_{2k-1}^{\text{aux},-}(\omega^k) \in I_j\,, \quad \text{that is, if } X^{\Phi(\eta_k^{\text{aux}})}(\omega_1^k) \in I_j\,, \\ 0 & \text{else}\,; \end{cases} \tag{52}$$

$$Y_{k,2}^{j,-}(\omega^k) := \begin{cases} 1 & \text{if } p_{2k}^{\text{aux},-}(\omega^k) \in I_j\,, \quad \text{that is, if } X^{\eta_k^{\text{aux}}}(\omega_2^k) \in I_j\,, \\ 0 & \text{else}\,; \end{cases} \tag{53}$$

$$Y_{k,1}^{j,+}(\omega^k) := \begin{cases} 1 & \text{if } p_{N+2k-1}^{\text{aux},+}(\omega^k) \in I_j\,, \\ 0 & \text{else}\,; \end{cases} \tag{54}$$

$$Y_{k,2}^{j,+}(\omega^k) := \begin{cases} 1 & \text{if } p_{N+2k}^{\text{aux},+}(\omega^k) \in I_j\,, \\ 0 & \text{else}\,. \end{cases} \tag{55}$$

By definition 8 of the auxiliary process, the $Y_k^{j,\pm}$ are independent for different values of k, that is consecutive updates with birth and death are independent. Therefore, the difference between positive and negative masses in the interval I_j after K update steps in one realization of the auxiliary process is obtained as:

$$\sum_{k=1}^{K} \left(Y_{k,1}^{j,+}(\omega^k) + Y_{k,2}^{j,+}(\omega^k) - Y_{k,1}^{j,-}(\omega^k) - Y_{k,2}^{j,-}(\omega^k) \right) = N \int_{I_j} \mathrm{d}p \, \left(\eta_{N,K}^{\text{aux}}(p) - \eta_{N,0}^{\text{aux}}(p) \right) . \tag{56}$$

By the definition of the average probability distribution η^{aux} of the auxiliary process, it is:

$$\mathbb{E}\left(\sum_{k=1}^{K} \left(Y_{k,1}^{j,+}(\omega^k) + Y_{k,1}^{j,+}(\omega^k) - Y_{k,1}^{j,-}(\omega^k) - Y_{k,1}^{j,-}(\omega^k) \right) \right) = N \int_{I_j} \mathrm{d}p \, \left(\eta_K^{\text{aux}}(p) - \eta_0^{\text{aux}}(p) \right) . \tag{57}$$

Using independence between the different update steps, we have a law of large numbers argument for every interval I_j and every sequence of random variables,

$$(Z_k)_{k\in\{1,\dots,K\}} \in \left\{ \left(Y_{k,1}^{j,+}\right)_{k\in\{1,\dots,K\}}, \left(Y_{k,2}^{j,+}\right)_{k\in\{1,\dots,K\}}, \left(Y_{k,1}^{j,-}\right)_{k\in\{1,\dots,K\}}, \left(Y_{k,2}^{j,-}\right)_{k\in\{1,\dots,K\}} \right\} , \tag{58}$$

as follows:

$$\mathbb{E}\left(\left| \frac{1}{N} \sum_{k=1}^{K} Z_k - \mathbb{E}\left(\frac{1}{N} \sum_{k=1}^{K} Z_k \right) \right| \right) \leq \left(\mathrm{Var}\left(\frac{1}{N} \sum_{k=1}^{K} Z_k \right) \right)^{1/2} \leq \frac{1}{N} \sqrt{K} \frac{1}{2} . \tag{59}$$

The last estimate exploits the independence of random variables between consecutive steps of the sampling process, and the boundedness of the variance with $\mathrm{Var}(Y_{k,1/2}^{j,\pm}) \leq 1/4$ for all $k = 1, \dots, K$.

Using triangle inequality and linearity of the expectation value we obtain:

$$\mathbb{E}\left(\left|\frac{1}{N}\sum_{k=1}^{K}\left(Y_{k,1}^{j,+}+Y_{k,2}^{j,+}-Y_{k,1}^{j,-}-Y_{k,2}^{j,-}\right)-\mathbb{E}\left(\frac{1}{N}\sum_{k=1}^{K}\left(Y_{k,1}^{j,+}+Y_{k,2}^{j,+}-Y_{k,1}^{j,-}-Y_{k,2}^{j,-}\right)\right)\right|\right)\leq\frac{2\sqrt{K}}{N}\,. \tag{60}$$

Therefore, we obtain with equation (51):

$$\mathbb{E}\left(d(\eta_{N,K}^{\text{aux}},\eta_{K}^{\text{aux}})\right)\leq Const\cdot\left(1+\frac{1}{n}\right)\sum_{j=1}^{n}\frac{2\sqrt{K}}{N}+\mathbb{E}\left(d(\eta_{N,0}^{\text{aux}},\eta_{0}^{\text{aux}})\right)\,. \tag{61}$$

Choosing $n=K^{1/4}$ yields the estimate:

$$\mathbb{E}\left(d(\eta_{N,K}^{\text{aux}},\eta_{K}^{\text{aux}})\right)\leq Const\cdot\frac{K^{3/4}}{N}+\mathbb{E}\left(d(\eta_{N,0}^{\text{aux}},\eta_{0}^{\text{aux}})\right)\,, \tag{62}$$

which proves the lemma. □

C.3.2 Convergence of the microscopic to the auxiliary process $\left(\rho_{N,K}^{(1)}\xrightarrow[N\to\infty]{\text{in prob}}\eta_{N,K}^{\text{aux}}\right)$ – Control of error propagation with Grönwall's inequality

We now show that the propagation of errors, which build up over time due to the correlation of production degrees, can be controlled with Grönwall's inequality. In other words, the empirical density of the microscopic process $\rho_{N,K}^{(1)}$ converges to the empirical density of the auxiliary process $\eta_{N,K}^{\text{aux}}$, as $N\to\infty$ for any finite update step K, see lemma 4.

Lemma 2. *Let $\nu\in(L^\infty)^\star$ and $\psi\in(L^\infty)^\star$ be two one-particle probability densities, f some globally Lipschitz continuous function on* $[0,1]$. *Then:*

$$\left|\langle f\rangle_\nu-\langle f\rangle_\psi\right|\leq\|f\|_L d(\nu,\psi)\,. \tag{63}$$

Here $\langle\cdot\rangle_\nu$ and $\langle\cdot\rangle_\psi$ means averaging with respect to ν and ψ, respectively, and $\|f\|_L$ is the global Lipschitz constant of f.

Proof. Plugging in the definitions, one obtains:

$$\left|\langle f\rangle_\nu-\langle f\rangle_\psi\right|=\left|\int_0^1 \mathrm{d}p\; f(p)\left(\nu(p)-\psi(p)\right)\right|\,, \tag{64}$$

$$=\|f\|_L\left|\int_0^1 \mathrm{d}p\;\frac{f(p)}{\|f\|_L}\left(\nu(p)-\psi(p)\right)\right|\,. \tag{65}$$

Since $\left\|\frac{f(p)}{\|f\|_L}\right\|_L=1$ we can use it to test the supremum in the definition of the bounded Lipschitz distance $d(\cdot,\cdot)$ and obtain that the right hand side of the last equation is indeed bounded by $\|f\|_L d(\nu,\psi)$. □

Lemma 3. *Let $\nu \in (L^\infty)^\star$ and $\psi \in (L^\infty)^\star$ be two one particle probability densities. Then:*

$$\mathbb{E}(|X^\nu - X^\psi|) \leq d(\nu, \psi)\,, \tag{66}$$

see definition 4 of the quantile function X^ν and X^ψ.

Proof. Laisant's formula for inverse functions states that for any invertible function g, it is:

$$\int_a^b \mathrm{d}y\; g^{-1}(y) + \int_c^d \mathrm{d}x\; g(x) = bd - ac\;. \tag{67}$$

Applying Laisant's formula to the random variables X^ν and X^ψ, it follows that:

$$\mathbb{E}\left(\left|X^\nu - X^\psi\right|\right) = \int_0^1 \mathrm{d}\omega\;\left|X^\nu(\omega) - X^\psi(\omega)\right| = \int_0^1 \mathrm{d}p\;\left|\int_0^p \mathrm{d}y\;(\nu(y) - \psi(y))\right|\,, \tag{68}$$

see figure 6(A) for a sketch.

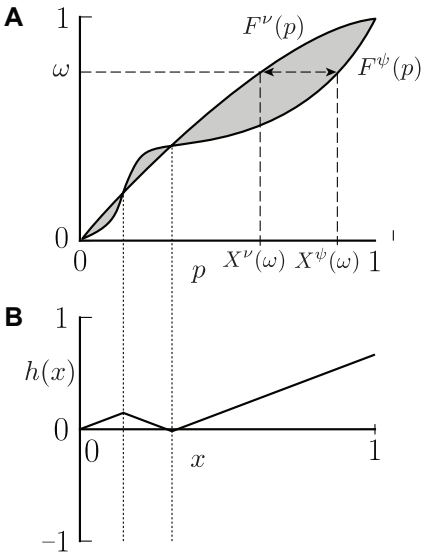


Figure 6. (A) Illustration of the application of Laisant's formula to the expected difference of two random variables, see equation (68). The expected difference of the two random variables X^ν and X^ψ is obtained as the area enclosed by the two curves $X^\nu(\omega)$ and $X^\psi(\omega)$ for $\omega \in [0, 1]$. Thus, the area is given by $\int_0^1 \mathrm{d}\omega\;\left|X^\nu(\omega) - X^\psi(\omega)\right|$. On the other hand, the area between the two curves can be computed from the inverse functions to X^ν and X^ψ, which are the cumulative distribution functions $F^\nu(p) = \int_0^p \mathrm{d}y\;\nu(y)$ and $F^\psi(p) = \int_0^p \mathrm{d}y\;\psi(y)$, respectively, see definition 4. Therefore, the area is also given by $\int_0^1 \mathrm{d}p\;\left|F^\nu(p) - F^\psi(p)\right|$. The rigorous argument follows with Laisant's formula. (B) Definition of the function h that is used to estimate $d(\nu, \psi)$. $h'(x) = 1$ if $F^\nu(p) > F^\psi(p)$ and $h'(x) = -1$ if $F^\nu(p) < F^\psi(p)$, and thus $\|h\|_L \leq 1$.

Now, let $h : [0,1] \to \mathbb{R}$ be given by $h(x) := \int_0^x \mathrm{d}p \ \mathrm{sgn}\left(\int_0^p \mathrm{d}y \ \nu(y) - \psi(y)\right)$ (here sgn is the signum function), that is $h'(x) = 1$ if $\int_0^p \mathrm{d}y \ \nu(y) > \int_0^p \mathrm{d}y \ \psi(y)$ and $h'(x) = -1$ if $\int_0^p \mathrm{d}y \ \nu(y) < \int_0^p \mathrm{d}y \ \psi(y)$, see figure 6(B) for a sketch. In particular, it is $\|h\|_L \leq 1$. Therefore, one may use h to estimate the supremum in the definition of the bounded Lipschitz metric $d(\cdot,\cdot)$ as:

$$d(\nu,\psi) \geq \int_0^1 \mathrm{d}p \ h(p)(\nu(p) - \psi(p)) \,. \tag{69}$$

Integration by parts yields:

$$d(\nu,\psi) \geq \int_0^1 \mathrm{d}p \ h'(p) \int_0^p \mathrm{d}y \ (\nu(y) - \psi(y)) \,, \tag{70}$$

$$= \int_0^1 \mathrm{d}p \ \left| \int_0^p \mathrm{d}y \ (\nu(y) - \psi(y)) \right| \,. \tag{71}$$

Since both ν and ψ are normalized to 1, the boundary termsarising in the integration by parts above vanish.

Together with equation (68) from above, one obtains the estimate of the lemma:

$$\mathbb{E}(|X^\nu - X^\psi|) \leq d(\nu,\psi) \,. \tag{72}$$

□

Lemma 4. *There exists a constant* $0 < Const < \infty$ *such that for a given update step K one estimates:*

$$\left|\mathbb{E}\left(d(\rho^{(1)}_{N,K}, \eta^{aux}_{N,K})\right)\right| \leq e^{Const \cdot \frac{K}{N}} \left(\mathbb{E}\left(d(\rho^{(1)}_{N,0}, \eta^{aux}_{N,0})\right) + \frac{K^{3/4}}{N}\right) \,. \tag{73}$$

Proof. The proof of the lemma is based on a discrete Grönwall's inequality.

We first estimate how the distance between a realization of the microscopic process $(\rho^{(1)}_{N,k})$ and a realization of the auxiliary process $(\eta^{\mathrm{aux}}_{N,k})$ propagates from at a certain update step k to step $k+1$. This distance measures the error that occurs upon neglecting correlations of the individuals' production degrees. This error propagates on average from one update step k to the next update step $k+1$ as follows:

$$\left|\mathbb{E}\left(d(\rho^{(1)}_{N,k+1}, \eta^{\mathrm{aux}}_{N,k+1})\right) - \mathbb{E}\left(d(\rho^{(1)}_{N,k}, \eta^{\mathrm{aux}}_{N,k})\right)\right| \leq \frac{Const}{N} \mathbb{E}\left(\left|X^{\rho^{(1)}_{N,k}} - X^{\eta^{\mathrm{aux}}_k}\right|\right) \,. \tag{74}$$

To see this estimate, we write:

$$
\begin{aligned}
&d(\rho^{(1)}_{N,k+1}, \eta^{\text{aux}}_{N,k+1}) \\
&\quad = \sup_{\|f\|_L=1} \left| \int_0^1 \mathrm{d}p\, f(p) \rho^{(1)}_{N,k+1}(p) - \int_0^1 \mathrm{d}p\, f(p) \eta^{\text{aux}}_{N,k+1}(p) \right| , \qquad (75) \\
&\quad \leq \sup_{\|f\|_L=1} \left| \int_0^1 \mathrm{d}p\, f(p) \left(\rho^{(1)}_{N,k}(p) - \eta^{\text{aux}}_{N,k}(p) \right) \right| \qquad (76) \\
&\qquad + \frac{Const}{N} \sup_{\|f\|_L=1} \int_0^1 \mathrm{d}p\, f(p) \int_0^1 \mathrm{d}\omega^{k+1} \left| \text{“realization (micro)” – “realization (aux)”} \right| .
\end{aligned}
$$

The latter estimate follows because the distance of the densities between consecutive update steps involves the change of at most four production degrees in the population and, thus, a change of probability mass of order $\mathcal{O}(1/N)$ from $\rho^{(1)}_{N,k}$ to $\rho^{(1)}_{N,k+1}$ and from $\eta^{\text{aux}}_{N,k}$ to $\eta^{\text{aux}}_{N,k+1}$. We further estimate:

$$
\begin{aligned}
&d(\rho^{(1)}_{N,k+1}, \eta^{\text{aux}}_{N,k+1}) \\
&\leq d(\rho^{(1)}_{N,k}, \eta^{\text{aux}}_{N,k}) \\
&\quad + \frac{Const}{N} \sup_{\|f\|_L=1} \int_0^1 \mathrm{d}p\, f(p) \int_0^1 \mathrm{d}\omega^{k+1} \left| X^{\rho^{(1)}_{N,k}}(\omega_1^{k+1}) - X^{\Phi(\eta^{\text{aux}}_{k+1})}(\omega_1^{k+1}) \right| \\
&\quad + \frac{Const}{N} \sup_{\|f\|_L=1} \int_0^1 \mathrm{d}p\, f(p) \int_0^1 \mathrm{d}\omega^{k+1} \left| X^{\rho^{(1)}_{N,k} - \delta(p - p^{\text{micro},-}_{2k+1}(\omega_1^{k+1}))}(\omega_2^{k+1}) - X^{\eta^{\text{aux}}_{k+1}}(\omega_2^{k+1}) \right| \\
&\quad + \frac{Const}{N} \sup_{\|f\|_L=1} \int_0^1 \mathrm{d}p\, f(p) \int_0^1 \mathrm{d}\omega^{k+1} \left| \omega_3^{k+1} \left(X^{\rho^{(1)}_{N,k}}(\omega_1^{k+1}) - X^{\Phi(\eta^{\text{aux}}_{k+1})}(\omega_1^{k+1}) \right) \right. \\
&\qquad\qquad \left. + (1 - \omega_3^{k+1}) \left(R(\langle p \rangle_{\rho^{(1)}_{N,k}}) - R(\langle p \rangle_{\eta^{\text{aux}}_{k+1}}) \right) \right| \\
&\quad + \frac{Const}{N} \sup_{\|f\|_L=1} \int_0^1 \mathrm{d}p\, f(p) \int_0^1 \mathrm{d}\omega^{k+1} \left| \omega_4^{k+1} \left(X^{\rho^{(1)}_{N,k}}(\omega_1^{k+1}) - X^{\Phi(\eta^{\text{aux}}_{k+1})}(\omega_1^{k+1}) \right) \right. \\
&\qquad\qquad \left. + (1 - \omega_4^{k+1}) \left(R(\langle p \rangle_{\rho^{(1)}_{N,k}}) - R(\langle p \rangle_{\eta^{\text{aux}}_{k+1}}) \right) \right| .
\end{aligned} \qquad (77)
$$

This estimate can be further summarized and estimated as follows:

$$
\begin{aligned}
&d(\rho^{(1)}_{N,k+1}, \eta^{\text{aux}}_{N,k+1}) \\
&\leq d(\rho^{(1)}_{N,k}, \eta^{\text{aux}}_{N,k}) + \frac{Const}{N}\mathbb{E}\left(\left|X^{\rho^{(1)}_{N,k}} - X^{\eta^{\text{aux}}_{k+1}}\right|\right) + \frac{Const}{N}\mathbb{E}\left(\left|R(\langle p\rangle_{\rho^{(1)}_{N,k}}) - R(\langle p\rangle_{\eta^{\text{aux}}_{k+1}})\right|\right), \qquad (78)\\
&\leq d(\rho^{(1)}_{N,k}, \eta^{\text{aux}}_{N,k}) + \frac{Const}{N}\mathbb{E}\left(\left|X^{\rho^{(1)}_{N,k}} - X^{\eta^{\text{aux}}_{k}}\right|\right). \qquad (79)
\end{aligned}
$$

The last estimate follows with the triangle inequality, the definition of the average auxiliary density (8), and lemma 3. By applying lemma 2 to the last line above, estimate (74) follows.

We now determine how the growth of the average error from update step k to $k+1$ depends upon the error at step k. By applying lemma 3 to the estimate (74), which involves the average density of the auxiliary process and not the empirical density, one obtains (also note the different meanings of the expectation values taken above and below):

$$
\begin{aligned}
&\left|\mathbb{E}\left(d(\rho^{(1)}_{N,k+1}, \eta^{\text{aux}}_{N,k+1})\right) - \mathbb{E}\left(d(\rho^{(1)}_{N,k}, \eta^{\text{aux}}_{N,k})\right)\right| \\
&\qquad \leq \frac{Const}{N}\mathbb{E}\left(d(\rho^{(1)}_{N,k}, \eta^{\text{aux}}_{k})\right), \qquad (80)\\
&\qquad \leq \frac{Const}{N}\mathbb{E}\left(d(\rho^{(1)}_{N,k}, \eta^{\text{aux}}_{N,k})\right) + \frac{Const}{N}\mathbb{E}\left(d(\eta^{\text{aux}}_{N,k}, \eta^{\text{aux}}_{k})\right), \qquad (81)\\
&\qquad \leq \frac{Const}{N}\mathbb{E}\left(d(\rho^{(1)}_{N,k}, \eta^{\text{aux}}_{N,k})\right) + \frac{Const}{N}\left(\frac{k^{3/4}}{N} + \mathbb{E}\left(d(\eta^{\text{aux}}_{N,0}, \eta^{\text{aux}}_{0})\right)\right). \qquad (82)
\end{aligned}
$$

The estimate in the first line above follows with the triangle inequality, and the second estimate follows with the law of large numbers argument from lemma 1. Essentially, the growth of the average error from update step k to $k+1$ can be attributed to the following sources: (i) propagation of errors from update step k, (ii) creation of "new" errors at the update step $k+1$ because of the law of large numbers, and (iii) propagation of initial errors.

From the growth of errors between two consecutive update steps, the growth of errors for any given finite number of update steps K can be controlled with Grönwall's inequality as we show next. Grönwall's inequality for differentiable functions u states that if $u'(t)$ is bounded by $u'(t) \leq \alpha u(t) + \alpha\beta$ with $\alpha, \beta \in \mathbb{R}$, then it follows that $u(t)$ is bounded by the solution of the differential equation given of the right-hand side ($u'(t) = \alpha u(t) + \alpha\beta$) as $u(t) \leq u(0)e^{\alpha t} + \beta(e^{\alpha t} - 1)$. In the spirit of a discrete version of Grönwall's inequality applied to the estimate (82), one finds a constant $0 < Const < \infty$ such that for a given update step K:

$$
\left|\mathbb{E}\left(d(\rho^{(1)}_{N,K}, \eta^{\text{aux}}_{N,K})\right)\right| \leq e^{Const\cdot\frac{K}{N}}\left(\mathbb{E}\left(d(\rho^{(1)}_{N,0}, \eta^{\text{aux}}_{N,0})\right) + \frac{K^{3/4}}{N}\right), \qquad (83)
$$

which concludes the proof of lemma 4. □

C.3.3 Convergence of the macroscopic to the auxiliary process $\left(\eta^{\text{aux}}_{\kappa(t)} \xrightarrow[N\to\infty]{\text{in prob}} \rho_t\right)$ – Continuous time limit and control of time synchronization

We now show that the mean-field density of the macroscopic process (ρ_t) converges in probability to the average density of the auxiliary process $(\eta^{\text{aux}}_{\kappa(t)})$ as $N \to \infty$, see lemma 7.

Definition 9. *Let $\nu \in (L^\infty)^\star$ be a time-dependent density function. We define the time evolution operator:*

$$T_t(\nu) := \int_0^t \mathrm{d}t' \, \big[2\lambda(1 + s(b-c)\langle p\rangle_{\nu_{t'}})\big(\delta(p - R(\langle p\rangle_{\nu_{t'}})) - \nu(p,t')\big) + (1-2\lambda)sc\big(\langle p\rangle_{\nu_{t'}} - p\big)\nu(p,t')\big] \,, \tag{84}$$

with $\langle p\rangle_{\nu_{t'}} = \int_0^1 \mathrm{d}p \; p \; \nu(p,t')$.

With this definition, the time evolution of the production distribution, that is, the probability density of the macroscopic process, is given by (see definition (4) of the fitness and the mean-field equation (1)):

$$\rho(p,t) = T_t(\rho)(p) + \rho_0(p) \,. \tag{85}$$

Lemma 5. *Let $\nu \in (L^\infty)^\star$ and $\psi \in (L^\infty)^\star$ be one-particle probability densities. Then:*

$$d\left(T_t(\nu), T_t(\psi)\right) \le Const \cdot \int_0^t \mathrm{d}t' \; d(\nu_{t'}, \psi_{t'}) \,. \tag{86}$$

Proof. First, we rewrite $d\left(T_t(\nu), T_t(\psi)\right)$ as follows:

$$d\left(T_t(\nu), T_t(\psi)\right) = \|T_t(\nu) - T_t(\psi)\|_{BL} \tag{87}$$

$$= \Big\| \int_0^t \mathrm{d}t' \, \big[2\lambda(1 + s(b-c)\langle p\rangle_{\nu_{t'}})\big(\delta(p - R(\langle p\rangle_{\nu_{t'}}) - \nu(p,t')\big) + (1-2\lambda)sc\big(\langle p\rangle_{\nu_{t'}} - p\big)\nu(p,t') - 2\lambda(1 + s(b-c)\langle p\rangle_{\psi_{t'}})\big(\delta(p - R(\langle p\rangle_{\psi_{t'}}) - \psi(p,t')\big) + (1-2\lambda)sc\big(\langle p\rangle_{\psi_{t'}} - p\big)\psi(p,t')\big] \Big\|_{BL} \,, \tag{88}$$

$$= \Big\| \int_0^t \mathrm{d}t' \, \big[2\lambda(1 + s(b-c)\langle p\rangle_{\nu_{t'}})\big(\delta(p - R(\langle p\rangle_{\nu_{t'}}) - \nu(p,t')\big) + (1-2\lambda)sc\big(\langle p\rangle_{\nu_{t'}} - p\big)\nu(p,t') - 2\lambda(1 + s(b-c)\langle p\rangle_{\psi_{t'}})\big(\delta(p - R(\langle p\rangle_{\psi_{t'}}) - \psi(p,t')\big) + (1-2\lambda)sc\big(\langle p\rangle_{\psi_{t'}} - p\big)\psi(p,t') + 2\lambda(1 + s(b-c)\langle p\rangle_{\nu_{t'}})\big(\delta(p - R(\langle p\rangle_{\nu_{t'}}) - \psi(p,t')\big) + (1-2\lambda)sc\big(\langle p\rangle_{\nu_{t'}} - p\big)\psi(p,t') - 2\lambda(1 + s(b-c)\langle p\rangle_{\nu_{t'}})\big(\delta(p - R(\langle p\rangle_{\nu_{t'}}) - \psi(p,t')\big) + (1-2\lambda)sc\big(\langle p\rangle_{\nu_{t'}} - p\big)\psi(p,t')\big] \Big\|_{BL} \,. \tag{89}$$

We now estimate with triangle inequality from the last line above:

$$\begin{aligned}
&d\left(T_t(\nu), T_t(\psi)\right) \\
&\leq \Big\| \int_0^t \mathrm{d}t' \, \big[2\lambda(1+s(b-c)\langle p\rangle_{\nu_{t'}})\big(\psi(p,t')-\nu(p,t')\big) \\
&\qquad\qquad + (1-2\lambda)sc\big(\langle p\rangle_{\nu_{t'}} - p\big)\big(\nu(p,t')-\psi(p,t')\big)\big]\Big\|_{BL} \\
&\quad + \Big\| \int_0^t \mathrm{d}t' \, \big[2\lambda(1+s(b-c)\langle p\rangle_{\nu_{t'}})\delta(p-R(\langle p\rangle_{\nu_{t'}}) \\
&\qquad\qquad - 2\lambda(1+s(b-c)\langle p\rangle_{\psi_{t'}})\delta(p-R(\langle p\rangle_{\psi_{t'}}) \\
&\qquad\qquad + s(c-2\lambda b)\big(\langle p\rangle_{\nu_{t'}} - \langle p\rangle_{\psi_{t'}}\big)\psi(p,t')\big]\Big\|_{BL} \, . \qquad (90)
\end{aligned}$$

Since $c, b, s, \lambda, \langle p\rangle_{\nu_{t'}}, \langle p\rangle_{\psi_{t'}}, \|\nu\|_{BL}$ and $\|\psi\|_{BL}$ are uniformly bounded, it follows that there exists a constant $0 < Const < \infty$ such that:

$$d\left(T_t(\nu), T_t(\psi)\right) \leq Const \cdot \int_0^t \mathrm{d}t' \, \left[\left\|\psi(p,t')-\nu(p,t')\right\|_{BL} + \left|\langle p\rangle_{\nu_{t'}} - \langle p\rangle_{\psi_{t'}}\right|\right] . \qquad (91)$$

Using lemma 2 it follows that there exists (another) constant $0 < Const < \infty$ such that:

$$d\left(T_t(\nu), T_t(\psi)\right) \leq Const \cdot \int_0^t \mathrm{d}t' \, d(\nu_{t'}, \psi_{t'}) \, . \qquad (92)$$

□

To prepare the continuous time limit, we show that the average time τ_k, at which updates of the population occur, stays on average close to the real time. The proof proceeds by applying a law of large numbers argument. Recall from definition 2 that we denoted the random variable for the length of the time interval between the k^{th} and $(k+1)^{\text{th}}$ update step as $\Delta\tau_k$ with $\langle\Delta\tau_k\rangle = N^{-1}\left(1+s(b-c)\langle p\rangle_k\right)^{-1}$. We also defined $\tau_k = \sum_{l=1}^{k-1}\Delta\tau_l$ and $\langle\tau_k\rangle = \sum_{l=1}^{k-1}\langle\Delta\tau_l\rangle$; $M(t)$ denotes the maximal natural number such that $\langle\tau_{M(t)}\rangle \leq t$ and $\kappa(t)$ is the random variable given by the maximal number such that $\tau_{\kappa(t)} \leq t$.

Lemma 6. *Let $t > 0$. Then:*

$$\mathbb{P}(|\kappa(t) - M(t)| \geq N^{3/4}) \leq Const \cdot N^{-1/2} \, . \qquad (93)$$

Proof. The lemma is based on the law of large numbers. Since the $\kappa(t)$ is monotonously increasing, it follows that:

$$\begin{aligned}
\mathbb{P}(\kappa(t) < M(t) - N^{3/4}) &\leq \mathbb{P}(\tau_{M(t)-N^{3/4}} > t) \, , &(94)\\
&= \mathbb{P}(\tau_{M(t)-N^{3/4}} - \langle\tau_{M(T)-N^{3/4}}\rangle > t - \langle\tau_{M(t)-N^{3/4}}\rangle) \, , &(95)\\
&\leq \mathbb{P}\left(\left|\tau_{M(t)-N^{3/4}} - \langle\tau_{M(t)-N^{3/4}}\rangle\right| > \left|t - \langle\tau_{M(t)-N^{3/4}}\rangle\right|\right) \, . &(96)
\end{aligned}$$

Because the τ_k are independent of each other, it follows with Chebyshev's inequality that for any (possibly N-dependent) $\epsilon_N > 0$:

$$\mathbb{P}\left(\left|\tau_{M(t)-N^{3/4}} - \langle \tau_{M(t)-N^{3/4}}\rangle\right| > \epsilon_N\right) \leq \epsilon_N^{-2}\mathrm{Var}(\tau_{M(t)-N^{3/4}})\,, \tag{97}$$

$$= \epsilon_N^{-2} \sum_{k=1}^{M(t)-N^{3/4}} \mathrm{Var}\Delta\tau_k\,, \tag{98}$$

$$\leq Const \cdot \frac{N}{\epsilon_N^2 N^2}\,. \tag{99}$$

Since the average lengths of time intervals between two update steps, $\langle\Delta\tau_k\rangle$, are bounded for all k by some constant times N^{-1}, the respective variances are of order N^{-2}. The estimate in the last line above then follows by recalling that $M(t) \sim \mathcal{O}(N)$.

We choose $\epsilon_N := \left|t - \langle\tau_{M(t)-N^{3/4}}\rangle\right|$ and estimate:

$$\epsilon_N = \left|\langle\tau_{M(t)-N^{3/4}}\rangle - t\right|\,, \tag{100}$$

$$= \left|\sum_{k=1}^{M(t)-N^{3/4}} \langle\Delta\tau_k\rangle - t\right|\,, \tag{101}$$

$$\leq \left|\sum_{k=1}^{M(t)-N^{3/4}} \langle\Delta\tau_k\rangle - \sum_{k=1}^{M(t)} \langle\Delta\tau_k\rangle\right| + \frac{Const}{N}\,, \tag{102}$$

$$\leq N^{3/4}\frac{Const}{N} = Const \cdot N^{-1/4}\,. \tag{103}$$

Therefore, one obtains from Chebyshev's inequality with the chosen ϵ_N:

$$\mathbb{P}\left(\left|\tau_{M(t)-N^{3/4}} - \langle \tau_{M(t)-N^{3/4}}\rangle\right| > \epsilon_N\right) \leq Const \cdot N^{-1/2}\,. \tag{104}$$

From equation (96) one obtains:

$$\mathbb{P}(\kappa(t) < M(t) - N^{3/4}) \leq CN^{-1/2}\,. \tag{105}$$

In the same way one shows that:

$$\mathbb{P}(\kappa(t) > M(t) + N^{3/4}) \leq CN^{-1/2}\,, \tag{106}$$

and the lemma follows. □

After these preparatory steps, we now proceed with the following lemma:

Lemma 7. *For any $t > 0$, one estimates:*

$$\mathbb{E}\left(d(\eta^{aux}_{\kappa(t)}, \rho_t)\right) \leq Const(t) \cdot N^{-1/4}\,. \tag{107}$$

Proof. We apply the triangle inequality and estimate:

$$\mathbb{E}\left(d(\eta^{\text{aux}}_{\kappa(t)}, \rho_t)\right) \leq \mathbb{E}\left(d(\eta^{\text{aux}}_{\kappa(t)}, \eta^{\text{aux}}_{M(t)})\right) + d(\eta^{\text{aux}}_{M(t)}, \rho_t)\,. \tag{108}$$

Note that the expectation values above are taken with respect to sampling the update times. The first summand (i) addresses the distance of the average auxiliary distribution between different update steps; namely between a single realization of update steps ($\kappa(t)$) up to the given time t and the average number of update steps ($M(t)$) up to time t. By a law of large numbers argument, we show below that:

$$\mathbb{E}\left(d(\eta^{\text{aux}}_{\kappa(t)}, \eta^{\text{aux}}_{M(t)})\right) \leq Const \cdot N^{-1/4}\,, \tag{109}$$

The second summand (ii) governs the distance of the auxiliary process at average times to the macroscopic process at the real time t. We show below that the propagation of errors due to different timings of the auxiliary process and the macroscopic process are controlled by applying Grönwall's inequality, and estimate:

$$d(\eta^{\text{aux}}_{M(t)}, \rho_t) \leq \frac{1}{N} e^{Const \cdot t}\,. \tag{110}$$

(i) First, we estimate $\mathbb{E}\left(d(\eta^{\text{aux}}_{\kappa(t)}, \eta^{\text{aux}}_{M(t)})\right)$ in equation (109) by splitting up the expectation value as follows:

$$\begin{aligned}
&\mathbb{E}\left(d(\eta^{\text{aux}}_{\kappa(t)}, \eta^{\text{aux}}_{M(t)})\right) \\
&\quad\leq \sup\left\{\|\eta^{\text{aux}}_{\kappa(t)} - \eta^{\text{aux}}_{M(t)}\|_{BL} : |\kappa(t) - M(t)| \geq N^{3/4}\right\} \mathbb{P}(|\kappa(t) - M(t)| \geq N^{3/4}) \\
&\qquad + \sup\left\{\|\eta^{\text{aux}}_{\kappa(t)} - \eta^{\text{aux}}_{M(t)}\|_{BL} : |\kappa(t) - M(t)| \leq N^{3/4}\right\} \mathbb{P}(|\kappa(t) - M(t)| \leq N^{3/4})\,, \qquad (111) \\
&\quad\leq \sup\left\{\|\eta^{\text{aux}}_{\kappa(t)} - \eta^{\text{aux}}_{M(t)}\|_{BL}\right\} \mathbb{P}(|\kappa(t) - M(t)| \geq N^{3/4}) \\
&\qquad + \sup\left\{\|\eta^{\text{aux}}_{\kappa(t)} - \eta^{\text{aux}}_{M(t)}\|_{BL} : |\kappa(t) - M(t)| \leq N^{3/4}\right\}\,. \qquad (112)
\end{aligned}$$

Using the fact that for any probability densities ν and ψ it is $d_{BL}(\nu, \psi) \leq \|\nu\|_1 + \|\psi\|_1 = 2$, one estimates:

$$\sup\left\{\|\eta^{\text{aux}}_{\kappa(t)} - \eta^{\text{aux}}_{M(t)}\|_{BL}\right\} \leq 2\,. \tag{113}$$

From lemma 6, we obtain:

$$\mathbb{P}(|\kappa(t) - M(t)| \geq N^{3/4}) \leq Const \cdot N^{-1/2}\,. \tag{114}$$

Since $\|\eta^{\text{aux}}_k\|_{BL}$ and $\|\Phi(\eta^{\text{aux}}_k)\|_{BL}$ are bounded, it follows that:

$$\|\eta^{\text{aux}}_{k+1} - \eta^{\text{aux}}_k\|_{BL} = N^{-1}\|-\eta^{\text{aux}}_k + \Phi(\eta^{\text{aux}}_k)\|_{BL} \leq Const/N\,. \tag{115}$$

Therefore, one obtains:

$$\sup\left\{\|\eta^{\text{aux}}_{\kappa(t)} - \eta^{\text{aux}}_{M(t)}\|_{BL} : |\kappa(t) - M(t)| \leq N^{3/4}\right\} \leq Const \cdot N^{3/4} N^{-1} = Const \cdot N^{-1/4}\,, \tag{116}$$

and the estimate in equation (109) follows as:

$$\mathbb{E}\left(d(\eta^{\text{aux}}_{\kappa(t)}, \eta^{\text{aux}}_{M(t)})\right) \leq Const \cdot N^{-1/4}\,. \tag{117}$$

(ii) Second, we show the estimate (110) for $d(\eta^{\text{aux}}_{M(t)}, \rho_t)$. Recall that by definition 5, we have:

$$\Phi(\eta^{\text{aux}}_k)(p) = 2\lambda\delta\left(p - R(\langle p\rangle_{\eta^{\text{aux}}_k})\right) + (1-2\lambda)\frac{1 + s\left(b\langle p\rangle_{\eta^{\text{aux}}_k} - cp\right)}{1 + s(b-c)\langle p\rangle_{\eta^{\text{aux}}_k}}\eta^{\text{aux}}_k(p)\,. \tag{118}$$

We write the average density of the auxiliary process at update step k by applying definition 8 iteratively:

$$\begin{aligned}
&\eta^{\text{aux}}_k(p)\\
&= \eta^{\text{aux}}_{k-1}(p) - \frac{1}{N}\eta^{\text{aux}}_{k-1}(p) + \frac{1}{N}\Phi(\eta^{\text{aux}}_{k-1})(p)\,, && (119)\\
&= \eta^{\text{aux}}_0(p) + \frac{1}{N}\sum_{j=0}^{k-1}\left[\Phi(\eta^{\text{aux}}_j)(p) - \eta^{\text{aux}}_j(p)\right], && (120)\\
&= \eta^{\text{aux}}_0(p) + \sum_{j=0}^{k-1}\left[\frac{2\lambda}{N}\delta\left(p - R(\langle p\rangle_{\eta^{\text{aux}}_j})\right) + \frac{1-2\lambda}{N}\frac{1 + s\left(b\langle p\rangle_{\eta^{\text{aux}}_j} - cp\right)}{1 + s(b-c)\langle p\rangle_{\eta^{\text{aux}}_j}}\eta^{\text{aux}}_j(p) - \frac{1}{N}\eta^{\text{aux}}_j(p)\right], && (121)\\
&= \eta^{\text{aux}}_0(p) + \sum_{j=0}^{k-1} N^{-1}\left(1 + s(b-c)\langle p\rangle_{\eta^{\text{aux}}_j}\right)^{-1}\left[2\lambda\left(1 + s(b-c)\langle p\rangle_{\eta^{\text{aux}}_j}\right)\left(\delta(p - R(\langle p\rangle_{\eta^{\text{aux}}_j}) - \eta^{\text{aux}}_j\right)\right.\\
&\qquad\qquad \left. + (1-2\lambda)sc\left(\langle p\rangle_{\eta^{\text{aux}}_j} - p\right)\eta^{\text{aux}}_j(p)\right], && (122)\\
&= \eta^{\text{aux}}_0(p) + \int_0^{\langle\tau_k\rangle} \mathrm{d}t'\,\left[2\lambda(1 + s(b-c)\langle p\rangle_{\eta^{\text{aux}}_{M(t')}})\left(\delta\left(p - R(\langle p\rangle_{\eta^{\text{aux}}_{M(t')}})\right) - \eta^{\text{aux}}_{M(t')}(p)\right)\right.\\
&\qquad\qquad \left. + (1-2\lambda)sc\left(\langle p\rangle_{\eta^{\text{aux}}_{M(t')}} - p\right))\eta^{\text{aux}}_{M(t')}(p)\right], && (123)
\end{aligned}$$

where, in the last line, it was exploited that the j^{th} update step occurs after an average time $\langle\Delta\tau_j\rangle = N^{-1}\left(1 + s(b-c)\langle p\rangle_{\eta^{\text{aux}}_j}\right)^{-1}$. Furthermore, the L^1-norm of the above integrand is

bounded. Thus, one finds a constant $0 < Const < \infty$ such that:

$$\begin{aligned}
&\|\eta^{\mathrm{aux}}_{M(t)} - T_t(\eta^{\mathrm{aux}}_{M(t)}) - \eta^{\mathrm{aux}}_0\|_{BL} \\
&\quad \leq \Big\| \int_{\langle\tau_k\rangle}^{t} \mathrm{d}t' \, \Big[2\lambda(1 + s(b-c)\langle p\rangle_{\eta^{\mathrm{aux}}_{M(t')}}) \Big(\delta\big(p - R(\langle p\rangle_{\eta^{\mathrm{aux}}_{M(t')}})\big) - \eta^{\mathrm{aux}}_{M(t')}(p)\Big) \qquad (124) \\
&\qquad\qquad + (1-2\lambda)sc\big(\langle p\rangle_{\eta^{\mathrm{aux}}_{M(t')}} - p\big)\big)\eta^{\mathrm{aux}}_{M(t')}(p)\Big]\Big\|_{BL} \,, \\
&\quad \leq Const \cdot (t - \langle\tau_k\rangle) \,, \qquad (125) \\
&\quad \leq Const/N \,. \qquad (126)
\end{aligned}$$

Therefore, one estimates with the triangle inequality:

$$\begin{aligned}
&\|\eta^{\mathrm{aux}}_{M(t)} - \rho_t\|_{BL} \\
&\quad = \|\eta^{\mathrm{aux}}_{M(t)} - T_t(\eta^{\mathrm{aux}}_{M(t)}) - \eta^{\mathrm{aux}}_0 + T_t(\eta^{\mathrm{aux}}_{M(t)}) + \eta^{\mathrm{aux}}_0 - \rho_t\|_{BL} \,, \qquad (127) \\
&\quad \leq \|\eta^{\mathrm{aux}}_{M(t)} - T_t(\eta^{\mathrm{aux}}_{M(t)}) - \eta^{\mathrm{aux}}_0\|_{BL} + \|T_t(\eta^{\mathrm{aux}}_{M(t)}) + \eta^{\mathrm{aux}}_0 - T_t(\rho) - \rho_0\|_{BL} \,, \qquad (128) \\
&\quad \leq Const/N + \|T_t(\eta^{\mathrm{aux}}_{M(t)}) - T_t(\rho)\|_{BL} + d(\eta^{\mathrm{aux}}_0, \rho_0) \,, \qquad (129) \\
&\quad \leq Const/N + Const \cdot \int_0^t \mathrm{d}t' \, \big\|\eta^{\mathrm{aux}}_{t'} - \rho_{t'}\big\|_{BL} \,. \qquad (130)
\end{aligned}$$

The last line follows with lemma 5 and with $d(\eta^{\mathrm{aux}}_0, \rho_0) = 0$ because $\eta^{\mathrm{aux}}_0 = \rho_0$. By applying Grönwall's inequality to the last line above, one obtains the estimate in equation (110):

$$d(\eta^{\mathrm{aux}}_{M(t)}, \rho_t) \leq \frac{1}{N} e^{Const \cdot t} \,. \qquad (131)$$

Combining the estimates for summand (i) in equation (109) and summand (ii) in equation (110), lemma 7 follows. □

C.3.4 Proof of the theorem

Proof. We estimate the expectation value of the distance between the empirical one-particle density of the microscopic stochastic many-particle process and the mean-field distribution by the estimates obtained in lemmata 1, 4 and 7. In total, one finds a $0 < Const < \infty$ such that:

$$\begin{aligned}
\mathbb{E}\Big(d(\rho^{(1)}_N(t), \rho(t))\Big) &\leq \mathbb{E}\Big(d(\rho^{(1)}_{N,\kappa(t)}, \eta^{\mathrm{aux}}_{N,\kappa(t)})\Big) + \mathbb{E}\Big(d(\eta^{\mathrm{aux}}_{N,\kappa(t)}, \eta^{\mathrm{aux}}_{\kappa(t)})\Big) + \mathbb{E}\Big(d(\eta^{\mathrm{aux}}_{\kappa(t)}, \rho_t)\Big) \,, \qquad (132) \\
&\leq e^{Const \cdot \frac{M(t)}{N}} \left(\mathbb{E}\Big(d(\rho^{(1)}_{N,0}, \eta^{\mathrm{aux}}_{N,0})\Big) + \frac{M(t)^{3/4}}{N}\right) \\
&\quad + Const \cdot \frac{M(t)^{3/4}}{N} + \mathbb{E}\Big(d(\eta^{\mathrm{aux}}_{N,0}, \eta^{\mathrm{aux}}_0)\Big) + Const(t) \cdot N^{-1/4} \,. \qquad (133)
\end{aligned}$$

Note that $\eta^{\mathrm{aux}}_0 = \rho_0$. Note also that expectation values above are taken with respect to both sampling production degrees and sampling update times. Furthermore, for a given time t,

$M(t)$ is of order N. Thus, one estimates:

$$\mathbb{E}\left(d(\rho_N^{(1)}(t), \rho(t))\right) \leq Const(t) \cdot \left(\mathbb{E}\left(d(\eta_{N,0}^{\text{aux}}, \rho_0)\right) + N^{-1/4}\right), \tag{134}$$

with some constant $0 < Const < \infty$ that depends on the chosen time t. Applying Markov's inequality establishes an error estimate of the convergence with an (even N-dependent) ϵ_N:

$$\mathbb{P}\left(d(\rho_N^{(1)}(t), \rho(t)) > \epsilon_N\right) \leq \frac{\mathbb{E}\left(d(\rho_N^{(1)}(t), \rho(t))\right)}{\epsilon_N} \tag{135}$$

$$\leq \frac{Const(t)}{\epsilon_N} \cdot \left(\mathbb{E}\left(d(\eta_{N,0}^{\text{aux}}, \rho_0)\right) + N^{-1/4}\right), \tag{136}$$

which proves theorem 1.

□

D Analysis of the mean-field equation (1)

D.1 Moment and cumulant equations

Mean-field equation for moment and cumulant-generating functions

The moment-generating function $M(u,t)$ for the production degree p, which is the random variable of interest, and its corresponding cumulant-generating function $C(u,t)$ are defined as:

$$M(u,t) \coloneqq \int_0^1 \mathrm{d}p\; e^{up}\rho(p,t) = \mathcal{L}[\rho](-u,t)\,, \tag{137}$$

$$C(u,t) \coloneqq \ln\left(M(u,t)\right)\,, \tag{138}$$

with argument $u \in (-\infty,\infty)$ at time t. The moment-generating function M is the (one-sided) Laplace transform $\mathcal{L}$ of ρ with negative argument at time t. Moments and cumulants of the degree distribution ρ are obtained as:

$$M_k(t) \coloneqq \partial_u^k M(u,t)|_{u=0} \quad \text{and} \quad C_k(t) \coloneqq \partial_u^k C(u,t)|_{u=0}\,, \quad \text{for } k \geq 1\,. \tag{139}$$

For the mean production, that is for the expectation value of the production distribution, it holds that $\overline{p} = M_1 = C_1$ and the variance is given by $\mathrm{Var}(p) = \overline{p^2} - \overline{p}^2 = M_2 - M_1^2 = C_2$. By applying transformations (137, 138) to the mean-field equation (1) and plugging in the form of the fitness function in equation (4), one obtains:

$$\begin{aligned}
\partial_t M(u,t) &= (1-2\lambda)s\left(M_1(t)M(u,t) - \partial_u M(u,t)\right) \\
&\quad + 2\lambda\left(1 - sM_1(t)\right)\left(e^{uR(M_1(t))} - M(u,t)\right)\,, \\
\partial_t C(u,t) &= (1-2\lambda)s\left(C_1(t) - \partial_u C(u,t)\right) \\
&\quad + 2\lambda(1 - sC_1(t))\left(e^{uR(C_1(t))}e^{-C(u,t)} - 1\right)\,.
\end{aligned} \tag{140}$$

Solution strategy for the moment and cumulant-generating functions: Method of characteristics

This mean-field equation in moment/cumulant space (140) is more conveniently written as a semilinear partial differential equation (PDE) of first order in t and u, for example for C:

$$\partial_t C(u,t) + (1-2\lambda)s\partial_u C(u,t) = F(C,u,t)\,, \tag{141}$$

with $F(C,u,t) \coloneqq (1-2\lambda)sC_1(t) + 2\lambda(1-sC_1)\left(e^{uR(C_1(t))}e^{-C(u,t)} - 1\right)$ and initial condition $C(u,t=0) = C_0(u)$. This PDE admits the straight lines $r(u,t) = u - (1-2\lambda)st$ as characteristics. Restricted to these characteristic curves, the PDE reduces to a nonlinear

ordinary differential equation (ODE) of first order in time for $z(r,t) = C(u(r,t),t)$:

$$\frac{\mathrm{d}}{\mathrm{d}t} z(r,t) = \partial_t u(r,t) \partial_u C(u,t) + \partial_t C(u,t)\,, \tag{142}$$

$$= F(C(u(r,t),t), u(r,t), t)\,, \tag{143}$$

$$= F(z,r,t)\,, \tag{144}$$

with initial condition $z(r, t = 0) = C(u(r, t = 0), t = 0) = C_0(r)$. The solution for the cumulant-generating function is then obtained from the solution of the above ODE as $C(u,t) = z(r(u,t),t) = z(u - (1-2\lambda)st, t)$. For the two cases $\lambda = 0$ and $\lambda = 1/2$ with linear response function, an insightful, analytical solution of the mean-field equation (1) for the production distribution for all times t was found this way; see below.

Moment and cumulant equations

A different approach to characterize the dynamics of the quorum-sensing model is to analyze the equations of motions for the moments and cumulants. The moment equations are derived from equation (140) by applying the definition of the moments (139), which yields for $k \geq 1$,

$$\begin{aligned} \partial_t M_k(t) = {} & (1-2\lambda) s (M_1(t) M_k(t) - M_{k+1}(t)) \\ & + 2\lambda \left(1 - s M_1(t)\right) \left(R^k(M_1(t)) - M_k(t) \right) . \end{aligned} \tag{145}$$

The evolution equations for the first three cumulants are obtained as,

$$\partial_t C_1(t) = -(1-2\lambda) s C_2(t) + 2\lambda (1 - s C_1(t)) \left(R(C_1(t)) - C_1(t)\right) , \tag{146}$$

$$\partial_t C_2(t) = -(1-2\lambda) s C_3(t) + 2\lambda (1 - s C_1(t)) \left(-C_2(t) + (R(C_1(t)) - C_1(t))^2 \right) , \tag{147}$$

$$\begin{aligned} \partial_t C_3(t) = {} & -(1-2\lambda) s C_4(t) \\ & + 2\lambda (1 - s C_1(t)) \left(-C_3(t) - 3(R(C_1(t)) - C_1(t)) C_2(t) + (R(C_1(t)) - C_1(t))^3 \right) . \end{aligned} \tag{148}$$

For figure 2(E), the cumulant equations (146-148) were numerically integrated after applying a Gaussian approximation, that is a cumulant closure with $C_i(t) = 0$ for $i \geq 3$ and all t, and plotted for $\overline{p}_t = C_1(t)$.

D.2 Without sense-and-response to the environment: unimodal stationary distributions (homogeneity)

Homogeneous stationary distributions

For the case without sense-and-response through quorum sensing, $\lambda = 0$, it is readily seen from equation (1) that stationary production distributions are given by δ-peaks as $\rho_\infty(p) \coloneqq \rho(p, t \to \infty) = \delta(p - p_{\text{low}})$ for all $p_{\text{low}} \in [0, 1]$. However, the distribution with solely non-producers, $p_{\text{low}} = 0$, is the only asymptotically stable solution of the mean-field equation (1); see below.

When sense-and-response is absent, the analytical solution of the mean-field equation (1) for ρ can be obtained by applying the method of characteristics to equation (141) as outlined above. The implicit solution is given by:

$$C(u, t) = C_0(u - st) + st\langle\overline{p}\rangle_t\,, \quad \text{with } \langle\overline{p}\rangle_t \coloneqq 1/t \int_0^t \mathrm{d}t'\, \overline{p}_{t'} \tag{149}$$

as the temporal average of the mean production $\overline{p}_t$. Back-transformation and exploiting normalization of ρ yields:

$$\rho(p, t) = \rho_0(p)e^{-st(p-\langle\overline{p}\rangle_t)} = \rho_0(p)e^{-stp}/\mathcal{L}[\rho_0](st)\,. \tag{150}$$

For example, if the initial production distribution ρ_0 is a uniform distribution on $[0, 1]$, ρ evolves in time as $\rho(p, t) = st/(1 - e^{-st})e^{-stp}$, which is plotted in figure 2(A) (black, solid lines). Every production degree that is different from $p = 0$ decays exponentially fast and the time scale of the decay is set by the inverse of the value of that production degree. As $p \to 0$, this time scale diverges and, hence, the stationary distribution,

$$\rho_\infty(p) = \delta(p)\,, \tag{151}$$

is approached algebraically slowly; see figure 2(D).

Approach of homogeneous stationary distributions

To quantify the dependence of the time scales to approach stationarity on the initial distribution in more generality, we analyzed the temporal solution of the mean $\overline{p}_t$, which is obtained from the solution for the cumulant generating function as:

$$\overline{p}_t = -\partial_v \ln \mathcal{L}[\rho_0](v)|_{v=st}\,. \tag{152}$$

Therefore, the temporal evolution of the mean production depends only on the initial distribution ρ_0 via its Laplace transform $\mathcal{L}[\rho_0]$. For the asymptotic behavior of Laplace transforms it is known that if $\rho_0(p) \sim p^\mu$ as $p \to 0$ with $\mu > -1$, then $\mathcal{L}[\rho_0](v) \sim 1/v^{(\mu+1)}$ for $v \gg 1$ [133]. Therefore, it follows that the mean evolves as $\overline{p}_t \sim 1/t$ for $t \gg 1$ if the initial production distribution is a continuous probability density with non-vanishing weight at $p_{\text{low}} = 0$ (chosen for simplicity as the lowest production degree). The condition that the

exponent satisfies $\mu > -1$ is always fulfilled for a continuous probability distribution to ensure integrability at zero. In the same manner, the decay of the variance is shown to evolve algebraically as $\mathrm{Var}(p)(t) \sim 1/t^2$ for $t \gg 1$.

In contrast, if the lowest production degree is separated from all other degrees in the population by a gap $\Delta > 0$ in production space, mean and variance approach their stationary value exponentially fast at a time scale set by Δ. To see this qualitative difference in the approach of stationarity, we consider an initial probability distribution with probability mass $y_0 > 0$ at degree $p_{\mathrm{low}} = 0$ (chosen again for simplicity) and a remainder probability distribution $\widetilde{\rho}_0$ with support on $[\Delta, 1]$: $\rho_0(p) = y_0\delta(p) + (1 - y_0)\widetilde{\rho}_0(p)\mathbb{I}_{[\Delta,1]}(p)$ (here $\mathbb{I}_{[\Delta,1]}$ denotes the indicator function, which takes value 1 on the interval $[\Delta, 1]$ and 0 otherwise, and highlights the support of $\widetilde{\rho}_0$ on $[\Delta, 1]$). Using this form for ρ_0 and plugging in its Laplace transform into the solution for the mean in equation (152), one estimates $\overline{p}_t \lesssim (1 + \Delta)e^{-s\Delta \cdot t}$ for $t \gg 1$. This result generalizes the exponentially fast approach of stationarity that is known, for example, from the discrete Prisoner's dilemma in evolutionary game theory [26, 88–90].

In total, $\overline{p}_t$ vanishes exponentially fast if and only if the production degree at the smallest production degree is separated by a gap Δ from all other production degrees that are present in the population. On the other hand, if the lowest production degree is part of an interval with continuously distributed production degrees (that is, $\Delta = 0$), $\overline{p}_t$ decreases algebraically slowly.

D.3 With frequent sense-and-response to the environment: unimodal stationary distributions (homogeneity)

Homogeneous stationary distributions

For the case with sense-and-response through quorum sensing, $\lambda > 0$, one obtains from equation (1) or from the cumulant equations (146-148) that stationary production distributions are given by δ-peaks as:

$$\rho_\infty(p) = \delta(p - p^*)\,, \text{ with } R(p^*) = p^* \in [0, 1]\,. \tag{153}$$

In other words, fixed points of the response function give rise to homogeneous stationary distributions. Whether these stationary distributions are stable against small perturbations around stationarity depends on the stability of the fixed points (see linear stability analysis of homogeneous stationary distributions below). Whether they are approached for long times depends, in addition to the dependence on the stability of the fixed points, also on the initial distribution, and on the response function and the value of λ (see heterogeneous stationary distributions).

Linear stability analysis of homogeneous stationary distributions

Here, we supplement the statements from sections 2 and 4 on the stability of homogeneous stationary distributions in the linear approximation around stationarity if sense-and-response is present ($\lambda > 0$). For the sake of simplicity and feasibility, we carry out the stability analysis in the space of cumulants. To this end, we define the vector:

$$\mathbf{C}(t) = (C_1(t), C_2(t), C_3(t), \ldots)\,, \tag{154}$$

which is at stationarity (see equation (153)):

$$\mathbf{C}_\infty := \mathbf{C}(t \to \infty) = (C_{1,\infty}, C_{2,\infty}, C_{3,\infty}, \ldots) = (p^*, 0, 0, \ldots)\,. \tag{155}$$

With this notation, the equations of motion for the cumulants of ρ are given as follows:

$$\partial_t C_i(t) = F_i(\mathbf{C}(t))\,, \text{ for } i \geq 1\,. \tag{156}$$

Here, the functions F_i for $i \geq 0$ are defined by the right hand side of the cumulant equations (146-148). Upon introducing the distance $\Delta\mathbf{C}$ to the stationary vector $\mathbf{C}_\infty$, that is $\Delta\mathbf{C} = \mathbf{C} - \mathbf{C}_\infty$, one obtains the temporal behavior of $\Delta\mathbf{C}$ as:

$$\partial_t \Delta C_i(t) = F_i(\mathbf{C}_\infty + \Delta\mathbf{C}(t)) = \sum_{j=0}^{\infty} J_{ij}(\mathbf{C}_\infty)\Delta C_j(t) + \mathcal{O}(\|\Delta\mathbf{C}\|^2)\,, \text{ for } i \geq 0\,, \tag{157}$$

with Jacobian $J_{ij}(\mathbf{C}_\infty) = \frac{\partial F_i(\mathbf{C})}{\partial C_j}|_{\mathbf{C}=\mathbf{C}_\infty}$, whose entries are obtained after some algebra as:

$$J_{11} = -2\lambda(1 - sp^*)(1 - R'(p^*)) , \tag{158}$$
$$J_{i,i} = -2\lambda(1 - sp^*) , \text{ for } i \geq 2 , \tag{159}$$
$$J_{i,i+1} = -(1 - 2\lambda)s , \text{ for } i \geq 1 , \tag{160}$$
$$J_{i,j} = 0 , \text{ otherwise} . \tag{161}$$

The eigenvalues of the upper triangular matrix J determine the stability of the stationary distribution up to linear order in perturbations at the level of cumulants around stationarity. Because of the upper triangular structure of the Jacobian J, its eigenvalues are given by the diagonal entries of J:

$$\gamma_1 = -2\lambda(1 - sp^*)(1 - R'(p^*)) , \tag{162}$$
$$\gamma_i = -(1 - 2\lambda)s < 0 , \text{ for } i \geq 2 . \tag{163}$$

Thus, local stability of homogeneous stationary distributions ($\rho_\infty(p) = \delta(p - p^*)$ with $R(p^*) = p^*$) is determined by the stability of the fixed points, that is whether $R'(p^*)$ is less or greater than 1.

In total, homogeneous stationary distributions are unstable up to linear order in perturbations at the level of cumulants around stationarity if $R'(p^*) > 1$. In other words, stationary distributions located at a fixed point p^* are linearly unstable if p^* is an unstable fixed point of the response function ($R'(p^*) > 1$). On the other hand, linear stability of the response function at p^* ($R'(p^*) \leq 1$) yields linearly stable homogeneous stationary distributions located at p^*.

With sense-and-response ($\lambda = 1/2$) and linear response function ($R(p) = p$): Analytical solution and approach of homogeneous stationary distribution

For the choice of linear response function ($R(p) = p$, that is, $R'(p) = 1$ for all $p \in [0, 1]$) and $\lambda = 1/2$, the mean remains constant in time (see equation (146)). Furthermore, one obtains the analytical solution of the mean-field equation (1) by applying the method of characteristics (most conveniently in the space of moment generating functions) as:

$$M(u, t) = M_0(u)e^{-\overline{\phi}_0 t} + e^{u\overline{p}_0}\left(1 - e^{-\overline{\phi}_0 t}\right) , \tag{164}$$

which yields after back-transformation:

$$\rho(p, t) = y(t)\rho_0(p) + (1 - y(t))\delta(p - \overline{p}_0) , \quad \text{with } y(t) = \exp(-\overline{\phi}_0 t) . \tag{165}$$

The initial production distribution ρ_0 decays exponentially fast on a time scale that is set by the average initial fitness in the population $\overline{\phi}_0$, whereas a singular probability mass at the initial mean production degree $\overline{p}_0$ builds up concomitantly due to sense-and-response through quorum sensing. The population approaches the stationary distribution $\rho_\infty(p) = \delta(p - \overline{p}_0)$ exponentially fast.

With sense-and-response ($\lambda = 1/2$) and polynomial response function: Divergence of time scales at bifurcations of parameters of the response function

For $\lambda > s/2$, the approach of stationarity is typically exponentially fast. However, upon fine-tuning parameters of the response function one observes an algebraically slow approach of stationarity. We exemplify this qualitative change in the temporal evolution by setting the response probability to $\lambda = 1/2$ and by considering the following nonlinear response function, see supplementary figure A.3 (for the sake of readability, we label the argument of R by p instead of $\langle p \rangle$):

$$R(p) = p + A \cdot p(p - (p_{\mathrm{cr}} - \epsilon))(p - p_{\mathrm{cr}})(p - (p_{\mathrm{cr}} + \epsilon))(p - 1)\,, \tag{166}$$

with some real constant $A > 0$. The chosen response function (166) is a polynomial of 5th order with $R(0) = 0$ and $R(1) = 1$, and parameter $0 < p_{\mathrm{cr}} < 1$, which is set to $p_{\mathrm{cr}} = 1/2$ in supplementary figure A.3. The bifurcation parameter $0 \leq \epsilon \leq \min(p_{\mathrm{cr}}, 1 - p_{\mathrm{cr}})$ controls a supercritical pitchfork bifurcation of the response function (166) at $p^* = p_{\mathrm{cr}}$: Whereas $p^* = 0$ and $p^* = 1$ are unstable fixed points for all ϵ, the fixed points at $p^* = p_{\mathrm{cr}} \pm \epsilon$ are stable for $\epsilon > 0$ and merge with $p^* = p_{\mathrm{cr}}$ for $\epsilon = 0$. The fixed point $p^* = p_{\mathrm{cr}}$ is unstable for $\epsilon > 0$ and is a three-fold degenerate, stable fixed point for $\epsilon = 0$, see supplementary figure A.3(A, B).

For $\lambda = 1/2$ and upon plugging in the explicit form of the response function (166), the temporal evolution equation of the mean (146) is given by the ODE:

$$\partial_t C_1 = A(1 - sC_1)C_1(C_1 - (p_{\mathrm{cr}} - \epsilon))(C_1 - p_{\mathrm{cr}})(C_1 - (p_{\mathrm{cr}} + \epsilon))(C_1 - 1)\,, \tag{167}$$

with initial condition $C_1(t = 0) = \overline{p}_0$. From integrating this temporal evolution equation, one obtains the implicit solution for the mean $\overline{p} = C_1$ as:

$$t = \sum_{p^*} \alpha_{p^*} \int_{\overline{p}_0}^{\overline{p}_t} \frac{\mathrm{d}C_1}{C_1 - p^*}\,. \tag{168}$$

The sum is performed over all non-degenerate fixed points of the right hand side of the equation for the mean (167), that is over the roots $p^* \in \{0, p_{\mathrm{cr}} - \epsilon, p_{\mathrm{cr}}, p_{\mathrm{cr}} + \epsilon, 1, 1/s\}$ of both the response function (166) and the mean fitness $\overline{\phi}_t = 1 - s\overline{p}_t$. The coefficients α_{p^*} arise from the partial fraction decomposition with $\alpha_{p_{\mathrm{cr}}, p_{\mathrm{cr}} \pm \epsilon} \sim \mathcal{O}(1/\epsilon^2)$ and $\alpha_{0,1,1/s} \sim \mathcal{O}(\epsilon^0)$. Therefore, one concludes that:

$$|\overline{p}_t - \overline{p}_\infty| \sim e^{-t/\alpha}\,, \text{ for } \epsilon > 0\,, \tag{169}$$

for large times and with a decay constant α that diverges as the bifurcation is approached as $\alpha \sim 1/\epsilon^2$. In other words, stationarity is approached exponentially fast when all fixed points of the response function (166) are non-degenerate, see supplementary figure A.3(D) inset. Which of the two stable fixed points $p^* = p_{\mathrm{cr}} \pm \epsilon$ constitutes the stationary distribution $\rho_\infty(p) = \delta(p - p^*)$ depends on the initial distribution (and demographic fluctuations of the initial dynamics in the stochastic process). The prediction that the decay constant τ diverges

as the bifurcation of the response function is approached ($\epsilon \to 0$) is in good agreement with numerical simulations of the stochastic process, see supplementary figure A.3(D).

In contrast to the exponentially fast approach away from the bifurcation, stationarity is approached algebraically slowly at the bifurcation of the nonlinear response function, that is for $\epsilon = 0$. Since the stable fixed point $p^* = p_{\mathrm{cr}}$ is three-fold degenerate, one finds by integration of equation (168) the implicit solution for the mean as:

$$t = \sum_{p^* \neq p_{\mathrm{cr}}} \alpha_{p^*} \int_{\overline{p}_0}^{\overline{p}_t} \frac{\mathrm{d}C_1}{C_1 - p^*} + \sum_{i=1}^{z} \alpha_{p_{\mathrm{cr}}}^{(i)} \int_{\overline{p}_0}^{\overline{p}_t} \frac{\mathrm{d}C_1}{(C_1 - p_{\mathrm{cr}})^i} \,. \tag{170}$$

In addition to the sum over the non-degenerate fixed points ($p^* \neq p_{\mathrm{cr}}$), a second sum accounts for the degeneracy $z = 3$ of the fixed point p_{cr}, which is reflected by the singularities in the integrand up to order z. Consequently, the mean production approaches its stationary value as:

$$|\overline{p}_t - \overline{p}_\infty| \sim t^{-1/\nu} \,, \text{ for } \epsilon = 0 \,, \tag{171}$$

for large times with critical exponent $\nu = z - 1 = 2$, that is $-1/\nu = -1/2$. Supplementary figure A.3(C) shows the excellent agreement of our theoretical predictions with numerical simulations of the stochastic process for the algebraically slow approach of stationarity at the bifurcation.

D.4 With rare sense-and-response to the environment: bimodal stationary distributions (heterogeneity)

Heterogeneous stationary distributions

To analyze heterogeneous stationary distributions, we decompose the production distribution as follows:

$$\rho(p,t) = y(t)\rho_{\text{low}}(p,t) + (1-y(t))\rho_{\text{high}}(p,t)\,, \tag{172}$$

where ρ_{low} and ρ_{high} denote two probability distributions with support on the interval $[0,1]$. Their respective means are denoted as:

$$\overline{p}_{\text{low},t} = \int_0^1 \mathrm{d}p\; p\rho_{\text{low}}(p,t)\,, \text{ and } \quad \overline{p}_{\text{high},t} = \int_0^1 \mathrm{d}p\; p\rho_{\text{high}}(p,t)\,, \tag{173}$$

such that $\overline{p}_t = y(t)\overline{p}_{\text{low},t} + (1-y(t))\overline{p}_{\text{high},t}$; their stationary values are denoted as $\overline{p}_{\text{low},\infty} =: p_{\text{low}}$ and $\overline{p}_{\text{high},\infty} =: p_{\text{high}}$, respectively. We decompose the initial distribution $\rho_0(p) = y_0\rho_{\text{low},0}(p) + (1-y_0)\rho_{\text{high},0}(p)$ such that $\min(\text{supp}(\rho_{\text{low},0})) = \min(\text{supp}(\rho_0))$. For a numerical integration of the mean-field equation (1) that not only reproduces the stationary distribution, but also the temporal approach towards stationarity, it turns out suitable to choose the following decomposition: $\rho_{\text{low},0} = \rho_0$, $\rho_{\text{high},0} = \delta(\cdot - R(\overline{p}_0))$, and $y_0 = 1-\epsilon$ with $0 < \epsilon \lesssim 0.01$.

With decomposition (172), the mean-field equation (1) for ρ can be rewritten in terms of equations for ρ_{low}, ρ_{high}, and y as follows:

$$\partial_t \rho_{\text{low}}(p,t) = -s(1-2\lambda)\big(p - \overline{p}_{\text{low},t}\big)\rho_{\text{low}}(p,t)\,, \tag{174}$$

$$\begin{aligned}\partial_t \rho_{\text{high}}(p,t) = {} & -s(1-2\lambda)\big(p - \overline{p}_{\text{high},t}\big)\rho_{\text{high}}(p,t) \\ & + 2\lambda\frac{1-s\overline{p}_t}{1-y(t)}\big(\delta(p - R(\overline{p}_t)) - \rho_{\text{high}}(p,t)\big)\,,\end{aligned} \tag{175}$$

$$\partial_t y(t) = y(t)\left(-2\lambda(1-s\overline{p}_{\text{low},t}) + s(1-y(t))(\overline{p}_{\text{high},t} - \overline{p}_{\text{low},t})\right)\,. \tag{176}$$

We note that the decomposition (172) of ρ with equations (174-176) is not unique, but this choice of decomposition enables the characterization of heterogeneous stationary distributions and, thus, phenotypic heterogeneity.

The temporal evolution equation (174) for ρ_{low} has the form of the continuous replicator equation (see equation (1) with $\lambda = 0$) with renormalized selection strength $s(1-2\lambda)$. Following the analysis that resulted in equation (150), the solution for ρ_{low} is given by:

$$\begin{aligned}\rho_{\text{low}}(p,t) &= \rho_{\text{low},0}(p)e^{-s(1-2\lambda)tp}/\mathcal{L}[\rho_{\text{low},0}](s(1-2\lambda)t)\,, \\ \text{with} \quad \rho_{\text{low},0}(p) &= \rho_{\text{low}}(p,t=0)\,,\end{aligned} \tag{177}$$

if $\lambda \leq 1/2$. As shown in section 2, the condition $\lambda \leq 1/2$ is consistent with the condition for the upper threshold of the response probability $\lambda \leq s/2 < 1/2$, above which heterogeneous

stationary distributions cannot occur. For the mean $\overline{p}_{\mathrm{low},t}$, one obtains:

$$\overline{p}_{\mathrm{low},t} = -\partial_v \ln \mathcal{L}[\rho_{\mathrm{low},0}](v)|_{v=s(1-2\lambda)t} \ . \tag{178}$$

In other words, ρ_{low} approaches a stationary δ-distribution:

$$\begin{aligned} &\rho_{\mathrm{low}}(p, t \to \infty) = \rho_{\mathrm{low},\infty}(p) = \delta(p - p_{\mathrm{low}}) \,, \\ &\quad \text{with} \quad p_{\mathrm{low}} = \overline{p}_{\mathrm{low},\infty} = \min(\mathrm{supp}(\rho_{\mathrm{low},0})) = \min(\mathrm{supp}(\rho_0)) \,. \end{aligned} \tag{179}$$

The temporal evolution equation (175) for ρ_{high} has a similar form as the original mean-field equation (1): it involves the sense-and-response term with prefactor 2λ, and the replicator term with prefactor $1 - 2\lambda$. The sense-and-response term, however, couples to the full production distribution ρ through the argument $R(\overline{p}_t)$ in the δ-function and the prefactor $(1 - s\overline{p}_t)/(1 - y(t))$, whereas the replicator term does not couple to ρ_{low} or y. Equation (175) is most suitably analyzed in the space of moment and cumulant generating functions with ($u \in (-\infty, \infty)$):

$$M_{\mathrm{high}}(u, t) := \int_0^1 \mathrm{d}p \ e^{up} \rho_{\mathrm{high}}(p, t) \,, \tag{180}$$

$$C_{\mathrm{high}}(u, t) := \ln \Big(M_{\mathrm{high}}(u, t) \Big) \ . \tag{181}$$

The moments and cumulants of ρ_{high} are obtained as $M_{\mathrm{high},k}(t) := \partial_u^k M_{\mathrm{high}}(u, t)|_{u=0}$ and $C_{\mathrm{high},k}(t) := \partial_u^k C_{\mathrm{high}}(u, t)|_{u=0}$ for $k \geq 1$. With this notation, it is $\overline{p}_{\mathrm{high},t} = M_{\mathrm{high},1}(t) = C_{\mathrm{high},1}(t)$. By applying these transformations to the temporal evolution equation (175) of ρ_{high}, one obtains:

$$\begin{aligned} \partial_t M_{\mathrm{high}}(u, t) = &-(1 - 2\lambda)s \Big(\partial_u M_{\mathrm{high}}(u, t) - M_{\mathrm{high},1}(t) M_{\mathrm{high}}(u, t) \Big) \\ &+ 2\lambda \frac{1 - s\overline{p}_t}{1 - y(t)} \Big(e^{uR(\overline{p}_t)} - M_{\mathrm{high}}(u, t) \Big) \ , \end{aligned} \tag{182}$$

and,

$$\begin{aligned} \partial_t C_{\mathrm{high}}(u, t) = &-(1 - 2\lambda)s \Big(\partial_u C_{\mathrm{high}}(u, t) - C_{\mathrm{high},1}(t) \Big) \\ &+ 2\lambda \frac{1 - s\overline{p}_t}{1 - y(t)} \Big(e^{uR(\overline{p}_t)} e^{-C_{\mathrm{high}}(u,t)} - 1 \Big) \ , \end{aligned} \tag{183}$$

in which the coupling of ρ_{high} to ρ_{low} and y is apparent explicitly through the occurrence of the factor $1 - y(t)$ and implicitly through the occurrence of $\overline{p}_t = y(t)\overline{p}_{\mathrm{low},t} + (1 - y(t))\overline{p}_{\mathrm{high},t}$.

The corresponding equations of motion for the first three cumulants are, thus, obtained as:

$$\partial_t C_{\text{high},1}(t) = -(1-2\lambda)sC_{\text{high},2}(t) + 2\lambda\frac{1-s\overline{p}_t}{1-y(t)}\left(R(\overline{p}_t) - C_{\text{high},1}(t)\right)\,, \tag{184}$$

$$\begin{aligned}\partial_t C_{\text{high},2}(t) = {}& -(1-2\lambda)sC_{\text{high},3}(t) \\ & + 2\lambda\frac{1-s\overline{p}_t}{1-y(t)}\left(-C_{\text{high},2}(t) + (R(\overline{p}_t) - C_{\text{high},1}(t))^2\right)\,,\end{aligned} \tag{185}$$

$$\begin{aligned}\partial_t C_{\text{high},3}(t) = {}& -(1-2\lambda)sC_{\text{high},4}(t) \\ & + 2\lambda\frac{1-s\overline{p}_t}{1-y(t)}\Big(-C_{\text{high},3}(t) - 3(R(\overline{p}_t) - C_{\text{high},1}(t))C_{\text{high},2}(t) \\ & \qquad + (R(\overline{p}_t) - C_{\text{high},1}(t))^3\Big)\,.\end{aligned} \tag{186}$$

At stationarity, it is $\partial_t y(t) = 0$ and $y(t) \equiv y_\infty$ with (see equation (176); recall also that $\overline{p}_\infty = y_\infty p_{\text{low}} + (1-y_\infty)p_{\text{high}}$):

$$\begin{aligned}2\lambda(1-sp_{\text{low}}) &= s(1-y_\infty)(p_{\text{high}} - p_{\text{low}})\,, \\ \text{or, equivalently,}\quad (1-2\lambda)(1-sp_{\text{low}}) &= 1 - s\overline{p}_\infty\,.\end{aligned} \tag{187}$$

Assuming that a stationary value $0 < y_\infty < 1$ exists, it fulfils the self-consistency relation:

$$y_\infty = 1 - \frac{2\lambda}{s}\frac{1-sp_{\text{low}}}{p_{\text{high}} - p_{\text{low}}} = \frac{p_{\text{high}} - \overline{p}_\infty}{p_{\text{high}} - p_{\text{low}}}\,. \tag{188}$$

Note that we denoted y_∞ simply as y in sections 2 – 4. If $0 < y_\infty < 1$ exists, it follows that the stationary solution for ρ_{high} can be obtained via equation (182) in terms of the stationary moment generating function $M_{\text{high},\infty}(u) = M_{\text{high}}(u, t\to\infty)$ with:

$$\begin{aligned}\partial_u M_{\text{high},\infty}(u) - p_{\text{low}}M_{\text{high},\infty}(u) &= (p_{\text{high}} - p_{\text{low}})e^{uR(\overline{p}_\infty)}\,, \\ \text{and}\quad p_{\text{high}} &= \partial_u M_{\text{high},\infty}(u)|_{u=0}\,,\end{aligned} \tag{189}$$

where the relation between $p_{\text{low}}, p_{\text{high}}$, and y_∞ in equation (187) was exploited and the definition $p_{\text{high}} = \overline{p}_{\text{high},\infty}$ translates into the boundary condition. In total, one obtains $M_{\text{high},\infty}(u) = e^{up_{\text{high}}}$ with the self-consistency relation $p_{\text{high}} = R(\overline{p}_\infty)$. In other words, ρ_{high} approaches a stationary δ-distribution:

$$\rho_{\text{high}}(p, t\to\infty) = \rho_{\text{high},\infty}(p) = \delta(p - p_{\text{high}})\,, \tag{190}$$

$$\text{with}\quad p_{\text{high}} = \overline{p}_{\text{high},\infty} = R(\overline{p}_\infty) = R(2\lambda/s + (1-2\lambda)p_{\text{low}})\,. \tag{191}$$

For $p_{\text{low}} = \min(\text{supp}(\rho_0)) = 0$, one recovers from equations (179, 188, 190) the heterogeneous stationary distribution (2).

For figure 2(F), equations (176), (178), and (184-186) were numerically integrated with $C_{\text{high},i}(t) = 0$ for $i \geq 3$ and for all t, and initial conditions $y_0 = 0.99$, $\rho_{\text{low},0} \sim \text{Uniform}(0,1)$,

and $\rho_{\text{high},0} \sim \delta(\cdot - R(0.5))$. The choice of initial conditions, however, is not important for the asymptotic behavior, see supplementary figure A.1.

Linear stability analysis of heterogeneous stationary distributions

Here, we supplement the statements from sections 2 and 4 on the stability of heterogeneous stationary distributions (2) in the linear approximation around stationarity. For the sake of simplicity and feasibility, we carry out the stability analysis in the space of cumulants. To this end, we define the vector:

$$\mathbf{c}(t) = (y(t), C_{\text{low},1}(t), C_{\text{high},1}(t), C_{\text{low},2}(t), C_{\text{high},2}(t), \ldots)\,, \tag{192}$$

$$= (c_0(t), c_1(t), c_2(t), \ldots)\,, \tag{193}$$

which is at stationarity:

$$\mathbf{c}_\infty := \mathbf{c}(t \to \infty)\,, \tag{194}$$

$$= (y, C_{\text{low},1}, C_{\text{high},1}, C_{\text{low},2}, C_{\text{high},2}, \ldots)\,, \tag{195}$$

$$= (y, p_{\text{low}}, p_{\text{high}}, 0, 0, \ldots)\,, \tag{196}$$

$$= (c_0, c_1, c_2, \ldots)\,. \tag{197}$$

The cumulants of ρ_{low} are obtained in the same way as for ρ_{high}, that is as $C_{\text{low},k}(t) := \partial_u^k C_{\text{low}}(u,t)|_{u=0}$ for $k \geq 1$ from $M_{\text{low}}(u,t) := \int_0^1 \mathrm{d}p\, e^{up}\rho_{\text{low}}(p,t)$ and $C_{\text{low}}(u,t) := \ln(M_{\text{low}}(u,t))$ for $u \in (-\infty, \infty)$. With this notation, the equations of motion for $y(t)$ in equation (176) and the cumulants of ρ_{low} and ρ_{high}, respectively, are cast into the compact form:

$$\partial_t c_i(t) = F_i(\mathbf{c}(t))\,, \text{ for } i \geq 0\,. \tag{198}$$

Upon introducing the distance $\Delta\mathbf{c}$ to the stationary vector $\mathbf{c}_\infty$, that is $\Delta\mathbf{c} = \mathbf{c} - \mathbf{c}_\infty$, one obtains the temporal behavior of $\Delta\mathbf{c}$ as follows:

$$\partial_t \Delta c_i(t) = F_i(\mathbf{c}_\infty + \Delta\mathbf{c}(t)) = \sum_{j=0}^{\infty} J_{ij}(\mathbf{c}_\infty)\Delta c_j(t) + \mathcal{O}(\|\Delta\mathbf{c}\|^2)\,, \text{ for } i \geq 0\,, \tag{199}$$

and with Jacobian $J_{ij}(\mathbf{c}_\infty) = \frac{\partial F_i(\mathbf{c})}{\partial c_j}|_{\mathbf{c}=\mathbf{c}_\infty}$.

The entries of the Jacobian J are obtained after some algebra as:

$$J_{00} = -sy(p_{\text{high}} - p_{\text{low}})\,, \qquad J_{01} = -sy(1-y)\frac{1 - sp_{\text{high}}}{1 - sp_{\text{low}}}\,, \qquad J_{02} = sy(1-y)\,, \tag{200}$$

$$J_{10} = 0\,, \qquad J_{11} = 0\,, \qquad J_{12} = 0, \tag{201}$$

and, furthermore,

$$J_{20} = -s(1-2\lambda)(p_{\text{high}} - p_{\text{low}})^2 R'(\overline{p}_\infty)\,, \tag{202}$$
$$J_{21} = s(1-2\lambda)y(p_{\text{high}} - p_{\text{low}})R'(\overline{p}_\infty)\,, \tag{203}$$
$$J_{22} = s(1-2\lambda)(p_{\text{high}} - p_{\text{low}})((1-y)R'(\overline{p}_\infty) - 1)\,, \tag{204}$$

and,

$$J_{i,i+2} = -s(1-2\lambda)\,,\ \text{for } i \geq 1\,, \tag{205}$$
$$J_{2i,2i} = -s(1-y)(p_{\text{high}} - p_{\text{low}})\,,\ \text{for } i \geq 2\,, \tag{206}$$
$$J_{i,j} = 0\,,\ \text{otherwise}\,. \tag{207}$$

The eigenvalues of the matrix J determine the stability of the heterogeneous stationary distribution up to linear order in perturbations at the level of cumulants around stationarity. Its eigenvalues are given by:

- the two eigenvalues $\gamma_{1,2}$ of the 2×2 matrix,

$$\tilde{J} = \begin{pmatrix} J_{00} & J_{02} \\ J_{20} & J_{22} \end{pmatrix} \tag{208}$$
$$= \begin{pmatrix} -sy(p_{\text{high}} - p_{\text{low}}) & sy(1-y) \\ -s(1-2\lambda)(p_{\text{high}} - p_{\text{low}})^2 R'(\overline{p}_\infty) & s(1-2\lambda)(p_{\text{high}} - p_{\text{low}})((1-y)R'(\overline{p}_\infty) - 1) \end{pmatrix},$$

- one eigenvalue 0,
- and infinitely many pairs of eigenvalues with values 0 and $-s(1-y)(p_{\text{high}} - p_{\text{low}}) < 0$ (because $p_{\text{high}} - p_{\text{low}} > 0$ and $1 - y > 0$ for the considered bimodal distributions).

For simplicity of the discussion, we assume $p_{\text{low}} = \min(\text{supp}(\rho_0)) = 0$ in the following, and also introduce the parameter $\beta = 2\lambda/s$ as in section 2. The two eigenvalues $\gamma_{1,2}$ of $\tilde{J}$ are given by:

$$\gamma_{1,2} = \frac{1}{2}\text{Tr}(\tilde{J}) \pm \left(\frac{1}{4}\text{Tr}(\tilde{J})^2 - \text{Det}(\tilde{J})\right)^{1/2}\,, \tag{209}$$
$$\text{with}\quad \text{Tr}(\tilde{J}) = s(1-2\lambda)(\beta R'(\beta) - R(\beta)) - s(R(\beta) - \beta)\,, \tag{210}$$
$$\text{and}\quad \text{Det}(\tilde{J}) = s^2(1-2\lambda)R(\beta)(R(\beta) - \beta))\,. \tag{211}$$

Linear stability for small λ.
Under the assumptions $R(0) = 0$ and $1 < R'(0) < \infty$, one checks that for $0 < \lambda \ll 1$ the eigenvalues of the Jacobian $\tilde{J}$ in equation (209) are given by:

$$\gamma_{1,2} = -\lambda(R'(0) - 1) + \mathcal{O}(\lambda^3) \pm i\lambda\left((R'(0) - 1)(3R'(0) + 1) + \mathcal{O}(\lambda)\right)^{1/2}\,, \tag{212}$$

and, thus, $\text{Re}(\gamma_{1,2}) < 0$ as $\lambda \searrow 0$.

Therefore, for small response probabilities, the heterogeneous stationary distribution (equation (2)) is stable up to linear order in perturbations at the level of cumulants around stationarity (here shown under the assumptions $p_{\mathrm{low}} = \min(\mathrm{supp}(\rho_0)) = 0$, $R(0) = 0$, and $1 < R'(0) < \infty$).

Linear stability for the specific response function $R(\beta) = \beta + \kappa \cdot \sin(\pi\beta)$.
Upon choosing the response function $R(\beta) = \beta + \kappa \cdot \sin(\pi\beta)$ with $\beta \in [0, 1]$ (that is $\lambda \in [0, s/2]$) and with $\kappa \in [0, 1/\pi]$, one checks that all eigenvalues of the Jacobian $\tilde{J}$ in equation (209) have negative real part.

Therefore, for the special choice of the response function that up-regulates the cellular autoinducer production for all sensed average productions in the population, all heterogeneous stationary distributions (equation (2)) for choices of the parameters $\lambda \in [0, s/2]$ and $\kappa \in [0, 1/\pi]$ are stable up to linear order in perturbations at the level of cumulants around stationarity.

Bibliography

[1] P. W. Anderson. More is different. *Science* **177**(4047), 393–396, 1972. DOI: 10.1126/science.177.4047.393.

[2] S. N. Bose. Plancks Gesetz und Lichtquantenhypothese. *Zeitschrift für Physik* **26**(1), 178–181, 1924. DOI: 10.1007/BF01327326.

[3] A. Einstein. Quantentheorie des einatomigen idealen Gases. *Sitzungsberichte der Preußischen Akademie der Wissenschaften*, 261–267, 1924. DOI: 10.1002/3527608958.ch29.

[4] A. Einstein. Quantentheorie des einatomigen idealen Gases. Zweite Abhandlung. *Sitzungsberichte der Preußischen Akademie der Wissenschaften*, 3–14, 1925. DOI: 10.1002/3527608958.ch28.

[5] M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. A. Cornell. Observation of Bose-Einstein condensation in a dilute atomic vapor. *Science* **269**(5221), 198–201, 1995. DOI: 10.1126/science.269.5221.198.

[6] K. B. Davis, M.-O. Mewes, M. R. Andrews, N. J. van Druten, D. S. Durfee, D. M. Kurn, and W. Ketterle. Bose-Einstein condensation in a gas of sodium atoms. *Physical Review Letters* **75**, 3969–3973, 1995. DOI: 10.1103/PhysRevLett.75.3969.

[7] D. Vorberg, W. Wustmann, R. Ketzmerick, and A. Eckardt. Generalized Bose-Einstein condensation into multiple states in driven-dissipative systems. *Physical Review Letters* **111**, 240405, 2013. DOI: 10.1103/PhysRevLett.111.240405.

[8] C. M. Waters and B. L. Bassler. Quorum sensing: Cell-to-cell communication in bacteria. *Annual Review of Cell and Developmental Biology* **21**(1), 319–346, 2005. DOI: 10.1146/annurev.cellbio.21.012704.131001.

[9] L. Keller and M. G. Surette. Communication in bacteria: An ecological and evolutionary perspective. *Nature Reviews Microbiology* **4**(4), 249–258, 2006. DOI: 10.1038/nrmicro1383.

[10] D. Garmyn, L. Gal, R. Briandet, M. Guilbaud, J.-P. Lemaitre, A. Hartmann, and P. Piveteau. Evidence of autoinduction heterogeneity via expression of the Agr system of *Listeria monocytogenes* at the single-cell level. *Applied and Environmental Microbiology* **77**(17), 6286–6289, 2011. DOI: 10.1128/AEM.02891-10.

[11] J. Grote, D. Krysciak, and W. R. Streit. Phenotypic heterogeneity, a phenomenon that may explain why quorum sensing not always results in truly homogenous cell behavior. *Applied and Environmental Microbiology* **81**(16), 2015. DOI: 10.1128/AEM.00900-15.

[12] M. Ackermann. A functional perspective on phenotypic heterogeneity in microorganisms. *Nature Reviews Microbiology* **13**(8), 497–508, 2015. DOI: 10.1038/nrmicro3491.

[13] A. Eldar and M. B. Elowitz. Functional roles for noise in genetic circuits. *Nature* **467**(7312), 167–173, 2010. DOI: 10.1038/nature09326.

[14] K. Fujimoto and S. Sawai. A design principle of group-level decision making in cell populations. *PLOS Computational Biology* **9**(6), e1003110–, 2013. DOI: 10.1371/journal.pcbi.1003110.

[15] A. Campa, T. Dauxois, and S. Ruffo. Statistical mechanics and dynamics of solvable models with long-range interactions. *Physics Reports* **480**(3-6), 57–159, 2009. DOI: 10.1016/j.physrep.2009.07.001.

[16] C. Gardiner. *Stochastic methods: A handbook for the natural and social sciences*. Springer, Berlin, 2009. DOI: n/a.

[17] N. G. Van Kampen. *Stochastic processes in physics and chemistry*. Elsevier, Amsterdam, 2007. DOI: 10.1016/b978-0-444-52965-7.x5000-4.

[18] M. F. Weber and E. Frey. Master equations and the theory of stochastic path integrals. *Reports on Progress in Physics* **80**(4), 046601, 2017. DOI: 10.1088/1361-6633/aa5ae2.

[19] S. Grosskinsky, F. Redig, and K. Vafayi. Condensation in the inclusion process and related models. *Journal of Statistical Physics* **142**(5), 952–974, 2011. DOI: 10.1007/s10955-011-0151-9.

[20] J. Knebel, M. F. Weber, T. Krüger, and E. Frey. Evolutionary games of condensates in coupled birth-death processes. *Nature Communications* **6**, 6977, 2015. DOI: 10.1038/ncomms7977.

[21] J. Hofbauer and K. Sigmund. *Evolutionary games and population dynamics*. Cambridge University Press, Cambridge, UK, 1998. DOI: 10.1017/CBO9781139173179.

[22] E. Frey. Evolutionary game theory: Theoretical concepts and applications to microbial communities. *Physica A: Statistical Mechanics and its Applications* **389**(20), 4265–4298, 2010. DOI: 10.1016/j.physa.2010.02.047.

[23] M. A. Nowak. *Evolutionary dynamics*. Harvard University Press, Cambridge, MA, 2006. DOI: n/a.

[24] J. Maynard Smith. *Evolution and the theory of games*. Cambridge University Press, Cambridge, UK, 1982. DOI: 10.1017/CBO9780511806292.

[25] M. A. Nowak and K. Sigmund. Evolutionary dynamics of biological games. *Science* **303**(5659), 793–799, 2004. DOI: 10.1126/science.1093411.

[26] M. A. Nowak, A. Sasaki, C. Taylor, and D. Fudenberg. Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**(6983), 646–650, 2004. DOI: 10.1038/nature02414.

[27] B. Sinervo and C. M. Lively. The rock-paper-scissors game and the evolution of alternative male strategies. *Nature* **380**(6571), 240–243, 1996. DOI: 10.1038/380240a0.

[28] B. Kerr, M. Riley, M. Feldman, and B. Bohannan. Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature* **418**(6894), 171–174, 2002. DOI: 10.1038/nature00823.

[29] T. Reichenbach, M. Mobilia, and E. Frey. Mobility promotes and jeopardizes biodiversity in rock-paper-scissors games. *Nature* **448**(7157), 1046–1049, 2007.

[30] M. F. Weber, G. Poxleitner, E. Hebisch, E. Frey, and M. Opitz. Chemical warfare and survival strategies in bacterial range expansions. *Journal of The Royal Society Interface* **11**(96), 20140172, 2014. DOI: 10.1098/rsif.2014.0172.

[31] A. Szolnoki, M. Mobilia, L.-L. Jiang, B. Szczesny, A. M. Rucklidge, and M. Perc. Cyclic dominance in evolutionary games: A review. *Journal of The Royal Society Interface* **11**(100), 20140735, 2014. DOI: 10.1098/rsif.2014.0735.

[32] P. D. Taylor and L. B. Jonker. Evolutionary stable strategies and game dynamics. *Mathematical Biosciences* **40**(1), 145–156, 1978. DOI: 10.1016/0025-5564(78)90077-9.

[33] P. Schuster and K. Sigmund. Replicator dynamics. *Journal of Theoretical Biology* **100**(3), 533–538, 1983. DOI: 10.1016/0022-5193(83)90445-9.

[34] J. von Neumann and O. Morgenstern. *Theory of games and economic behavior*. 3rd edition. Princeton University Press, Princeton, 1953. DOI: 10.1515/9781400829460.

[35] J. Maynard Smith and G. R. Price. The logic of animal conflict. *Nature* **246**(5427), 15–18, 1973. DOI: 10.1038/246015a0.

[36] J. Knebel, T. Krüger, M. F. Weber, and E. Frey. Coexistence and survival in conservative Lotka-Volterra networks. *Physical Review Letters* **110**, 168106, 2013. DOI: 10.1103/PhysRevLett.110.168106.

[37] T. Chawanya and K. Tokita. Large-dimensional replicator equations with antisymmetric random interactions. *Journal of the Physical Society of Japan* **71**(2), 429–431, 2002. DOI: 10.1143/JPSJ.71.429.

[38] H.-P. Breuer and F. Petruccione. *The theory of open quantum systems*. Oxford University Press, Oxford, 2002. DOI: 10.1093/acprof:oso/9780199213900.001.0001.

[39] W. Wustmann. *Statistical mechanics of time-periodic quantum systems*. PhD thesis. Technische Universität Dresden, 2010.

[40] C. W. Gardiner and P. Zoller. *Quantum noise: A handbook of Markovian and non-Markovian quantum stochastic methods with applications to quantum optics*. Springer, Berlin Heidelberg, 2004. DOI: n/a.

[41] R. Blümel, A. Buchleitner, R. Graham, L. Sirko, U. Smilansky, and H. Walther. Dynamical localization in the microwave interaction of Rydberg atoms: The influence of noise. *Physical Review A* **44**, 4521–4540, 1991. DOI: 10.1103/PhysRevA.44.4521.

[42] S. Kohler, T. Dittrich, and P. Hänggi. Floquet-Markovian description of the parametrically driven, dissipative harmonic quantum oscillator. *Physical Review E* **55**, 300–313, 1997. DOI: 10.1103/PhysRevE.55.300.

[43] H.-P. Breuer, W. Huber, and F. Petruccione. Quasistationary distributions of dissipative nonlinear quantum oscillators in strong periodic driving fields. *Physical Review E* **61**, 4883–4889, 2000. DOI: 10.1103/PhysRevE.61.4883.

[44] M. Grifoni and P. Hänggi. Driven quantum tunneling. *Physics Reports* **304**(5–6), 229–354, 1998. DOI: 10.1016/S0370-1573(98)00022-2.

[45] R. Zwanzig. *Nonequilibrium statistical mechanics*. Oxford University Press, 2001. DOI: n/a.

[46] G. Floquet. Sur les équations différentielles linéaires à coefficients périodiques. *Annales de l'École Normale Supérieure* **12**, 47–88, 1883. DOI: n/a.

[47] W. Pauli. *Festschrift zum 60. Geburtstage A. Sommerfeld*. Hirzel, Leipzig, 1928. P. 30. DOI: n/a.

[48] L. Mandel and E. Wolf. *Optical coherence and quantum optics*. Cambridge University Press, Cambridge, UK, 1995. DOI: 10.1017/CBO9781139644105.

[49] C. H. Durney, S. O. Case, M. Pleimling, and R. K. P. Zia. Saddles, arrows, and spirals: Deterministic trajectories in cyclic competition of four species. *Physical Review E* **83**, 051108, 2011. DOI: 10.1103/PhysRevE.83.051108.

[50] R. K. P. Zia. General properties of a system of S species competing pairwise. *pre-print arXiv:1101.0018*, 2010.

[51] H. Kuhn and A. Tucker. *Linear inequalities and related systems*. Princeton University Press, Princeton, NJ, 1956. DOI: 10.1515/9781400881987.

[52] U. C. Täuber. *Critical dynamics*. Cambridge University Press, Cambridge, MA, 2014. DOI: 10.1017/CBO9781139046213.

[53] F. P. Kelly. *Reversibility and stochastic networks*. Cambridge University Press, New York, 2011. DOI: n/a.

[54] O. Taussky and H. Zassenhaus. On the similarity transformation between a matrix and its transpose. *Pacific Journal of Mathematics*, 893–896, 1959. DOI: 10.2140/pjm.1959.9.893.

[55] T. Muir. *A treatise on the theory of determinants*. Macmillan, 1882. DOI: n/a.

[56] C. E. Cullis. *Matrices and determinoids*. Vol. I and II. Cambridge University Press, Cambridge, UK, 1913. DOI: n/a.

[57] B. A. Hense and M. Schuster. Core principles of bacterial autoinducer systems. *Microbiology and Molecular Biology Reviews* **79**(1), 153–169, 2015. DOI: 10.1128/MMBR.00024-14.

[58] M. J. Gray, N. E. Freitag, and K. J. Boor. How the bacterial pathogen *Listeria monocytogenes* mediates the switch from environmental Dr. Jekyll to pathogenic Mr. Hyde. *Infection and Immunity* **74**(5), 2505–2512, 2006. DOI: 10.1128/IAI.74.5.2505-2512.2006.

[59] E. P. da Silva and E. C. P. De Martinis. Current knowledge and perspectives on biofilm formation: The case of *Listeria monocytogenes*. *Applied Microbiology and Biotechnology* **97**(3), 957–968, 2013. DOI: 10.1007/s00253-012-4611-1.

[60] K. B. Xavier and B. L. Bassler. LuxS quorum sensing: More than just a numbers game. *Current Opinion in Microbiology* **6**(2), 191–197, 2003. DOI: 10.1016/S1369-5274(03)00028-6.

[61] C. Anetzberger, T. Pirch, and K. Jung. Heterogeneity in quorum sensing-regulated bioluminescence of *Vibrio harveyi*. *Molecular Microbiology* **73**(2), 267–277, 2009. DOI: 10.1111/j.1365-2958.2009.06768.x.

[62] C. Fuqua and E. P. Greenberg. Listening in on bacteria: Acyl-homoserine lactone signalling. *Nature Reviews Molecular Cell Biology* **3**(9), 685–695, 2002. DOI: 10.1038/nrm907.

[63] K. Papenfort and B. L. Bassler. Quorum sensing signal-response systems in Gram-negative bacteria. *Nature Reviews Microbiology* **14**(9), 576–588, 2016. DOI: 10.1038/nrmicro.2016.89.

[64] J. W. Williams, X. Cui, A. Levchenko, and A. M. Stevens. Robust and sensitive control of a quorum-sensing circuit by two interlocked feedback loops. *Molecular Systems Biology* **4**(1), 2008. DOI: 10.1038/msb.2008.70.

[65] J. Q. Boedicker, M. E. Vincent, and R. F. Ismagilov. Microfluidic confinement of single cells of bacteria in small volumes initiates high-density behavior of quorum sensing and growth and reveals its variability. *Angewandte Chemie International Edition* **48**(32), 5908–5911, 2009. DOI: 10.1002/anie.200901550.

[66] P. D. Pérez and S. J. Hagen. Heterogeneous response to a quorum-sensing signal in the luminescence of individual *Vibrio fischeri*. *PLOS ONE* **5**(11), 1–9, 2010. DOI: 10.1371/journal.pone.0015473.

[67] J. Grote, D. Krysciak, A. Schorn, R. I. Dahlke, L. Soonvald, J. Müller, B. A. Hense, M. Schwarzfischer, M. Sauter, C. Schmeisser, and W. R. Streit. Evidence of autoinducer-dependent and -independent heterogeneous gene expression in *Sinorhizobium fredii* NGR234. *Applied and Environmental Microbiology* **80**(18), 5572–5582, 2014. DOI: 10.1128/AEM.01689-14.

[68] B. B. Pradhan and S. Chatterjee. Reversible non-genetic phenotypic heterogeneity in bacterial quorum sensing. *Molecular Microbiology* **92**(3), 557–569, 2014. DOI: 10.1111/mmi.12575.

[69] C. Anetzberger, U. Schell, and K. Jung. Single cell analysis of *Vibrio harveyi* uncovers functional heterogeneity in response to quorum sensing signals. *BMC Microbiology* **12**(1), 1–10, 2012. DOI: 10.1186/1471-2180-12-209.

[70] L. Plener, N. Lorenz, M. Reiger, T. Ramalho, U. Gerland, and K. Jung. The phosphorylation flow of the *Vibrio harveyi* quorum sensing cascade determines levels of phenotypic heterogeneity in the population. *Journal of Bacteriology* **197**(10), 2015. DOI: 10.1128/JB.02544-14.

[71] G. Cárcamo-Oyarce, P. Lumjiaktase, R. Kümmerli, and L. Eberl. Quorum sensing triggers the stochastic escape of individual cells from *Pseudomonas putida* biofilms. *Nature Communications* **6**, 2015. DOI: 10.1038/ncomms6945.

[72] D. Garmyn, L. Gal, J.-P. Lemaitre, A. Hartmann, and P. Piveteau. Communication and autoinduction in the species *Listeria monocytogenes*. *Communicative & Integrative Biology* **2**(4), 371–374, 2009. DOI: 10.4161/cib.2.4.8610.

[73] J. Pérez-Velázquez, M. Gölgeli, and R. Garcia-Contreras. Mathematical modelling of bacterial quorum sensing: A review. *Bulletin of Mathematical Biology* **78**(8), 1585–1639, 2016. DOI: 10.1007/s11538-016-0160-6.

[74] A. B. Goryachev, D.-J. Toh, K. B. Wee, T. Lee, H.-B. Zhang, and L.-H. Zhang. Transition to quorum sensing in an *Agrobacterium* population: A stochastic model. *PLOS Computational Biology* **1**(4), 2005. DOI: 10.1371/journal.pcbi.0010037.

[75] J. D. Dockery and J. P. Keener. A mathematical model for quorum sensing in *Pseudomonas aeruginosa*. *Bulletin of Mathematical Biology* **63**(1), 95–116, 2001. DOI: 10.1006/bulm.2000.0205.

[76] M. Antoni and S. Ruffo. Clustering and relaxation in Hamiltonian long-range dynamics. *Physical Review E* **52**, 2361–2374, 1995. DOI: 10.1103/PhysRevE.52.2361.

[77] Y. Y. Yamaguchi, J. Barré, F. Bouchet, T. Dauxois, and S. Ruffo. Stability criteria of the Vlasov equation and quasi-stationary states of the HMF-model. *Physica A: Statistical Mechanics and its Applications* **337**(1–2), 36–66, 2004. DOI: 10.1016/j.physa.2004.01.041.

[78] J. Barré, F. Bouchet, T. Dauxois, and S. Ruffo. Out-of-equilibrium states as statistical equilibria of an effective dynamics in a system with long-range interactions. *Physical Review Letters* **89**, 110601, 2002. DOI: 10.1103/PhysRevLett.89.110601.

[79] M. Y. Choi and J. Choi. Stability and ensemble inequivalence in a globally coupled system. *Physical Review Letters* **91**, 124101, 2003. DOI: 10.1103/PhysRevLett.91.124101.

[80] P. de Buyl, D. Mukamel, and S. Ruffo. Statistical mechanics of collisionless relaxation in a non-interacting system. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **369**(1935), 439–452, 2010. DOI: 10.1098/rsta.2010.0251.

[81] R. Pakter and Y. Levin. Non-equilibrium dynamics of an infinite range XY model in an external field. *Journal of Statistical Physics* **150**(3), 531–539, 2013. DOI: 10.1007/s10955-012-0576-9.

[82] A. Ruparell, J. F. Dubern, C. A. Ortori, F. Harrison, N. M. Halliday, A. Emtage, M. M. Ashawesh, C. A. Laughton, S. P. Diggle, P. Williams, D. A. Barrett, and K. R. Hardie. The fitness burden imposed by synthesising quorum sensing signals. *Scientific Reports* **6**, 33101, 2016. DOI: 10.1038/srep33101.

[83] S. P. Diggle, A. S. Griffin, G. S. Campbell, and S. A. West. Cooperation and conflict in quorum-sensing bacterial populations. *Nature* **450**(7168), 411–414, 2007. DOI: 10.1038/nature06279.

[84] X. He, W. Chang, D. L. Pierce, L. O. Seib, J. Wagner, and C. Fuqua. Quorum sensing in *Rhizobium sp.* strain NGR234 regulates conjugal transfer (tra) gene expression and influences growth rate. *Journal of Bacteriology* **185**(3), 809–822, 2003. DOI: 10.1128/JB.185.3.809-822.2003.

[85] P. A. P. Moran. Random processes in genetics. *Mathematical Proceedings of the Cambridge Philosophical Society* **54**, 60–71, 1958. DOI: 10.1017/S0305004100033193.

[86] W. J. Ewens. *Mathematical population genetics*. 2nd edition. Springer, New York, 2004. DOI: 10.1007/978-0-387-21822-9.

[87] R. A. Blythe and A. J. McKane. Stochastic models of evolution in genetics, ecology and linguistics. *Journal of Statistical Mechanics: Theory and Experiment* **2007**(07), P07018, 2007. DOI: 10.1088/1742-5468/2007/07/P07018.

[88] A. Traulsen, J. C. Claussen, and C. Hauert. Coevolutionary dynamics: From finite to infinite populations. *Physical Review Letters* **95**, 238701, 2005. DOI: 10.1103/PhysRevLett.95.238701.

[89] A. Melbinger, J. Cremer, and E. Frey. Evolutionary game theory in growing populations. *Physical Review Letters* **105**, 178101, 2010. DOI: 10.1103/PhysRevLett.105.178101.

[90] M. Assaf, M. Mobilia, and E. Roberts. Cooperation dilemma in finite populations under fluctuating environments. *Physical Review Letters* **111**, 238101, 2013. DOI: 10.1103/PhysRevLett.111.238101.

[91] E. Kussell and S. Leibler. Phenotypic diversity, population growth, and information in fluctuating environments. *Science* **309**(5743), 2075–2078, 2005. DOI: 10.1126/science.1114383.

[92] M. Acar, J. T. Mettetal, and A. van Oudenaarden. Stochastic switching as a survival strategy in fluctuating environments. *Nature Genetics* **40**(4), 471–475, 2008. DOI: 10.1038/ng.110.

[93] K. Axelrod, A. Sanchez, and J. Gore. Phenotypic states become increasingly sensitive to perturbations near a bifurcation in a synthetic gene network. *eLife* **4**, e07935, 2015. DOI: 10.7554/eLife.07935.

[94] B. Drees, M. Reiger, K. Jung, and I. B. Bischofs. A modular view of the diversity of cell-density-encoding schemes in bacterial quorum-sensing systems. *Biophysical Journal* **107**(1), 266–277, 2014. DOI: 10.1016/j.bpj.2014.05.031.

[95] T. Maire and H. Youk. Molecular-level tuning of cellular autonomy controls the collective behaviors of cell populations. *Cell Systems* **1**(5), 349–360, 2015. DOI: 10.1016/j.cels.2015.10.012.

[96] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics* **22**(4), 403–434, 1976. DOI: 10.1016/0021-9991(76)90041-3.

[97] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry* **81**(25), 2340–2361, 1977. DOI: 10.1021/j100540a008.

[98] G. Wild and A. Traulsen. The different limits of weak selection and the evolutionary dynamics of finite populations. *Journal of Theoretical Biology* **247**(2), 382–390, 2007. DOI: 10.1016/j.jtbi.2007.03.015.

[99] H. Hinrichsen. Non-equilibrium critical phenomena and phase transitions into absorbing states. *Advances in Physics* **49**(7), 815–958, 2000. DOI: 10.1080/00018730050198152.

[100] M. Kadar. *Statistical physics of particles*. 3rd edition. Cambridge University Press, Cambridge, UK, 2007. DOI: 10.1017/CBO9780511815898.

[101] I. M. Bomze. Dynamical aspects of evolutionary stability. *Monatshefte für Mathematik* **110**(3), 189–206, 1990. DOI: 10.1007/BF01301675.

[102] J. Oechssler and F. Riedel. Evolutionary dynamics on infinite strategy spaces. *Economic Theory* **17**(1), 141–162, 2001. DOI: 10.1007/PL00004092.

[103] J. Hofbauer and K. Sigmund. Evolutionary game dynamics. *Bulletin of the American Mathematical Society* **40**(4), 479–519, 2003. DOI: 10.1090/S0273-0979-03-00988-1.

[104] R. Cressman. Stability of the replicator equation with continuous strategy space. *Mathematical Social Sciences* **50**(2), 127–147, 2005. DOI: 10.1016/j.mathsocsci.2005.03.001.

[105] B. J. McGill and J. S. Brown. Evolutionary game theory and adaptive dynamics of continuous traits. *Annual Review of Ecology, Evolution, and Systematics* **38**, 403–435, 2007. DOI: 10.1146/annurev.ecolsys.36.091704.175.

[106] V. Elgart and A. Kamenev. Rare event statistics in reaction-diffusion systems. *Physical Review E* **70**, 041106, 2004. DOI: 10.1103/PhysRevE.70.041106.

[107] D. A. Kessler and N. M. Shnerb. Extinction rates for fluctuation-induced metastabilities: A real-space WKB approach. *Journal of Statistical Physics* **127**(5), 861–886, 2007. DOI: 10.1007/s10955-007-9312-2.

[108] M. Assaf and B. Meerson. Extinction of metastable stochastic populations. *Physical Review E* **81**, 021116, 2010. DOI: 10.1103/PhysRevE.81.021116.

[109] P. Hänggi. Escape from a metastable state. *Journal of Statistical Physics* **42**(1), 105–148, 1986. DOI: 10.1007/BF01010843.

[110] M. R. Parsek and E. Greenberg. Sociomicrobiology: The connections between quorum sensing and biofilms. *Trends in Microbiology* **13**(1), 27–33, 2005. DOI: 10.1016/j.tim.2004.11.007.

[111] C. Schmeisser, H. Liesegang, D. Krysciak, N. Bakkou, A. Le Quéré, A. Wollherr, I. Heinemeyer, B. Morgenstern, A. Pommerening-Röser, M. Flores, R. Palacios, S. Brenner, G. Gottschalk, R. A. Schmitz, W. J. Broughton, X. Perret, A. W. Strittmatter, and W. R. Streit. *Rhizobium sp.* strain NGR234 possesses a remarkable number of secretion systems. *Applied and Environmental Microbiology* **75**(12), 4035–4045, 2009. DOI: 10.1128/AEM.00515-09.

[112] E. A. Yates, B. Philipp, C. Buckley, S. Atkinson, S. R. Chhabra, R. E. Sockett, M. Goldner, Y. Dessaux, M. Cámara, H. Smith, and P. Williams. N-acylhomoserine lactones undergo lactonolysis in a pH-, temperature-, and acyl chain length-dependent manner during growth of *Yersinia pseudotuberculosis* and *Pseudomonas aeruginosa*. *Infection and Immunity* **70**(10), 5635–5646, 2002. DOI: 10.1128/IAI.70.10.5635-5646.2002.

[113] J. T. Byers, C. Lucas, G. P. C. Salmond, and M. Welch. Nonenzymatic turnover of an *Erwinia carotovora* quorum-sensing signaling molecule. *Journal of Bacteriology* **184**(4), 1163–1171, 2002. DOI: 10.1128/jb.184.4.1163-1171.2002.

[114] A. W. Decho, P. T. Visscher, J. Ferry, T. Kawaguchi, L. He, K. M. Przekop, R. S. Norman, and R. P. Reid. Autoinducers extracted from microbial mats reveal a surprising diversity of N-acylhomoserine lactones (AHLs) and abundance changes that may relate to diel pH. *Environmental Microbiology* **11**(2), 409–420, 2009. DOI: 10.1111/j.1462-2920.2008.01780.x.

[115] C. Grandclément, M. Tannières, S. Moréra, Y. Dessaux, and D. Faure. Quorum quenching: Role in nature and applied developments. *FEMS Microbiology Reviews* **40**(1), 86, 2016. DOI: 10.1093/femsre/fuv038.

[116] L. R. Hmelo. Quorum sensing in marine microbial environments. *Annual Review of Marine Science* **9**(1), 257–281, 2017. DOI: 10.1146/annurev-marine-010816-060656.

[117] B. LaSarre and M. J. Federle. Exploiting quorum sensing to confuse bacterial pathogens. *Microbiology and Molecular Biology Reviews* **77**(1), 73–111, 2013. DOI: 10.1128/MMBR.00046-12.

[118] T. G. Platt and C. Fuqua. What's in a name? The semantics of quorum sensing. *Trends in Microbiology* **18**(9), 383–387, 2010. DOI: 10.1016/j.tim.2010.05.003.

[119] J. E. González and M. M. Marketon. Quorum sensing in nitrogen-fixing *Rhizobia*. *Microbiology and Molecular Biology Reviews* **67**(4), 574–592, 2003. DOI: 10.1128/MMBR.67.4.574-592.2003.

[120] A. Novick and M. Weiner. Enzyme induction as an all-or-none phenomenon. *Proceedings of the National Academy of Sciences of the United States of America* **43**(7), 553–566, 1957. DOI: 10.1073/pnas.43.7.553.

[121] E. M. Ozbudak, M. Thattai, H. N. Lim, B. I. Shraiman, and A. van Oudenaarden. Multistability in the lactose utilization network of Escherichia coli. *Nature* **427**(6976), 737–740, 2004. DOI: 10.1038/nature02298.

[122] M. Kaern, T. C. Elston, W. J. Blake, and J. J. Collins. Stochasticity in gene expression: From theories to phenotypes. *Nature Reviews Genetics* **6**(6), 451–464, 2005. DOI: 10.1038/nrg1615.

[123] D. Dubnau and R. Losick. Bistability in bacteria. *Molecular Microbiology* **61**(3), 564–572, 2006. DOI: 10.1111/j.1365-2958.2006.05249.x.

[124] W. K. Smits, O. P. Kuipers, and J.-W. Veening. Phenotypic variation in bacteria: The role of feedback regulation. *Nature Reviews Microbiology* **4**(4), 259–271, 2006. DOI: 10.1038/nrmicro1381.

[125] A. Raj and A. van Oudenaarden. Nature, nurture, or chance: Stochastic gene expression and its consequences. *Cell* **135**(2), 216–226, 2008. DOI: 10.1016/j.cell.2008.09.050.

[126] S. Pollak, S. Omer-Bendori, E. Even-Tov, V. Lipsman, T. Bareia, I. Ben-Zion, and A. Eldar. Facultative cheating supports the coexistence of diverse quorum-sensing alleles. *Proceedings of the National Academy of Sciences of the United States of America* **113**(8), 2152–2157, 2016. DOI: 10.1073/pnas.1520615113.

[127] A. A. Dandekar, S. Chugani, and E. P. Greenberg. Bacterial quorum sensing and metabolic incentives to cooperate. *Science* **338**(6104), 264–266, 2012. DOI: 10.1126/science.1227289.

[128] T. Czárán and R. F. Hoekstra. Microbial communication, cooperation and cheating: Quorum sensing drives the evolution of cooperation in bacteria. *PLOS ONE* **4**(8), 1–10, 2009. DOI: 10.1371/journal.pone.0006655.

[129] E. C. Carnes, D. M. Lopez, N. P. Donegan, A. Cheung, H. Gresham, G. S. Timmins, and C. J. Brinker. Confinement-induced quorum sensing of individual *Staphylococcus aureus* bacteria. *Nature Chemical Biology* **6**(1), 41–45, 2010. DOI: 10.1038/nchembio.264.

[130] W. Braun and K. Hepp. The Vlasov dynamics and its fluctuations in the $1/N$ limit of interacting classical particles. *Communications in Mathematical Physics* **56**(2), 101–113, 1977. DOI: 10.1007/BF01611497.

[131] R. L. Dobrushin. Vlasov equations. *Functional Analysis and Its Applications* **13**(2), 115–123, 1979. DOI: 10.1007/BF01077243.

[132] H. Spohn. *Large scale dynamics of interacting particles*. Springer, Heidelberg, 1991. DOI: 10.1007/978-3-642-84371-6.

[133] G. Doetsch. *Einführung in Theorie und Anwendung der Laplace Transformation*. Vol. 3. Birkhaeuser Verlag Basel und Stuttgart, 1976. DOI: 10.1007/978-3-0348-5188-6.

Acknowledgments

First and foremost, I would like to use this opportunity to express my deepest gratitude to my supervisor Erwin Frey for giving me the opportunity to work with him and in his group at the Ludwig-Maximilians-Universität in Munich. I have been appreciating your constant support and your trust; especially during early stages of my PhD and in difficult situations. I thank you for making it possible for me to participate in the scientific life, whether it be by connecting me with other researchers to discuss our scientific results, by giving me the opportunity to present our results at numerous conferences and meetings, or by supporting me with applications for several summer schools (Hopkins Microbiology Course 2013 in Monterey, QBio Summer School 2014 at the KITP in Santa Barbara, Boulder Summer School 2015 in Boulder) and other scientific trips from which I greatly benefited both scientifically and personally in several aspects. In particular, I have been enjoying the freedom to follow all of my scientific interests and to explore new project ideas with you. Discussing with you about science and beyond has always been both insightful and really good fun.

Special thanks also go to Markus F. Weber with whom I have shared an office for several years. Not only did you help me with computing techniques and numerical issues of all possible kinds, and proofreading of my texts at any stages, but you also helped me with the layout of this thesis. Moreover, I always appreciated the joyful and creative atmosphere between us, in our office even during intense times of our research, and beyond. From a scientific point of view, I am particularly proud that we have developed together our scientific standards and an open feedback culture. Along the same lines, I want to sincerely thank Torben Krüger for his patience in explaining me analysis, already during my time as a master student, and for our ongoing discussions about physics. Our very first common project on the antisymmetric Lotka-Volterra equation has started during a Mensa discussion already half a decade ago! I would like to thank Matthias Lechner with whom I have had the pleasure to cooperate and to work during the master projects of the great students Stefano Duca and Matthias Bauer. Our project with Matthias Bauer flourished in several directions and parts of our results are now close to publication. I would like to thank Matthias Bauer especially for his patience and endurance with this project. Furthermore, I had the luck to meet Peter Pickl in his lecture on effective equations in physics. It was great for me to see how, starting out from chatting about issues on the lecture, a fruitful collaboration has begun and has been continuing up to date on the quorum-sensing model. It is always a pleasure for me to discuss with you.

Furthermore, I want to thank my former colleague Steffen Rulands, with whom I had the pleasure to accompany Raphaela Gessele and Frank Schlosser during their master projects. During my time as a PhD student, I very much enjoyed to interact with formidable students, such as Volker Karle, Daniel Kunz, Lukas Darnstädt, Tobias Hermann, Tobias Göppel, and Hannes Herrmann during their study or bachelor projects; and Florian Gartner, Philipp Geiger, and Oriol Servera during their master projects. It is exciting for me to see how your scientific and non-scientific paths move forward and how our paths have already crossed again! I thank the experimentalists Madeleine Opitz, Andreas Mader, and Sara Kesel for stimulating my curiosity about microbiology, their guidance, and their patience during my hands-on trials at the bench in the lab. I also want to thank all of the present and previous members of Erwin's chair for creating and shaping such a joyful atmosphere in the office, during Mensa times, and during free-time activities. Special thanks in this direction (besides the aforementioned people) go to Matthias Rank, Isabella Graf, Felix Kempf, Silke Bergeler, Marianne Bauer, Patrick Wilke, Emanuel Reithmann, Laeschkir Hassan, Karl Wienand, Lorenz Huber, Jonas Denk, Jacob Halatek, Chase Broedersz, Severin Schink, Louis Reese, Cornelius Weig, Cristina Pop, and Carolina Lesperance.

My thanks also go to Falk Töppel, Jasmin Reusch, and again to Markus F. Weber for their comments on the text of this thesis or on parts of it. Finally, I would like to thank my friends who supported me during the time as a PhD student in a direct or indirect way, both for creating a relaxing distance to my PhD studies and, at the same time, for giving me inspiration for my PhD projects.

Außerdem möchte ich die Gelegenheit nutzen, meinen Eltern, meinem Bruder Paul und meinen Großeltern von ganzem Herzen dafür zu danken, dass sie mich mit ihrer liebenswerten Art und ihrer Herzlichkeit bei all meinen Vorhaben über all die Jahre und in all möglichen Formen unterstützt haben und mir immer Rückhalt gegeben haben. Herzlichsten Dank!