Evolutionary dynamics from a variational principle

Peter Klimek,[1] Stefan Thurner,[1,2,*] and Rudolf Hanel[1]

[1]*Section for Science of Complex Systems, Medical University of Vienna, Spitalgasse 23, A 1090 Vienna, Austria*
[2]*Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA*

We demonstrate with a thought experiment that fitness-based population dynamical approaches to evolution are not able to make quantitative, falsifiable predictions about the long-term behavior of some evolutionary systems. A key characteristic of evolutionary systems is the ongoing endogenous production of new species. These novel entities change the conditions for already existing species. Even *Darwin's Demon*, a hypothetical entity with exact knowledge of the abundance of all species and their fitness functions at a given time, could not prestate the impact of these novelties on established populations. We argue that fitness is always *a posteriori* knowledge—it measures but does not explain why a species has reproductive success or not. To overcome these conceptual limitations, a variational principle is proposed in a spin-model-like setup of evolutionary systems. We derive a functional which is minimized under the most general evolutionary formulation of a dynamical system, i.e., evolutionary trajectories causally emerge as a minimization of a functional. This functional allows the derivation of analytic solutions of the asymptotic diversity for stochastic evolutionary systems within a mean-field approximation. We test these approximations by numerical simulations of the corresponding model and find good agreement in the position of phase transitions in diversity curves. The model is further able to reproduce stylized facts of timeseries from several man-made and natural evolutionary systems. Light will be thrown on how species and their fitness landscapes dynamically coevolve.

I. INTRODUCTION

Evolutionary dynamics appear in a multitude of different contexts. Evolution basically describes how sets of elements, such as biological species, goods and services in an economy, groups of living beings, or chemical compounds, change over time. Examples are abundant in various areas. Chemical compounds react with other compounds to produce new chemicals. Integrated circuits performing specific computational tasks can be combined to create another circuit for a different computational task. Prey and predator may coevolve by succinctly acquiring new traits and thereby develop into new species. In the following we will use *species* for elements in whatever context, chemicals, goods, biological species, etc. The removal or addition of a single species in an evolutionary system may have dramatic consequences. For example, in starfish removal experiments (e.g., Mukkaw Bay in Washington [1]) starfish are removed from an ecosystem with the consequences that mussel populations explode and drive out most other species, while the urchin population destroys coral reefs. In 1904 English physicist John Ambrose Fleming accidentally manufactured the first vacuum tube which triggered a cascade of technological and economic coevolutions and adaptations; in 2004 the semiconductor industry was a market of $213 billion and enabled the generation of approximately $1200 billion in electronic system businesses and $5000 billion in turn in services which amounts to 10% of world GDP [2]. Typically in evolutionary systems species are endogenously added or removed from a large system of mutually influencing species. Two species influence each other if the existence of one species has a positive or negative effect on the change of abundance of the other. The possibilities for interactions in evolutionary systems involve different natural, economic or social laws on a variety of time or length scales. The collective result of these ("microscopic") interactions between elements leads to ubiquitous well-known macro phenomena in evolutionary systems, such as punctuated equilibria, booms of diversification, breakdowns and crashes, or seemingly unpredictable responses to external perturbations. Maybe one of the most exciting questions in natural sciences today is to understand if evolutionary dynamics can be understood by a common underlying principle and—if yes—how such a principle might look like. Such a principle must be general enough to capture the multitude of different phenomena, and at the same time must be in a form which can be applied easily to specific problems.

In the present understanding of evolution the concept of *fitness* is of central importance. Usually the relative abundance of species (with respect to other species) is described by replicator equations (first-order differential equations) and variants such as Lotka-Volterra equations [3–5]. Their mutual influence is quantified by a rate specifying how the abundance of one species changes upon contact with another. In biology this rate is called Malthusian fitness, in chemistry one refers to it as the reaction rate, in economics it is related to production functions. Similar proliferation rates could also be introduced for technological, financial or even historical contexts. In the following we subsume them all under the term *fitness*.

A distance between two species can be defined as the minimal number of evolutionary steps needed for one species to evolve into the other one [6] (in biology this distance is often the number of single-point mutations two species differ in). In this way a metric is given on the space of all possible species. A *fitness landscape* assigns to each point in

_____
*thurner@univie.ac.at

this space (that is to each species) its reproductive success or fitness. Evolution is sometimes pictured as an optimization problem where species evolve and coevolve via adaptations toward peaks in this landscape [7]. These peaks represent regions of high reproductive success, whereas valleys correspond to low fitnesses. As these peaks and valleys are spread over larger regions of species space, the fitness landscape becomes less rugged. If one species moves toward a peak it may change the fitness of other species which are then moving on this fitness landscape too [8]—one speaks of coevolutionary cascades. The higher the degree of "ruggedness" (i.e., the smaller the average distance between adjacent adaptive peaks or valleys), the higher the probability that these cascades of coevolution last for long times. With lower ruggedness, however, the probability increases that each species reaches an adaptive peak and the evolutionary dynamics comes to a halt.

The fitness concept turned out to be of practical use in specific fields such as viral competition experiments [9,10] where certain assumptions such as high turnover rates, constant death rates or restricted population sizes are met. Here fitness functions could be used to quantify the relative replicative capacity of two competing virus variants.

The concept of fitness is limited however. To see this consider the following thought experiment. Suppose one—say a demon—would have exact knowledge about the abundance and fitness of each biological species in the universe. "Knowing the fitness of a species" means knowledge of the functional dependence of its proliferation rate on the entire current environment (i.e., all other species). The omniscient hypothetical entity in possession of this knowledge could be called *Darwin's Demon* for obvious reasons. The demon may be pictured as a superbiologist, able to measure each species' abundance as well as the dependence of its proliferation rate on each other species in each habitat. That is, he knows the set of all existing species and can measure their associated fitness landscape to an arbitrary degree of exactness. What can the demon *predict* about the future course of evolutionary events, such as biodiversity in 100 million years or the time to the next mass extinction event? Surprisingly little, for the following reasons. A key characteristic of evolutionary systems is its potential to generate innovations, i.e., new species. In biological systems this can happen through mutations, in technological or economical ones through spontaneous ideas of an inventor, etc. Once a new species is created it becomes part of the environment and thereby potentially changes the conditions for all already existing species and for those yet to arrive. To now assess the fitness of a new species one has to measure how it spreads in an environment it is now part of. The demon has information related to a different environment, one which only existed *before* arrival of the new species. Thus the demon may have an exact description of the *current* biosphere, but with the advent of each new species this description loses accuracy. Fitness thus always encodes *a posteriori* knowledge, and cannot be used to make falsifiable predictions. It is not fruitful to predict future fitness of species from their present fitness. Instead one has to understand how species and their fitness landscapes coconstruct each other, how they *coevolve*.

To make headway in understanding the phenomenology of evolution, i.e., in identifying principles which guide evo-

lutionary dynamics, a series of quantitative models have been suggested [8,11–18]. Here explicit assumptions are made about how new species come into being, how they interact with each other and under which conditions or under which selective forces they vanish. Each of these models focuses on particular aspects of evolution. For example in Kauffman *NK* models [8] species are bit-strings with randomly assigned fitness values. Arthur [11] focuses on technological evolution with integrated circuits as species whose fitness is examined by how well they execute certain computational tasks. Jain and Krishna [12] consider ecological systems and elucidate the interplay between interaction topology and survival of species. In most of these models some *ad hoc* assumptions about the mechanisms have to be made. In the model by Jain and Krishna, for example, species are actively removed and added to the system, innovations are externally enforced and not endogenously produced. In [11] the output of randomly assembled circuits is compared to a prespecified list of desired computational tasks (such as bitwise addition). In the *NK* model evolutionary interactions are constrained to actions on bit-strings. Although these assumptions are certainly reasonable in the specific contexts of their models, it is not at all clear whether conclusions derived on the basis of these assumptions are universally valid in different contexts.

To arrive at a general evolutionary description (without *ad hoc* specifications) one has to identify principles which are abstract enough to be applicable in each evolutionary context but which must be specific enough to make useful quantitative predictions. To meet these requirements, evolution can be pictured as a three-step production/destruction process. Step 1: New species come into being through recombination of already present species. That is, each species arises only under the condition that a given (and maybe not unique) set of other species or environmental factors exists. For example, to assemble an MP3-player all parts—including software—are needed. Sodium chloride can be produced by sodium hydroxide in solution with hydrochloric acid. With a substantial degree of oversimplification one can say that apes in combination steppe formation give rise to mankind. Step 2: The new species becomes part of the system and can now be combined with other, already existing species. One can legally download music for the MP3-player and listen to it, sodium chloride reacts with e.g., calcium carbonate in the Solvay process, mankind burns forests to create fields for agriculture. Step 3: As a consequence, through this recombination yet new species may come into being and other already existing ones may vanish or be destroyed. To be more precise, the abundance of the new species has now to be included in the already existing species' fitness function, i.e., their proliferation rates have to be remeasured or calculated. In this way—as a secondary effect—the extinction or expansion of other species may depend on the advent of new ones. For example, MP3 currently drives CDs out of the market but can be combined with cell phones to give smartphones. Soda ash can be used to remove sulfur dioxide from flue gases in power stations which might help to reduce the ongoing Holocene extinction event of biological species [19] possibly influenced by the advent of mankind. In previous work models incorporating these types of production and

destruction processes have been shown to reproduce a wide range of evolutionary phenomena, including booms of diversification [20], breakdowns of diversity [21] or punctuated equilibria [22]. Such processes further allow to understand stylized facts in time-series data on evolutionary systems, such as scale-free distributions of species lifetimes, the number of species per genus or the size of extinction events in fossil data [23], or GDP and business failures in economic markets [24].

In this work we propose a variational principle from which dynamics—identical to the dynamics of the production/destruction processes described above—can be derived. To this end we define the *evolutionary potential* of a species. This function measures in how many productions and destructions a species would (no longer) take part if it would enter (be removed from) the system. With this potential one obtains two formal representations of the system's dynamic: (i) The potential can be used to explicitly deduce a set of dynamical *update equations* of system diversity for production/destruction processes. (ii) Using this evolutionary potential and a measure for ongoing productions and destructions one can derive a *balance function*. The evolutionary process solving the dynamical update equations (i) always minimizes the balance function (ii). The balance function further allows asymptotic solutions for the system diversity (mean-field approximation). These analytic solutions are in good agreement with numerical simulations of the full model of productions and destructions. This is to a certain degree unexpected since the dynamics is dominated by strong and nonlinear interactions.

This description of evolutionary systems allows to understand how the set of existing species and their fitness landscapes coconstruct each other from first principles, as opposed to research strategies portrayed by Darwin's Demon, where snapshots of regions of fitness landscapes are empirically explored. Accordingly the focus shifts from predicting *microscopic* properties such as individual proliferation rates to estimating the occurrence of global, *macroscopic* events. Here we develop a formalism capable of showing (i) macroevolutionary phase transitions (booms of diversification, extinction events or punctuated equilibria), (ii) fat-tailed statistics of many evolutionary observables as found in real-world systems (such as species lifetimes, metabolic reaction frequencies or economic productivity) and (iii) how to derive these evolutionary dynamic equations from a variational principle given (statistical) knowledge of the topology of production/destruction rules only.

This work is structured as follows. In Sec. II we develop a general framework for evolutionary systems via a variational principle. We discuss deterministic and stochastic implementations and obtain asymptotic diversity solutions in a mean-field approximation. In Sec. III we motivate and define the choice of evolutionary interactions as production and destruction rules as in [24]. We treat the special cases of systems with only productive interactions in Sec. IV and the pure destructive case in Sec. V. Then we discuss the full model of productions *and* destructions in Sec. VI. We discuss empirical relevance of this work in Sec. VII and turn to a conclude in Sec. VIII.

II. GENERAL FORMULATION OF DIVERSITY DYNAMICS

A. Dynamical systems

The abundance of species i is given by a binary state variable $\sigma_i(t) \in \{0,1\}$. If species i exists at time t, $\sigma_i(t)=1$, otherwise $\sigma_i(t)=0$. The system can be populated by N species (N arbitrarily large, even infinite). A particular configuration of the system is characterized by the N-dimensional vector in phase space $\vec{\sigma}(t)=[\sigma_i(t)] \in \Gamma=(0,1)^N$. The system's *diversity* $D(t)$ is given by $D(t)=\frac{1}{N}\Sigma_i\sigma_i(t)$.

At each time, species i may experience three scenarios, (i) annihilation $\sigma_i(t)=1 \rightarrow \sigma_i(t+1)=0$, (ii) nothing $\sigma_i(t)=\sigma_i(t+1)$ or (iii) creation $\sigma_i(t)=0\rightarrow\sigma_i(t+1)=1$. Suppose that there exists a function $f_i[\vec{\sigma}(t)]:(0,1)^N \rightarrow \mathbb{R}$ indicating which of the transitions (i)–(iii) takes place. Specifically, let $f_i[\vec{\sigma}(t)]$ indicate the following transitions

$$(\text{i})f_i[\vec{\sigma}(t)] < 0 \Rightarrow \sigma_i(t+1) = 0,$$

$$(\text{ii})f_i[\vec{\sigma}(t)] = 0 \Rightarrow \sigma_i(t+1) = \sigma_i(t),$$

$$(\text{iii})f_i[\vec{\sigma}(t)] > 0 \Rightarrow \sigma_i(t+1) = 1, \quad (1)$$

For (i) or (iii) a transition occurs if $\sigma_i(t)=1$ or 0, respectively. That is, if $f_i(\vec{\sigma}(t))\geq 0$ the system evolves according to

$$\sigma_i(t+1) = \sigma_i(t) + \Delta\sigma_i(t) \quad \text{with}$$

$$\Delta\sigma_i(t) = \text{sgn}\{[1-\sigma_i(t)]f_i[\vec{\sigma}(t)]\}. \quad (2)$$

$\Delta\sigma_i(t)$ can only be nonzero if $\sigma_i(t)=0$ *and* $f_i[\vec{\sigma}(t)]>0$. Similarly, for $f_i[\vec{\sigma}(t)]\leq 0$ $\Delta\sigma_i(t)=\text{sgn}\{-\sigma_i(t)f_i[\vec{\sigma}(t)]\}$. Let us define the ramp function $R(x)$ by $R(x)\equiv x$ iff $x\geq 0$ and $R(x)\equiv 0$ iff $x<0$. Using these definitions we can generically map the indicator function f_i from Eq. (1) onto the update equation

$$\sigma_i(t+1) = \sigma_i(t) + \Delta\sigma_i(t),$$

$$\Delta\sigma_i(t) = \text{sgn}[1-\sigma_i(t)]R\{f_i[\vec{\sigma}(t)]\} - \sigma_i(t)R\{-f_i[\vec{\sigma}(t)]\}. \quad (3)$$

B. Variational principle for deterministic diversity dynamics

We introduce a distance function to quantify the number of state changes in the system. Consider a virtual displacement of $\sigma_i(t)$, $\sigma_i'(t)=\sigma_i(t)+\delta\sigma_i(t)$. A quadratic distance function is given by

$$K_i(\sigma_i'(t),\sigma_i(t)) \equiv \frac{\mu}{2}[\sigma_i'(t) - \sigma_i(t)]^2, \quad (4)$$

where $\mu>0$.[1] Analogously a *potential* V_i is defined by

---

[1]Note the similarity to kinetic energy in classical mechanics.

TABLE I. We exhaustively insert all possible values for $\sigma_i(t)$ and $\sigma_i'(t)$ in Eq. (3) and B_i for nonzero values of $f_i[\vec{\sigma}(t)]$. For convenience we choose $|f_i[\vec{\sigma}(t)]|=1$ and confine the threshold to $0<\mu<2$. We mark the values of $\sigma_i'(t)$ for which B_i is a minimum by an underscore. For these marked values $\sigma_i'(t)$ is always equal to $\sigma_i(t+1)$.

$\sigma_i(t)$	$f_i[\vec{\sigma}(t)]$	$\sigma_i'(t)$	K_i	V_i	B_i	$\sigma_i(t+1)$
0	-1	$\underline{0}$	0	0	0	0
0	-1	1	$\frac{\mu}{2}$	1	$1+\frac{\mu}{2}$	0
0	1	0	0	1	1	1
0	1	$\underline{1}$	$\frac{\mu}{2}$	0	$\frac{\mu}{2}$	1
1	-1	$\underline{0}$	$\frac{\mu}{2}$	0	$\frac{\mu}{2}$	0
1	-1	1	0	1	1	0
1	1	0	$\frac{\mu}{2}$	1	$1+\frac{\mu}{2}$	1
1	1	$\underline{1}$	0	0	0	1

$$V_i[\sigma_i'(t),\vec{\sigma}(t)] \equiv \left| [1-\sigma_i'(t)]R\{f_i[\vec{\sigma}(t)]\} - \sigma_i'(t)R\{-f_i[\vec{\sigma}(t)]\} \right|, \tag{5}$$

which "counts" the number of possible interactions for the displaced state $\sigma_i'(t)$. Depending on $\sigma_i'(t)$, Eq. (5) will reduce to $V_i[\sigma_i'(t),\vec{\sigma}(t)] = |R\{\pm f_i[\vec{\sigma}(t)]\}|$. A possible intuition behind Eq. (5) is that f_i acts as a "field" on $\sigma_i(t)$ which is "probed" by $\sigma_i'(t)$. We occasionally drop the σ-dependence for a clearer notation. Finally we define the *balance function*, $B_i \equiv K_i+V_i$.

K_i measures the *actual* activity in the system—it counts all state changes. The potential V_i counts the *potential* activity in the newly obtained states. B_i therefore contains the full dynamical information of Eq. (1) which can now be expressed through a *variational principle*.

Given $\vec{\sigma}(t)$, the solution $\sigma_i(t+1)$ of Eq. (3) is identical to the value of $\sigma_i'(t)$ for which B_i assumes its minimum, i.e.,

$$\sigma_i(t+1) = \arg\min_{\sigma_i'(t)}\{B_i[\sigma_i'(t),\vec{\sigma}(t)]\}, \tag{6}$$

with $\arg\min_x[f(x)]$ denoting the value of x for which $f(x)$ takes its minimum.

This is proved by exhaustive insertion. First, consider the case $f_i[\vec{\sigma}(t)]=0$. From Eq. (3) it follows that $\Delta\sigma_i(t)=0$ and $V_i=0$. The only possible term contributing to B_i is K_i; $K_i=0$ if $\sigma_i'(t)=\sigma_i(t)$ and $K_i=\frac{\mu}{2}>0$ otherwise. The balance function B_i takes its minimum, $B_i=0$ at $\sigma_i'(t)=\sigma_i(t)=\sigma_i(t+1)$. Similar reasoning can be applied to the cases of nonzero $f_i[\vec{\sigma}(t)]$, see Table I. This now clarifies the role of the parameter μ. It can be seen as an inertial threshold; the dynamics of Eq. (3) only takes place if the "field" $f_i[\vec{\sigma}(t)]$ describing a certain state-change exceeds the barrier set by μ. There is always a choice for μ such that Eq. (3) holds.

C. Stochastic diversity dynamics

There exists a natural stochastic variant of diversity dynamics. In Eq. (3) a state transition $\sigma_i(t) \to \sigma_i(t+1)$ is *determined* by $\Delta\sigma_i(t) \in \{-1,0,1\}$. For the stochastic case we specify *transition probabilities* for this evolution.

From the variational principle Eq. (6) it follows that Eq. (3) always minimizes the balance function B_i. In the stochastic variant we assume that the lower B_i, the higher is the probability to find the system in the respective configuration $\sigma_i(t)$. In analogy to spin systems this probability is a Boltzmann factor

$$p[\sigma_i(t)] \propto e^{-\beta B_i[\vec{\sigma}(t)]}, \tag{7}$$

with $\beta \equiv 1/T$ the inverse temperature. To obtain transition probabilities we demand detailed balance

$$\frac{p[\sigma_i(t) \to \hat{\sigma}_i(t)]}{p[\hat{\sigma}_i(t) \to \sigma_i(t)]} = \frac{p[\hat{\sigma}_i(t)]}{p[\sigma_i(t)]} = e^{-\beta(\hat{B}_i - B_i)}, \tag{8}$$

with $\hat{B}_i \equiv B_i[\hat{\sigma}_i(t),\sigma(t)_{j\neq i}]$. There are several ways to choose transition probability such that Eq. (8) is satisfied, here we use Metropolis transition probabilities $p[\sigma_i(t) \to \hat{\sigma}_i(t)]=1$ if $\hat{B}_i - B_i < 0$ and $p[\sigma_i(t) \to \hat{\sigma}_i(t)]=\exp[-\beta(\hat{B}_i - B_i)]$ otherwise. The stochastic diversity dynamics is fully specified by setting[2] $\sigma_i(t+1)=\hat{\sigma}_i(t)$.

Whereas in the deterministic case the balance function B_i is minimized, the stochastic diversity dynamics shows "disordering effects" due to nonzero temperature T as given in Eq. (7). We quantify this with Boltzmann-Gibbs entropy.

D. Mean-field approximation

Denote the expectation value of $\sigma_i(t)$ by $q_i(t)=\langle\sigma_i(t)\rangle$ and assume that the probability distribution factorizes, i.e., $p[\vec{\sigma}(t)]=\Pi_i p_i[\sigma_i(t)]$ with $p_i[\sigma_i(t)]=[1-q_i(t)]\delta_{\sigma_i(t),0} + q_i(t)\delta_{\sigma_i(t),1}$. In this mean-field approximation the Boltzmann-Gibbs entropy s for species i is

$$s[\sigma_i(t)] = -\langle\ln p_i[\sigma_i(t)]\rangle \equiv s[q_i(t)],$$

$$s[q_i(t)] = -[1-q_i(t)]\ln[1-q_i(t)] - q_i(t)\ln q_i(t). \tag{9}$$

The 'free energy' functional $\phi[\sigma_i(t)]$ for the system turns for this approximation into

$$\phi[q_i(t)] = \langle B_i\rangle_{p[\vec{\sigma}(t)]} - \frac{s[q_i(t)]}{\beta}. \tag{10}$$

The asymptotic state of species i, $q_i(t \to \infty) \equiv q_i$, is identified by a minimum in free energy. The necessary condition for this, $\partial\phi(q_i)/\partial q_i=0$, is $\frac{\partial\langle B_i\rangle}{\partial q_i} + \frac{1}{\beta}\ln(\frac{q_i}{1-q_i})=0$, and

$$q_i = \frac{1}{2}\left\{\tanh\left[-\frac{\beta}{2}\frac{\partial\langle B_i\rangle}{\partial q_i}\right] + 1\right\}. \tag{11}$$

The self-consistent solution of Eq. (11) yields the asymptotic configuration.

---

[2] One can also define dynamics "backward" by the transition probability $p[\hat{\sigma}_i(t) \to \sigma_i(t)]$ which can be interpreted as inferring $\vec{\sigma}(t)$ from the knowledge of $\vec{\sigma}(t+1)$.

III. GENERAL FORMULATION OF EVOLUTIONARY INTERACTIONS

Traditionally in the master equations framework[3] interactions are classified by transfer rates for abundances of species. The transfer rates measure how the change in abundance of a given species i is related to the abundance of other species $j_1, j_2, \ldots$. Depending on how i and $j_1, j_2, \ldots$ are chosen, one obtains different systems of differential equations which can be related to a specific form of evolutionary interactions. If species i with abundance x_i replicates with rate f_i, the interaction is of type *replication* and is represented as $x_i \xrightarrow{f_i} 2x_i$, (replicator or Crow-Kimura equation [4]). *Competition* is a mechanism where the replication rate of species i also depends on other species j through a transfer rate p_{ij}, $x_i + x_j \xrightarrow{p_{ij}} x_i$. This type of interactions is used in the game dynamical equation [25], which is a special case of the frequency dependent replicator equation [5,26]. The mechanism *mutation* assigns a mutation or transfer rate q_{ij} between two species according to $x_i \xrightarrow{q_{ij}} x_j$, together with replication and competition we obtain the replicator-mutator Eq [27], of which the quasispecies equation [28]. is a special case. Replication can take place without replicators, species are then produced by *recombination* processes. In the case of three species i, j, and k with a recombination rate α_{ijk}, this mechanism is $x_j + x_k \xrightarrow{\alpha_{ijk}} x_i$. The corresponding dynamical system is called *catalytic network* see e.g., [29]. It is formally possible to express replication, mutation and competition as special cases of the recombination mechanism in the sense that the replicator and replicator-mutator equation can be derived as special cases from the catalytic network equation [30,31]. Thus recombination mechanisms provide a unifying description of the other evolutionary interactions above–an observation we use as a starting point for our model.

In the general form of a recombination process an arbitrary number of species $j_1, j_2, \ldots, j_n$ influences a given species i. We distinguish two types of interactions of this form, (i) *constructive* interactions or productions where species i benefits from species $j_1, j_2, \ldots, j_n$ and (ii) *destructive* interactions or destructions where the j's are causing harm to i. In the master equation framework constructive interactions correspond to positive transfer rates, destructions to negative ones. We denote the set of species $j_1, j_2, \ldots, j_n = \mathbf{j}$. If the set of all species is $\mathcal{N}$, $\mathbf{j}$ is an element of the set of all subsets of $\mathcal{N}$, i.e., $\mathbf{j}$ is an element of the power set of $\mathcal{N}$, $\mathcal{P}(\mathcal{N})$. A recombination always maps an element from $\mathcal{P}(\mathcal{N})$ to an element from $\mathcal{N}$ via a transfer rate $\alpha_{i,\mathbf{j}}$, i.e., by a map $\alpha : \mathcal{P}(\mathcal{N}) \to \mathcal{N}$. From now on italic indices refer to elements of $\mathcal{N}$, e.g., $i \in \mathcal{N}$, while bold-face indices refer to elements of the power set, $\mathbf{j} \in \mathcal{P}(\mathcal{N})$. Transfer rates are represented by their sign. For convenience define binary state variables for sets of nodes, let $\sigma_{\mathbf{j}}(t) = \Pi_{i \in \mathbf{j}} \sigma_i(t)$. The most general form of evolutionary interactions can then be written as

TABLE II. Summary of the traditional evolutionary interaction mechanisms: replication, competition, mutation and recombination. We indicate how the constructive/destructive set $\mathbf{j}$ has to be specified in Eq. (12) in order to recover the various mechanisms in our model.

Mechanism				Power set notation
Replication	x_i	$\xrightarrow{f_i}$	$2x_i$	$\mathbf{j} = \{i\}$
Competition	$x_i + x_j$	$\xrightarrow{p_{ij}}$	x_i	$\mathbf{j} = \{i, j\}$
Mutation	x_j	$\xrightarrow{q_{ij}}$	x_i	$\mathbf{j} = \{j\}$
Recombination	$x_j + x_k$	$\xrightarrow{\alpha_{ijk}}$	x_i	$\mathbf{j} = \{j, k\}$

$$\sigma_{\mathbf{j}} \xrightarrow{\alpha_{i,\mathbf{j}}} \sigma_i. \qquad (12)$$

We summarize in Table II how the evolutionary interaction mechanisms of replication, competition, mutation and recombination are contained in Eq. (12) for special choices of $\mathbf{j}$. If there is a constructive interaction between $\mathbf{j}$ and i, i.e., $\mathbf{j}$ is a *constructive set*, we capture it in the *production rule table* $\alpha^+ : \mathcal{P}(\mathcal{N}) \to \mathcal{N}$ with $\alpha^+_{i,\mathbf{j}} = 1$, otherwise $\alpha^+_{i,\mathbf{j}} = 0$. Similarly, if the interaction in Eq. (12) is destructive, i.e., $\mathbf{j}$ is a *destructive set*, we record this in the *destruction rule table* $\alpha^- : \mathcal{P}(\mathcal{N}) \to \mathcal{N}$ with $\alpha^-_{i,\mathbf{j}} = 1$, otherwise $\alpha^-_{i,\mathbf{j}} = 0$. At some points in this work we will assume that the rule tables $\alpha^{\pm}$ are random tensors. In this case they are given by two parameters $n^{\pm}$ and $r^{\pm}$. n^+ is the cardinality of constructive sets, $|\mathbf{j}| = n^+$ in Eq. (12) and for random α^+ each species i has on average the same number r^+ of constructive sets, $\langle \Sigma_{\mathbf{j}} \alpha^+_{i,\mathbf{j}} \rangle_i = r^+$. Similarly α^- is given by n^- and r^-.

In a biological setting one can identify with our species σ_i an individual gene, RNA mutants or proteins. A set of species $\sigma_{\mathbf{j}}$ can be identified with a combination of chemical elements or proteins building a complex. The interactions could also include a γ-quant acting on a gene and producing a mutation. How to translate this model into an economic setting is explained in great detail in [24]. In this model a mutation is any kind of process by which the existence of a certain species j alone leads to a proliferation of another, different species i. Consequently a recombination is any kind of process through which the simultaneous existence of a set of species leads to the proliferation of another one. We do not make any further assumptions or specifications about the details of the recombination mechanisms involved, of which a number of variants have been studied in the literature [32–35]. In the setting of binary state variables for species abundances used here, these variants of recombination mechanisms are mimicked through their topological structure encoded in $\alpha^{\pm}$. This phenomenologic description level comes with the benefit of leading to a general and analytically tractable model of evolution.

Note that our formalism possesses structural similarities with other work where population genetic equations have been mapped onto spin-Hamiltonians, see e.g., [36] which in turn builds on [37]. In [36] a species is represented as a spin and collected into genomes which are chains of spins of

---

[3] As is typical for traditional evolutionary biology.
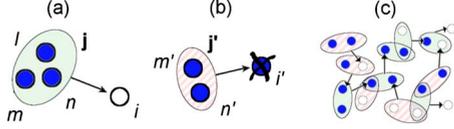
FIG. 1. (Color online) A graphical representation of constructive, destructive and combined interactions. (a) The constructive set **j** (shaded area) contains three species (circles) $\mathbf{j}=\{l,m,n\}$. They produce species i, i.e., $\alpha_{i,\mathbf{j}}^+=1$. The constructive set is active (indicated by the filled circles) we have $\sigma_{\mathbf{j}}(t)=1$ and $\sigma_i(t)=0$, by Eq. (13) $\sigma_i(t+1)=1$. (b) The destructive set $\mathbf{j}'$ (striped area) of cardinality two is active since each of its contained species m',n' is active and interacting with species i' through $\alpha_{i',\mathbf{j}'}^-=1$. Following Eq. (18), $\sigma_{i'}(t)=1$ will be deactivated, $\sigma_{i'}(t+1)=0$. (c) A pictorial description of a network with both constructive and destructive interactions at a point in time.

length N. The evolution of the relative frequencies of these genomes is then studied using Crow-Kimura [4] or quasispecies [28] models for the underlying dynamics. These models are mapped onto Hamilton-Jacobi equations admitting analytic solutions for the steady state distribution of genome frequencies for a certain class of fitness functions. In contrast to this work our dynamical principle Eq. (6) does not require any further assumptions about the dynamics other than the existence of an indicator function as given in Eq. (1). We are interested in the specific interaction structure as given in Table II, our dynamical equations Eq. (3) therefore lack the constant renormalization as present in the Crow-Kimura or quasispecies model (i.e., the comparison of individual growth rates versus population average of fitness). This means the dynamics studied here is rather different.

IV. CONSTRUCTIVE INTERACTIONS

A. Constructive dynamical system

We first consider a system with constructive interactions only. We read Eq. (12) as "from $\sigma_j(t)=1$ follows that $\sigma_i(t+1)=1$." In a chemical setting the chemical compounds contained in **j** react to give compound i, in an economic setting the goods **j** can be assembled to produce good i, see Fig. 1(a). The constructive dynamical system characterized by Eq. (12) is given by

$$\Delta\sigma_i(t)=\text{sgn}\left\{[1-\sigma_i(t)]\sum_{\mathbf{j}\in\mathcal{P}(\mathcal{N})}\alpha_{i,\mathbf{j}}^+\sigma_{\mathbf{j}}(t)\right\},\quad(13)$$

i.e., $f_i[\vec\sigma(t)]$ from Eq. (1) becomes $f_i^+[\vec\sigma(t)]=\Sigma_{\mathbf{j}}\alpha_{i,\mathbf{j}}^+\sigma_{\mathbf{j}}(t)$.

B. Deterministic constructive diversity dynamics

In the limiting case of $T=0$ i.e., $\beta\to\infty$ the system deterministically obeys the dynamics of Eq. (13). The behavior of $D(t\to\infty)$ is well understood; this case is identical to the model studied in [20,21] for random interaction topologies α^+ given by n^+,r^+ here. $D(\infty)$ was computed as a function of n^+,r^+ and $D(0)$. It was shown that this system has a phase transitions formally equivalent to the phase transition of a van der Waals Gas. There exists a critical diversity of initial species $D_c(0)$ above which the system is driven toward an

almost fully populated state; below this threshold the dynamics freezes. All these findings are identical what we find here in the $T=0$ case.

C. Stochastic constructive diversity dynamics

We next turn to nonzero temperature T. The crucial feature distinguishing deterministic and stochastic diversity dynamics is the *dependence on the initial conditions*. In the presence of stochastic perturbations the final diversity is not a function of the initial diversity $D(0)$. We employ a mean-field approach by assuming that the expectation value of a product equals the product of expectation values, $\langle g_1(\sigma)g_2(\sigma)\rangle=\langle g_1(\sigma)\rangle\langle g_2(\sigma)\rangle$. The expectation value of the constructive potential $\langle V_i^+\rangle_{p[\vec\sigma(t)]}$ of species i is

$$\langle V_i^+\rangle_{p[\vec\sigma(t)]}=[1-q_i(t)]\sum_{\mathbf{j}\in\mathcal{P}(\mathcal{N})}\alpha_{i,\mathbf{j}}^+\prod_{j\in\mathbf{j}}q_j(t),\quad(14)$$

quantifying what *could* be produced given the actual configuration of the system.[4]

As mentioned above, μ plays the role of a threshold. For $0<\mu<2$ a species gets activated by one constructive set, for $2<\mu<4$ at least two constructive sets are needed and etc. From now on we fix the threshold $\mu=1$. The contribution to free energy is $K_i^+(t)=\frac{1}{2}[\Delta\sigma_i(t)]^2$. We can estimate the expectation value $\langle K_i(t)\rangle_{p[\vec\sigma(t)]}$ by making use of the dynamical relation Eq. (13), and get for the mean-field assumption

$$\langle K_i^+\rangle_{p[\vec\sigma(t)]}=\frac{1}{2}\left\{[1-q_i(t)]\sum_{\mathbf{j}\in\mathcal{P}(\mathcal{N})}\alpha_{i,\mathbf{j}}^+\prod_{j\in\mathbf{j}}q_j(t)\right\}^2.\quad(15)$$

Using this in Eq. (11) gives us the mean-field solution for arbitrary interaction topologies α^+. To compute it explicitly we assume random interaction topologies. The aim is to derive an expression for $\frac{\partial\langle B_i\rangle}{\partial q_i}$ in the limit $t\to\infty$. Note that $\frac{\partial\langle V_i\rangle}{\partial q_i}=-\Sigma_{\mathbf{j}\in\mathcal{P}(\mathcal{N})}\alpha_{i,\mathbf{j}}^+\Pi_{j\in\mathbf{j}}q_j$. Due to the randomness in α^+ the same average 'field' is exerted on each species. With $q\equiv\langle q_i\rangle_i$ we get $\frac{\partial\langle V_i\rangle}{\partial q_i}=-r^+q^{n^+}$. We apply the same reasoning to the distance-contribution $\langle K_i^+\rangle_{p[\sigma(t)]}$. By first carrying out the derivation and then putting in the assumptions about α^+, we get

$$\frac{\partial\langle B_i\rangle}{\partial q_i}=-r^+q^{n^+}-(1-q)(r^+q^{n^+})^2,\quad(16)$$

and the self-consistent solution for the asymptotic abundance q,

$$q=\frac{1}{2}\left\{\tanh\left\{\frac{\beta}{2}[r^+q^{n^+}+(1-q)(r^+q^{n^+})^2]\right\}+1\right\},\quad(17)$$

from which the diversity follows as $D(t\to\infty)=Nq$. We compare predictions of Eq. (17) with simulations results from a

---

[4] Note that the structure of V_i^+ has a strong similarity to the potential of the paradigmatic Ising model. Our model diverges in the following ways: (i) interactions are defined not between nodes but between constructive sets and nodes and (ii) interactions are not symmetric, the action of **j** on i does not equal the action of i on **j**.
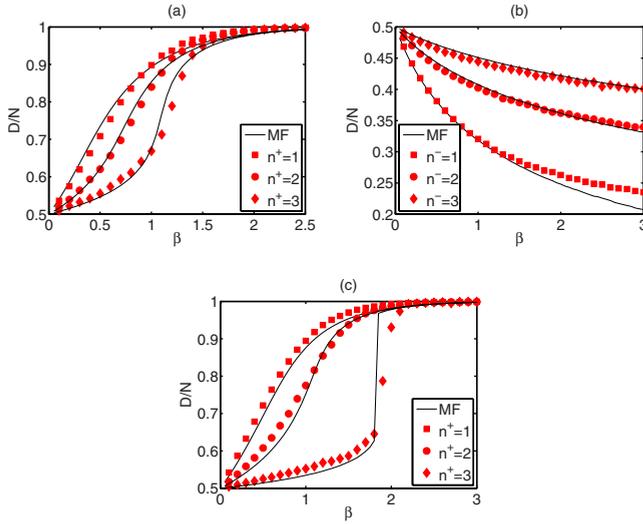
FIG. 2. (Color online) Diversity as a function of inverse temperature β for various dynamical systems obtained from a mean-field approach (MF, lines) and Metropolis simulations (symbols). (a) Constructive dynamics with $r^+ = 2$. (b) Destructive dynamics with $r^- = 2$. (c) Combined dynamics with $r^+ = 3$, $r^- = 1$ and $n^- = 2$.

Metropolis algorithm. The latter was implemented in the following way: We constructed a random α^+ and initialized the system with a random initial condition $\vec{\sigma}(0)$. After initialization the algorithm applies the following procedure to each species once within one timestep (random sequential update):

(i) Pick a species i randomly.

(ii) Calculate $B_i = K_i + V_i$, according to Eqs. (4) and (5) with $\sigma_i'(t) = \sigma_i(t)$.

(iii) Calculate $B_i = K_i + V_i$ with $\sigma_i'(t) = 1 - \sigma_i(t)$.

(iv) Calculate $\Delta B = B_i[\sigma_i'(t) = 1 - \sigma_i(t)] - B_i[\sigma_i'(t) = \sigma_i(t)]$

(v) If $\Delta B < 0$ set $\sigma_i(t+1) = 1 - \sigma_i(t)$.

(vi) If $\Delta B > 0$ set $\sigma_i(t+1) = 1 - \sigma_i(t)$ with probability $e^{-\beta \Delta B}$.

We executed the algorithm for one particular realization of α^+ for 10^3 timesteps and averaged over this time-span after discarding transient behavior (typically about 50 iterations). We performed simulations for system sizes of $N = 10^2 - 10^4$ without noticing size effects on the results. However, the time-to-converge depends on N. We show the degree of agreement of simulations and Eq. (17) in Fig. 2(a).

V. DESTRUCTIVE DYNAMICS

A. Destructive dynamical systems

Assume now that only destructive interactions take place, e.g., two chemicals catalyzing the consumption of another chemical species, or biological species gaining (in symbiosis) an evolutionary advantage over another species. Equation (12) is now read as "from $\sigma_{\mathbf{j}}(t) = 1$ follows $\sigma_i(t+1) = 0$," see Fig. 1(b). To formulate this as a dynamical system as in Eq. (3) set $f_i[\vec{\sigma}(t)] \rightarrow f_i^-[\vec{\sigma}(t)] = -\sum_{\mathbf{j}} \alpha_{i,\mathbf{j}}^- \sigma_{\mathbf{j}}(t)$ to get

$$\Delta \sigma_i(t) = \text{sgn}\left(-\sigma_i(t) \sum_{\mathbf{j} \in \mathcal{P}(\mathcal{N})} \alpha_{i,\mathbf{j}}^- \sigma_{\mathbf{j}}(t)\right). \tag{18}$$

We discuss the deterministic ($T = 0$) and stochastic ($T > 0$) scenario.

B. Deterministic destructive diversity dynamics

In the deterministic case the asymptotic diversity $D(t \rightarrow \infty)$ is a function of the initial diversity. Let us discuss the case of a completely random destructive rule table α^-. By denoting $q(t) = D(t)/N$ we can derive an update equation for $q(t)$ following the same reasoning as in [20], $q(t+1) = q(t) - \Delta^- q(t)$ with $\Delta^- q(t) = r^- q(t)[q^{n^-}(t) - q^{n^-}(t-1)]$. In the limit of sparse rule densities r^- this leads to $q(t \rightarrow \infty) = q(0) - n^- r^- q^{n^- + 1}$. In contrast to constructive dynamics, destructive dynamics do not exhibit a phase transition. With more species being destroyed the number of deactivated destructive sets increases even faster, thus the process comes to a halt without reaching a strongly unpopulated state. Note that other kinds of destructive dynamics may exhibit phase transitions. For example, if one requires for each species to be abundant at least one productive set to be abundant, that is species become extinct once they are not actively produced, the removal of a small number of species may trigger a cascade of extinction events which erases the entire population [21].

Stochastic destructive diversity dynamics

For a stochastic variant of the destructive dynamical system of Eq. (18) we repeat the analysis of the constructive case. We start with the corresponding destructive potential and distance terms,

$$\langle V_i^- \rangle_{p[\vec{\sigma}(t)]} = q_i(t) \sum_{\mathbf{j} \in \mathcal{P}(\mathcal{N})} \alpha_{i,\mathbf{j}}^- \prod_{j \in \mathbf{j}} q_j(t),$$

$$\langle K_i^- \rangle_{p(\vec{\sigma}(t))} = \frac{1}{2}\left[q_i(t) \sum_{\mathbf{j} \in \mathcal{P}(\mathcal{N})} \alpha_{i,\mathbf{j}}^- \prod_{j \in \mathbf{j}} q_j(t)\right]^2. \tag{19}$$

We proceed with the derivation of the destructive balance function B_i and get

$$\frac{\partial \langle B_i \rangle}{\partial q_i} = r^- q^{n^-} + q(r^- q^{n^-})^2, \tag{20}$$

and the self-consistent solution for the asymptotic abundance q,

$$q = \frac{1}{2}\left\{\tanh\left\{-\frac{\beta}{2}[r^- q^{n^-} + q(r^- q^{n^-})^2]\right\} + 1\right\}. \tag{21}$$

We compare this prediction to results of a Metropolis simulation in Fig. 2(b). As is seen in the $n^- = 1$ case, the deviation between Eq. (21) and simulations increases with β. For higher n^- and r^- the same extent of deviation occurs for a higher value of β. The mean-field approximation starts to significantly differ from simulations once entropic effects become negligible and the system's evolution approaches the deterministic scenario, that is $e^{-\beta \Delta B_i} \lesssim 1/N$ $\forall i$ (on average less than one random state flip per iteration). To approximate $V_i^-(t) = \sigma_i(t) \sum_{\mathbf{j}} \alpha_{i,\mathbf{j}}^- \prod_{j \in \mathbf{j}} \sigma_j(t)$ at any time t we have to consider the species which have *not* been deactivated at $t-1$—the system possesses memory. This is not captured in the mean-field approximation $\langle V_i^- \rangle = r^- q^{n^-}$ where we assume the populated species to be randomly distributed over N possible spe-

cies at each time t. In the destructive case the mean-field approach thus works best whenever the random fluctuations are large enough to "smear out" this memory effect, otherwise the system is better approximated by the deterministic description.

VI. COMBINED DYNAMICS

A. Combined dynamical systems

We now study the interplay of both constructive and destructive dynamics [22,24]; the situation is sketched in Fig. 1(c). Destructive interactions represent an implicit selection mechanism [24]. Each species may be targeted (influenced) by constructive *and* destructive interactions. Assume that each interaction has equal influence. If the constructive forces outweigh the destructive ones the species prefers to be active and vice versa. For some systems other choices of weighting could be more appropriate (e.g., assuming that *one* destructive interaction outweighs any number of constructive ones—"it is easier to destroy than to build"). It is straightforward to incorporate alternative weighting schemes in the present framework.

To combine constructive and destructive interactions we add their indicator functions,

$$f_i[\vec{\sigma}(t)] = f_i^+[\vec{\sigma}(t)] + f_i^-[\vec{\sigma}(t)] = \sum_{\mathbf{j}} \alpha_{i,\mathbf{j}}^+ \sigma_{\mathbf{j}}(t) - \sum_{\mathbf{j}} \alpha_{i,\mathbf{j}}^- \sigma_{\mathbf{j}}(t),$$
(22)

and get for the dynamical equation

$$\Delta \sigma_i(t) = \text{sgn}[[1 - \sigma_i(t)]R\{f_i[\vec{\sigma}(t)]\} - \sigma_i(t)R\{-f_i[\vec{\sigma}(t)]\}].$$
(23)

The purely destructive or constructive dynamical systems are recovered by setting $\alpha^\pm = 0$.

B. Deterministic combined diversity dynamics

To obtain an estimate for the asymptotic diversity, we again use an update equation and combine the finding for the constructive and destructive cases. If we denote the average in-(de)crements in the constructive (destructive) scenario by $\Delta q^+(t)$ [$\Delta q^-(t)$], we study the update equation $q(t+1) = q(t) + \Delta q^+(t) - \Delta q^-(t)$. This equation is solved by using the same Ansatz as in [20], yielding $q(t \to \infty) = q(0) - n^- r^- q^{n^- + 1} + n^+ r^+ (1-q) q.^{n^+}$

C. Stochastic combined diversity dynamics

Let us calculate $\langle B_i \rangle$ for the stochastic scenario. The expectation value of the distance contribution, $\langle K_i \rangle_{p[\vec{\sigma}(t)]}$, is more involved now. Constructive (destructive) dynamics take place under the condition that $f_i[\vec{\sigma}(t)] \geq 0 (\leq 0)$. Start with an expression for the probability that $f_i[\vec{\sigma}(t)]$ is positive (negative), $p^\pm$. Consider random interaction topologies specified by $r^\pm$ and $n^\pm$. Define $p(k, r^+)$ as the probability that there are exactly k active constructive interactions, that is $p(k, r^+) \equiv \binom{r^+}{k} q^{n^+ k}(1 - q^{n^+})^{r^+ - k}$. Analogously, $q(l, r^-)$ is the

probability that exactly l out of r^- destructive interactions are active. Then

$$p^+ = \sum_{k=1}^{r^+} p(k, r^+) \sum_{l=0}^{\min(k-1, r^-)} q(l, r^-),$$

$$p^- = \sum_{l=1}^{r^-} q(l, r^-) \sum_{k=0}^{\min(l-1, r^+)} p(k, r^+).$$
(24)

The average distance follows as

$$\langle K_i \rangle_{p(\sigma)} = \frac{1}{2}[(1 - q_i)p^+ + q_i p^-]^2,$$
(25)

and, abbreviating $f_i[\vec{\sigma}(t)] \equiv f_i$, the potential is

$$\langle V_i \rangle_{p(\sigma)} = |(1 - q_i)R(f_i) - q_i R(-f_i)|.$$
(26)

Taking the derivative with respect to q_i the mean-field result is

$$\frac{\partial \langle B_i \rangle}{\partial q_i} = -r^+ q^{n^+} + r^- q^{n^-} - [(1-q)p^+ + qp^-](p^+ - p^-),$$
(27)

with the self-consistent solution for the asymptotic abundance q

$$q = \frac{1}{2}\left\{ \tanh\left[\frac{\beta}{2}\{r^+ q^{n^+} - r^- q^{n^-} + [(1-q)p^+ + qp^-](p^+ - p^-)\} \right] + 1 \right\}.$$
(28)

Again we compare the mean-field prediction to results of a Metropolis simulation of the full model in Fig. 2(c).

VII. DISCUSSION ON EMPIRICAL RELEVANCE

A. Economical setting

We interpret the model in different evolutionary contexts and compare its behavior to measured data. In an economic setting one can identify the number of active interactions as a measure for the productive output of an economy—for example the GDP [24]. An interaction is defined to be active iff $\sigma_i(t) = \sigma_{\mathbf{j}}(t) = \alpha_{i,\mathbf{j}}^\pm = 1$. We show in Fig. 3(a) a comparison between the actual distribution of percent increments of the GDP of the U.K. and the number of active productions from the combined stochastic model for two different parameter settings. In one setting $\beta = 15$, $r^\pm = 5$, $n^\pm = 2$ is used, the other has a denser interaction topology, $\beta = 15$, $r^+ = 8$, $r^- = 12$, $n^\pm = 2$. Both model and real-world GDP timeseries produce fat-tailed distributions, with power exponents in the range between -2 and -4. These features are also found in GDP timeseries of other countries and for a wide range of model parameters, see e.g., [24].

B. Chemical setting

Another possible interpretation of the combined stochastic system is a chemical reaction network. In this case chemical
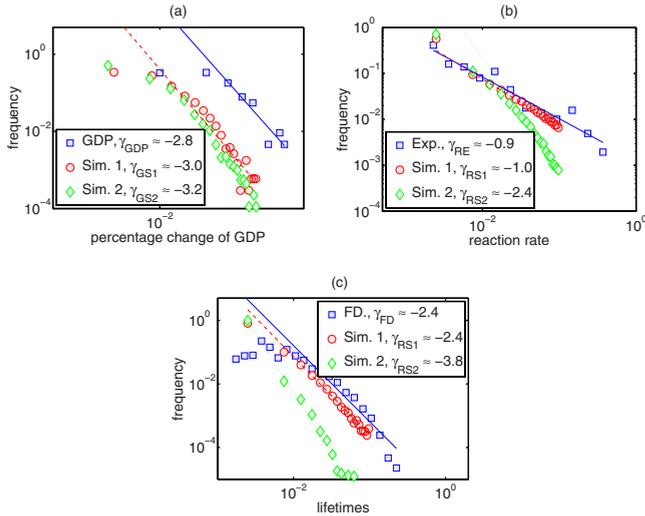
FIG. 3. (Color online) We compare the distribution of systemic observables (squares and fit as solid line) of evolutionary systems with those of the combined stochastic model for two different parameter settings: *Simulation 1* with $\beta=15$, $r^{\pm}=5$, $n^{\pm}=2$ (circles and fit as dashed line) and *Simulation 2* with $\beta=15$, $r^{+}=8$, $r^{-}=12$, $n^{\pm}=2$ (diamonds and fit as dotted line). Each distribution has been normalized (sum over all data points equals one). (a) The percent change of GDP of the U.K. since 1950 is compared to the model. (b) The reaction rate distribution in the model and in the metabolic network of *E. coli* is shown. (c) Species lifetime distributions as found in fossil data are well reproduced with the model.

species $\mathbf{j}=\{j_1, j_2, \ldots\}$ are producing or degrading chemical i. There are $N(r^{+}+r^{-})$ reactions. A reaction rate is defined as the frequency with which a certain reaction is active and a reaction is active if $\alpha_{i,\mathbf{j}}^{\pm}=1$, $\sigma_i(t)=1$ and $\sigma_{\mathbf{j}}(t)=1$. This is compared to reaction rates in the metabolic network of *E. coli* [38] in Fig. 3(b). Distributions of reaction rates in both cases, model and living organism, are fat-tailed. Least-squares fits to model power-laws yield exponents in the range of −1 to −3, depending on parameters. This compares well to the value of −1 found for *E. coli*.

C. Biological evolution setting

Translated into a macroecological setting, one can compare the distribution of lifetimes of species in the combined stochastic model (number of iterations a given species is unintermitted abundant) with the distribution of species lifetimes in fossil data [39] in Fig. 3(c). Again one finds power laws in the model with exponents between −2 and −4, which matches well with the paleontologic data, which suggest slopes between −2 and −3. Note that there is a strong dependence on the values used for the fit. We work with an intermediate choice in Fig. 3(c).

VIII. DISCUSSION

We propose a general framework to systematically study a large class of dynamical evolutionary systems defined on an arbitrary large number of species. The trajectory of existence of each species is governed by a function incorporating information of the surroundings—the existence of other species. We show how to express the resulting system dynamics via a variational principle. We discuss deterministic and stochastic variants. For the latter we derive a closed expression for the asymptotic diversity of evolutionary systems within a mean-field approximation. We discuss the quality of this approximation with respect to Metropolis simulations of the full model. Although the model explicitly introduces strong correlations between species' abundances, the mean-field approximation for asymptotic diversities match the simulation data surprisingly well. The model can be seen as a generalization of several previous models, which are contained as special cases. The deterministic constructive case is identical to the random catalytic networks studied in [20]. In the model of Solé and Manrubia [14] only linear interactions are allowed [i.e., $|j|=1$ in Eq. (12)] and new species are created not through endogenous recombinations, but by an explicit mutation mechanism. As discussed in [24], $f_i[\vec{\sigma}(t)]$ in our combined stochastic model plays the identical role as the randomly assigned fitness values in the Bak-Sneppen model [13]. To recover the *NK*-model [8] as a special case associate each species with a bit-string. A random fitness value is then assigned to each species' bit-string, in some variants of the model also in dependence of a given number of bits of other species' strings. However, fitness in our framework is a topological property of the entire system plus the set of abundant species, whereas in *NK*-models fitness is basically a mapping of random numbers to bit-strings.

We find that the model of constructive and destructive interactions reproduces stylized facts of man-made (economies) and natural evolutionary systems (metabolic networks, macroecology) across different orders of magnitude. We belief this adds empirical substance to our claim that we have identified a crucial and ubiquitous building block of evolutionary systems with recombinatory, nonlinear interactions within a simple binary framework. The model systematically expands on the idea that the concept of fitness is an *a posteriori* concept. Fitness in the traditional sense can of course be reconstructed for every timestep in our model. It is nothing but the coevolving network of rates of the actually active (productive) processes at a given time, see [24] for more details. It becomes clear that fitness cannot be used as concept with much predictive value, even if "Darwin's Demon" knowing all mutual influences at a given time would exist. The proposed model is free of Darwin's Demon.

[1] R. T. Paine, Am. Nat. **103**, 91 (1969).
[2] Semiconductor Industry Association Factsheet, http://www.sia-online.org/cs/industry_res ources/industry_fact_sheet
[3] V. Volterra, ICES J. Mar. Sci. **3**, 3 (1928).
[4] J. F. Crow and M. Kimura, *An Introduction to Population Genetics* (Burgess Publishing Co., Minneapolis, MN, 1970).
[5] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge University Press, Cambridge, England, 1998).
[6] J. Maynard Smith, Nature (London) **225**, 563 (1970).
[7] S. Wright, Proceedings of the Sixth International Congress on Genetics, Vol. 1, pp. 356–366 (1932).
[8] S. A. Kauffman, J. Theor. Biol. **22**, 437 (1969).
[9] A. F. M. Marée, W. Keulen, C. A. B. Boucher, and R. J. De Boer, J. Virol. **74**, 11067 (2000).
[10] S. Bonhoeffer, A. D. Barbour, and R. J. DeBoer, Proc. Biol. Sci. **269**, 1887 (2002).
[11] B. Arthur, Complexity **11**, 23 (2006).
[12] S. Jain and S. Krishna, Phys. Rev. Lett. **81**, 5684 (1998); Proc. Natl. Acad. Sci. U.S.A. **99**, 2055 (2002).
[13] P. Bak and K. Sneppen, Phys. Rev. Lett. **71**, 4083 (1993).
[14] R. V. Solé and S. C. Manrubia, Phys. Rev. E **54**, R42 (1996).
[15] M. E. J. Newman R. G. Palmer, *Modeling Extinction* (Oxford University Press, New York, NY, 2003).
[16] P. Dittrich and P. Speroni di Fenizio, Bull. Math. Biol. **69**, 1199 (2007).
[17] S. A. Kauffman, *Origins of Order: Self-Organization and Selection in Evolution* (Oxford University Press, New York, NY, 1993).
[18] L. Kadanoff, S. Coppersmith, and M. Aldana, in *Perspectives and Problems in Nonlinear Science. A Celebratory Volume in Honor of Lawrence Sirovich*, Springer Applied Mathematical Sciences Series, edited by E. Kaplan, J. E. Marsden, and K. R. Sreenivasan (Springer-Verlag, Berlin, 2003).
[19] S. L. Pimm, G. J. Russell, J. L. Gittleman, and T. M. Brooks, Science **269**, 347 (1995).
[20] R. Hanel, S. A. Kauffman, and S. Thurner, Phys. Rev. E **72**, 036117 (2005).
[21] R. Hanel, S. A. Kauffman, and S. Thurner, Phys. Rev. E **76**, 036110 (2007).
[22] S. Thurner, R. Hanel, P. Klimek, Physica A **389**, 747 (2010).
[23] P. Klimek, S. Thurner, and R. Hanel, J. Theor. Biol. **256**, 142 (2009).
[24] S. Thurner, P. Klimek, and R. Hanel, New J. Phys. (to be published).
[25] P. Schuster and K. Sigmund, J. Theor. Biol. **100**, 533 (1983).
[26] P. Taylor and L. Jonker, Math. Biosci. **40**, 145 (1978).
[27] K. P. Hadeler, SIAM J. Appl. Math. **41**, 1 (1981).
[28] M. Eigen, J. McCaskill, and P. Schuster, Adv. Chem. Phys. **75**, 149 (1989).
[29] P. F. Stadler, W. Fontana, and J. H. Miller, Physica D **63**, 378 (1993).
[30] W. Fontana, G. Wagner, and L. Buss, Artif. Life **1**, 211 (1994).
[31] K. M. Page and M. A. Nowak, J. Theor. Biol. **219**, 93 (2002).
[32] R. Burger, *The Mathematical Theory of Selection, Recombination, and Mutation* (Wiley, New York, 2000).
[33] E. Cohen, D. A. Kessler, and H. Levine, Phys. Rev. Lett. **94**, 098102 (2005).
[34] J.-M. Park and M. W. Deem, Phys. Rev. Lett. **98**, 058101 (2007).
[35] M. Baake and E. Baake, Can. J. Math. **55**, 3 (2003).
[36] D. B. Saakian, O. Rozanova, and A. Akmetzhanov, Phys. Rev. E **78**, 041908 (2008).
[37] E. Baake, M. Baake, and H. Wagner, Phys. Rev. Lett. **78**, 559 (1997).
[38] M. Emmerling *et al.*, J. Bacteriol. **184**, 152 (2002).
[39] J. J. Sepkoski, Jr., Contrib. Biol. Geol. **83**, 1 (1992).