An energy-based natural selection model

Noam Abadi\textsuperscript{a,b}, Guillermo Abramson\textsuperscript{a,b,c}

\textsuperscript{a}Centro Atómico Bariloche (CNEA), R8402AGP Bariloche, Argentina
\textsuperscript{b}Instituto Balseiro, Universidad Nacional de Cuyo, Argentina
\textsuperscript{c}Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

Abstract

Genetic information and environmental factors determine the path of an individual’s life and, therefore, the evolution of its entire species. We have succeeded in proposing and studying a model that captures this idea. In our model, a renewable resource extended throughout the environment provides the energy necessary to sustain life, including movement and reproduction. Since the resource doesn’t regrow immediately, it generates competition between individuals and therefore provides a natural selection pressure from which evolution of the genetic traits is observed. As a result of this, several phenomena characteristic of living systems emerge from this model without having to introduce them explicitly. These include speciation and punctuated equilibrium, competitive exclusion, and altruistic behavior from selfish rules.

Keywords: biological evolution, competition, natural selection, genetic algorithms

1. Introduction

Darwin’s theory \cite{1} was a huge step towards answering questions regarding the evolution of life and providing a unifying framework for its diversity as observed on Earth. Its central idea, and the source of its success, was to provide natural selection as the mechanism for evolution, a plausible way by which some naturally occurring variants of organisms increase their relative survival chances. However, the theory is so broad in its reach that many aspects of it are still under study. In particular, enormous advances in experimental molecular biology and the manipulation of genes, sustained over recent decades, have left many theoretical aspects lagging behind.

For example, it has become increasingly accepted that the evolution of the biosphere needs an integrated description of both the biological agents as well as their resources \cite{2}. Indeed, the main processes responsible for evolution, namely speciation and extinction, are the result of interactions, and interactions are often mediated by shared resources. Nevertheless, most theoretical approaches based on mathematical modeling, in particular those based on statistical physics, ignore the existence and dynamical role of resources (see, for example, the review by Drossel in \cite{3}).

Email addresses: noam.abadi@ib.edu.ar (Noam Abadi), abramson@cab.cnea.gov.ar (Guillermo Abramson)

Preprint submitted to Physica A May 19, 2021
The main area where modern understanding transcends Darwin’s is genetics. We understand many details of inheritance, from the molecular to the physiological levels. Yet, the relation between genetically inherited traits and the probability of producing offspring in a certain environment is not completely clear. Similarly, it is not obvious what general features of extinction, survival, and eventually speciation should arise from the complex interaction of the many mechanisms involved. Many of these problems have been studied in recent decades from a theoretical point of view, with a variety of tools rooted in the theory of complex systems [4–7].

Indeed, living systems are chemical, and their dynamics is driven by free energy. At the core of life, there seems to be just a chemical metabolism, which nowadays is a complex network of reaction pathways, involving both equilibrium and out-of-equilibrium processes that drive the flow of free energy through the system. Of necessity, thermodynamic principles, which apply to all physical and chemical systems, shape the dynamics that underly gene expression and regulation, metabolic pathways, potential energy storage, growth and differentiation, etc. [8, 9]. Moreover, there is no indication that individuality or replication played a role in early chemical (pre-biotic) organization [10], and it has been suggested that extremely simple cells, to the point of not qualifying as “organisms,” are the common ancestors of all modern life [11–13]. Even if they perhaps do not define life, individuality, replication, and selection arose at some moment in the evolution of living systems as persistent and universal distinctions.

In this work, we propose a model in which individual agents have traits, determined by genes, that determine the balance between energy input and output. We study how a population of such creatures evolves, in a scenario of competition for a spatially extended common resource that provides the necessary energy. The capacity to move is one such trait, and it allows creatures to change their position when they exhaust their local resource, albeit at a cost in energy. Likewise, body mass is energetically expensive to keep alive and also to carry around, but it guarantees a reserve in case of resource scarcity. Since both traits have benefits and drawbacks, we expect them to have optimal values, which the system should achieve in a stationary state, in a context of competition mediated by the common resource.

The model is similar to latent energy environment (LEE) models of Menczer and Belew [14, 15], but with two key differences. First, the resource is entirely equivalent to a creature’s energy instead of components that need to be combined with others to produce it. This allows the second distinction, which is that individuals have no neural network to act as a brain and allow them to learn throughout their lives. The advantage of this is that the model can address questions regarding primordial life forms. The model is also complementary to energy saving mechanism models, such as the variation range hypothesis of Trenchard and Perc [16]. In it, they approach the relation between energy and genetic information from the viewpoint of energy availability, while we attempt to study energy availability in terms of the genetic distribution.

As opposed to artificial selection, which evolves towards the optimization of a certain function (as implemented in many “artificial life” models, see [17]), natural selection occurs under a certain constraint, which we propose is the limitation of energy. Therefore, the model does not prescribe any landscape of fitness, nor any function to be optimized. Traits are inherited from one generation to the next with small, random mutations, that may change their performance and the corresponding energy balance. This makes the population undergo a natural selection process as the agents interact through the common
resource. In such a context, selection is not modeled, but an emergent mechanism that provides a certain organization to the ecosystem.

We show that phenomena commonly associated with evolution by natural selection appear to be intrinsically tied to it in our model. The first one is speciation, in the form of two possible sets of traits in a stationary state. A completely satisfactory definition of species and speciation is still an open problem in biology, as evidenced by the multiplicity of species concepts (see for example the discussions in [18, 19]). In this work it is not our intention to contribute to that debate. For the sake of the discussion, we adopt here a pragmatic approach and consider a species as the set of individuals that can live in the largest connected subset of the genetic space in which they themselves are found. We emphasize on can because in practice, we will only be able to sample this connected subset, and thus infer if a creature of certain genetic characteristics is expected to survive or not. Defining a species based on a sampled subset did not turn into a complicated classifying problem because we found that the stable species that emerge are disjoint (for fixed environmental conditions, there is a large distance between subsets of living creatures in the genetic space where no individuals survive). We have found this phenomenon in the model and, moreover, it is achieved by mechanisms akin to those of punctuated equilibrium [20]. The second one is that the main behavior throughout the population is more altruistic than what is expected from individual rules, as suggested by the selfish gene theory [21], both in the amount of resource consumed and life expectancy. We also provide a mean-field description of the resource dynamics, which shows important consequences of its spatial extension and limited growth.

2. The hydra model

We will proceed to formulate our model based on the following key ingredients of an evolutionary system as identified by Kaneko [22]: 1) Genotype, phenotype, and environment; 2) Fitness depends on the phenotype and the environment; 3) Only genes are transferred to the offspring; 4) Mutation of genes occurs randomly without any specific direction. Moreover, in the formulation that we describe below, fitness is not a prescribed function but an emergent feature of the dynamics, as will be discussed further on.

Consider a grid of size $L \times L$, and a resource $f$ extended over it. Let us suppose that, in the absence of consumers, its dynamics obeys locally a logistic equation [23]:

$$\frac{df(x)}{dt} = r f(x) \left( 1 - \frac{f(x)}{K} \right).$$

A population of individuals lives on the grid and moves around it, feeding off the resource to obtain the energy $E$ needed to stay alive and storing it in their bodies. We will assume a mass-energy equivalence which makes the amount of energy gained by a creature on feeding equal to the amount of resource lost in the cell where the intake happens.

---

1We call it the “hydra model” because of superficial similarity with the freshwater genus of the same name: tumbling motion, asexual reproduction by budding and potential immortality. Details of the model, either particular parameter values or results, do not pretend to correspond with actual specific features of the life history of Hydra spp.
The resource is renewable, and its complete dynamics depends on the movement of the creatures (local consumers) and their energy dynamics (which dictate their survival, reproduction and death), as described next. For each creature energy evolves according to intake and consumption

$$\frac{dE(t)}{dt} = \varphi(t) - C(t),$$  \hspace{1cm} (2)

where the feeding function \( \varphi \) is the amount of resource that the creature gets from the cell on which it is located, and the consumption \( C \) is the amount of energy it uses to fulfill its bodily needs. Equation (2) states that energy cannot be created inside a creature’s body, and therefore has little meaning without modeling \( \varphi \) and \( C \). To this end, we define a genotype \( G = \{M, P, s\} \), which consists of the adult body size, \( M \), the adult probability of moving to a nearest neighbor cell, \( P \), and the growth rate from birth to adulthood, \( s \). Upon reproduction, the genotype is transmitted to the next generation with small mutations, allowing the population to evolve. The genotype is expressed as a time (age) dependent phenotype \( F(t) = \{m(t), p(t)\} \), which is the creature’s size \( m(t) \) and its probability of moving \( p(t) \). These change as a creature ages to simulate physical growth, according to:

$$\frac{dm(t)}{dt} = s m(t) \left( 1 - \frac{m(t)}{M} \right), \quad m(0) = 0.05M, \quad (3a)$$

$$p(t) = \frac{m(t)}{M} P. \quad (3b)$$

Equations (3) are only meant to capture general behavior of living beings, such as bounded physical growth (the logistic form is one instance of Richard’s model, [24]) and a speed increasing with physical size as growth carries out. For example, here we have modeled the movement of the creature as proportional to its mass, corresponding to steps defined by a characteristic length (resembling a tumbling locomotion), but an exponent in the relation between \( p \) and \( m \) could be used to represent other geometries or movement characteristics.

The phenotype defines \( \varphi \) and \( C \) in the following way. The feeding function \( \varphi \) is such that a creature will eat enough from the resource \( f \) available at \( x \) to fill its energy storage (of maximum size \( m(t) \), the current mass) or, if the resource is not enough for this, whatever is available. A small amount of resource (0.01\(K\)) was made unavailable for consumption in each cell to prevent reaching the equilibrium \( f = 0 \), which is analytically unstable in Eq. (1), but in simulations would prevent regrowth and lead to resource depletion and therefore population extinction. Although this might seem like an artificial imposition, it can be taken to represent the fact that animals do not deplete their local resources, since plants survive herbivory or foraging, renewing themselves later on. We can therefore write the feeding function as:

$$\varphi(t) = \min \{ m(t) - E(t), f(x, t) - 0.01K \}. \quad (4)$$

The consumption \( C \) involves three terms, which represent different bodily needs. Firstly, we have metabolism, with an allometric dependence on mass that is found in living creatures across many orders of magnitude of body size [25], namely \( C_{\text{met}} \propto m^{3/4} \). The second one is kinetic energy, \( C_{\text{kin}} \propto mv^2 \), necessary for movement at velocity \( v(p) \). We model the velocity as a stochastic process taking the value 1 when the creature moves...
from one cell to another (with probability \( p \)) and 0 otherwise. Finally, we have the cost of growing, i.e. the power a creature consumes in increasing its mass from birth to adulthood, which comes from Eqs. (3). In mathematical form:

\[
C(t) = \mu m^{3/4} + \kappa m v^2(p) + \sigma m \left( 1 - \frac{m}{M} \right).
\]

Coefficients \( \mu \) and \( \kappa \) in the metabolic and kinetic terms are proportionality coefficients that may be interpreted as efficiencies; they are constant throughout the populations evolution and do not play a relevant role in the model. We have chosen as \( \mu = 0.1 \) and \( \kappa = 0.5 \).

The final mechanisms that need to be defined for the population to undergo natural selection are death and reproduction. Death happens when a creature empties its energy storage \( (E = 0) \) or spontaneously with an age-increasing probability to include the possibility of death by old age. The value of this probability was set to give an overall expected lifetime of \( \lambda = 1000 \) time steps; we will characterize it thoroughly in Section 3.3.

Reproduction proceeds asexually, simulated phenomenologically as follows. First, we require an adult individual \( (m(t) > 0.95M) \) that keeps its energy above a threshold \( T_R = 0.8M \) for \( \tau = 100 \) time steps. After this initial stage, the body size begins increasing, following a logistic dynamics similar to the one it followed for its own growth, but with a limit value of \( 2M \). This dynamics halts at \( m(t) = 1.05M \), when the offspring separates from the parent, producing two distinct individuals. The brood starts its autonomous existence with mass \( 0.05M \) and the parent is left with the remaining \( M \). The offspring inherits a genome \( G' \) which is a mutation of the parent’s genome \( G \), drawn from a random, uniform distribution centered at \( G \) with a maximum mutation of 5% in each gene. We would like to remark that most of the reproduction parameters were chosen arbitrarily and may not represent realistic values. The values were chosen so that the relevant phenomena could be analyzed in a reasonably clear way. The parameters are also left untouched by mutations since it is difficult to establish how their effects might be beneficial in one way but detrimental in others. In absence of optimal values, the system may be susceptible to never reaching a stationary state. We would also like to emphasize that the precise mathematical form of the presented equations are not of paramount importance. The relevance of this work is to show that, by imposing a restriction on a parameter (we have chosen energy for the sake of simplicity) instead of actively optimizing it (as artificial selection does), natural selection and many of its associated phenomena emerge.

3. Results

The hydra model was studied mainly through numerical simulations. No additional assumptions were made upon implementation aside from those described in Section 2. Results were obtained by measuring demographic characteristics like gene distributions and total population, as well as resource availability at periodic intervals during the simulation. The experiments were repeated several times (typically 20) to identify persistent features in the stochastic realizations. The resource recovery speed \( r \) was varied over sets of realizations to study the evolution under different conditions. In this section we present these results and our interpretation of them.
3.1. Population and genetics

Figure 1 (left panel) shows the temporal evolution of the population for different values of the resource recovery speed, \( r \). The simulations begin with 1000 individuals distributed randomly over both the physical and genetic spaces. The population dynamics shows three distinct regimes (note the logarithmic scale in the time axis). The first one is an eradication of any genes too unfit for survival, either because of competition with others or because of their own depletion of local resource and energy reservoir. The second one is a population explosion where the surviving genes reproduce while the capacity of the system allows them to do so. The last one is a stationary population value with no genetic change but constant individual renewal. For small values of the resource recovery speed (e.g. \( r = 0.0025 \)), the stable population is \( N = 0 \) for all realizations because the resource doesn’t renew itself fast enough to sustain the population. As the recovery speed increases (e.g. \( r = 0.0075 \)), some realizations provide settings where a non-zero stationary population emerges, and some have null populations. As the recovery speed continues to grow, null populations disappear and realizations converge to one of two stable, non-zero populations (e.g. \( r = 0.0175 \), shown on the right panel of Figure 1). Both initial transient regimes occur very quickly in terms of elapsed generations. The following stationary state is stable (albeit fluctuating) which makes the whole process reminiscent of the bursts in which evolution is thought to happen in nature [5, 20, 26–29].

From now on, let us focus on the stationary state unless stated otherwise. We have observed that the adult body mass obeys a weakly linear dependence on the recovery parameter \( r \). In a phenomenological fit \( M = M_0 + ar \), we obtained \( M_0 = 0.0717 \pm 0.0004 \) and \( a = 0.02 \pm 0.03 \). For values of \( r \) in the simulated range (\( r \approx 0.01 \), the error in the constant term is of the same order as the value of the linear one, making the latter negligible. In this sense, we can consider that the adult mass reaches a stationary
Figure 2: Left: Stationary population vs resource recovery rate. Both branches are linear in $r$. Since the bottom branch is constant in probability of taking a step, its slope can be calculated with the mean-field equation (18) obtained in Section 4. Right: Stationary adult speed vs resource recovery rate. The top branch is constant in $r$, $P = 1$ and corresponds to the bottom branch of the left figure. We found that the bottom branch obeys $P \propto r^{-1/2}$ and corresponds to the top branch of the left figure.

value independent of the recovery rate of the resource, within the precision achieved in the simulations. Meanwhile, the growth speed $s$ increased monotonously in all the simulations. The specific form in which this parameter changed was the same for all simulations regardless of the value of $r$, meaning that a faster growth speed always increases chances of survival. The effect of this gene seems to be purely in the amount of time over which a fixed amount of energy, determined by the mass gene, is expended on growth. Since the behavior of both of these parameters is untouched by the parameters we varied, we have mostly ignored them in this analysis. On the other hand, the probability of moving, $P$, presented non-trivial behavior that we will now present.

As mentioned above, stationary populations of different realizations with the same resource recovery speed converge to one of two values (fluctuating around them), as shown in the right panel of Fig. 1, corresponding to $r = 0.0175$. In Fig. 2 (left) we show how these two possible stationary populations vary with the recovery speed by measuring the mean population $N$ (averaged over time in the stationary regime for each realization). For equal recovery speeds, different populations can be easily explained by different energy requirements. In the same figure, realizations are colored by the mean adult probability of moving, $P$, (averaged over population and stationary). This shows that, indeed, the smaller populations always have $P = 1$ (black points) while the larger ones have $P < 1$ (colored points). Since body size is constant, smaller values of $P$ indicate lower energy consumption. Both branches show a linear relation to resource recovery, and we will indulge on this in Section 4. Figure 2 (right) shows the dependence of the average probability of moving, $P$, with the resource recovery, colored by the average population. The top branch (corresponding to the bottom one of Fig. 2 (left)) presents an almost perfectly constant behavior, $P = 1$, while the bottom one corresponds to an algebraic relation, $P \propto r^{-1/2}$.

Figure 2 shows that repeated realizations with the same initial and environmental conditions (resource properties) converge to one of two distinct genetic groups. The coexistence of these two sets is unstable, in accordance with the principle of competitive exclusion [30]. We can assert that we have obtained speciation, with species (in the sense
Figure 3: Power spectra of population fluctuations obtained from population dynamics in the stationary state for three different values of the resource recovery speed \( r \). All three obey a power law \( f^{-\alpha} \) with \( \alpha \approx 1 \). It is also interesting to note that at \( f \approx 10^{-2} \), a small peak emerges in the spectrum, which is related to the typical lifetimes of the creatures, as discussed in the text.

discussed in the Introduction) separated by their probability of movement. Moreover, since evolution occurs in fast bursts of transient dynamics, followed by long lasting stationary states (as discussed above in the context of Fig. 1), speciation (and not only population growth) proceeds as expected by the theory of punctuated equilibrium. Of course, we observe just one of the periods of stasis, mainly because the parameters defining the model are kept constant throughout the simulations, and possibly also because the low dimensionality of the genetic space. Should environmental conditions change, a new burst and a new consequent stationary state would ensue.

We studied different factors (genetic composition, physical distribution, resource consumption in the transitory states and departure from equilibrium) to try to determine if the surviving population would be large or small, but found no correlation. As will be discussed further on, the species with \( P = 1 \) (corresponding to the smaller population band) is less stable, it emerges less often in the simulations and occurs only in far from equilibrium scenarios.

Figure 3 shows the power spectrum obtained from the population dynamics. It displays a power-law behavior of the type \( f^{-\alpha} \) with \( \alpha \approx 1 \). This means that population fluctuations are very similar to the ones obtained in simpler models \([5]\) in which fitness has no physical interpretation but natural selection takes place nevertheless.

3.2. Altruistic behavior

The feeding function modeled by Eq. (4) might suggest that feeding would always be as large as possible, limited only by the capacity of the creature or the availability of the resource. However, interactions in the system manage to provide a setting in which individuals maintain the minimum energy needed to reproduce, instead of the maximum possible energy. Figure 4 portrays this effect by giving an estimation of the energy
distribution throughout the population. Observe that it stays just above the reproduction threshold, $T_R/M = 0.8$ (note the logarithmic scale), instead of being concentrated around $E/M = 1$. This has a beneficial effect on other creatures, which have more resource available for themselves, and on the resource, by preserving part of it for its recovery. Thus, we can assimilate this phenomenon to the emergence of a rudimentary altruistic behavior from egoistic rules, in the sense proposed by the “selfish gene” theory [21, 31].

3.3. Age distribution and death

We will now fulfill our promise to characterize spontaneous death more thoroughly. First, consider the probability of being dead at age $t$, which we will denote $P_D(t)$. It is the cumulative distribution of the probability of death at age $t$, $p_D(t) = dP_D(t)/dt$. The probability of death is also related to the rate of lethal events, $p_D(t)$. The latter represents dangers that may or may not kill an individual at each moment in time. Simply put, the rate of lethal events is to the probability of death as the outcome of a single coin toss is to the process of tossing a set of coins until they fall heads up. The probability of being dead is the proportion of coins that have stopped being tossed. Since it is a time-dependent Poisson process, we have:

$$p_D(t) = e^{-\int_0^t \rho_D(t')dt'},$$

$$= -\frac{d}{dt} \left( e^{-\int_0^t \rho_D(t')dt'} \right),$$

$$= \rho_D(t) e^{\int_0^t \rho_D(t')dt'},$$

$$= -\frac{d}{dt} \left( e^{\int_0^t \rho_D(t')dt'} \right).$$

Figure 4: Energy distribution measured over a stationary population. Three values of the recovery are shown, as indicated in the legend. Given that creatures eat until their energy reservoir is full when there is enough resource, one might expect this distribution to have a maximum at $E/M = 1$. However, note that most individuals have energy a little over the reproduction threshold $T_R/M = 0.8$. 

[21, 31]
and:

\[
\ln (1 - P_D(t)) = \int_0^t \rho_D(t') dt',
\]

\[
\Rightarrow \rho_D(t) = \frac{p_D(t)}{1 - P_D(t)},
\]

where in the left hand side of Eq. (8) we made use of the initial condition, \(P_D(0) = 0\).

As mentioned before, we assume that our creatures have a probability of dying spontaneously due to aging. This occurs in such a way as to provide a lifetime of \(\lambda = 1000\) time steps, and a slowly increasing \(P_D(t)\). We chose a logistic sigmoid function for this purpose:

\[
P_{\text{age}}D(t) = \frac{1}{1 + e^{-2(t-\lambda)/\lambda}},
\]

\[
\rho_{\text{age}}D(t) = \frac{2}{\lambda} \frac{1}{1 + e^{-2(t-\lambda)/\lambda}},
\]

\[
\rho_{\text{age}}E(t) = \frac{2}{\lambda} \frac{1}{1 + e^{-2(t-\lambda)/\lambda}}.
\]

Since individuals are also susceptible to dying from lack of energy, the total rate of deadly events is \(\rho_D(t) = \rho_{\text{age}}D(t) + \rho_{\text{age}}E(t)\). We can relate it to the age distribution in the population, which we will denote by \(z(t)\), where \(z(t)dt\) is the amount of creatures of age \(t\). If \(N\) is the total population, \(\int z(t)dt = N\). The age distribution evolves according to:

\[
z(t + dt) = z(t) - z(t)\rho_D(t)dt,
\]

\[
d_t z(t) = -z(t)\rho_D(t),
\]

\[
\Rightarrow \rho_D(t) = -d_t \ln (z(t)).
\]

Equation (15) expresses the total rate of deadly events in terms of age distribution. Age distributions are shown in Fig. 5 in the steady-state of evolution for three values of the resource recovery parameter \(r\). Since they are almost perfect exponentials, Eq. (15) implies that the death rate is constant. Its inverse is the average overall lifetime of an individual in the population. For all three cases, the obtained lifetime is \(\tau = 181 \pm 9\). We can therefore conclude that death is governed by lack of energy, not aging. Lifetime is enough for growth and reproduction, but well under the value \(\lambda = 1000\) available if only death by old age is considered. Again, the system has produced an altruistic behavior of sorts arising from selfish rules.

An interesting consequence of individual lifetime can be observed by analyzing the autocorrelation of the population. This can be calculated by applying the Wiener-Khinchin theorem [32] to the power spectrum shown in Fig. 3, obtaining the result shown in Fig. 6. In it, we observe that the system’s coherence drops quickly between \(t = 100\) and \(t = 300\), which we attribute to the fact that almost all creatures in the simulation die in this range of lifetimes. Observe that this value is compatible with the average lifetime \(\tau = 181 \pm 9\) calculated in a completely different way in Fig. 5. This means that large scale (population) coherence is a direct consequence of small scale (individual) dynamics. It also suggests that no intermediate scale structures appear (as could provide, for example, coexisting
Figure 5: Age distributions normalized to the total population in the stationary regime of the system’s evolution. Since they are almost perfect exponentials, the mortality rate given by using Eq. (15) is \( \rho_D = 0.055 \pm 0.003 \). Its inverse is the expected age, \( \tau = 181 \pm 9 \).

species), since these would interfere in the connection between the individual scale and the population. That is, the system loses coherence just on account of the creatures vital dynamics.

4. Mean-field dynamics

In this section we present a mean-field description of the resource dynamics, validated with the results obtained in simulations.

The average consumption rate of a creature \( c \) with genotype \( G = \{M, P, s\} \) and phenotype \( F = \{m, p\} \) is \( \langle C \rangle = C(F_c) = \mu m^\gamma + kmp + s \cdot m (1 - m/M) \). For a population of creatures \( P : c \in P \) of which a subset \( P(x) \) are at a site \( x \) at a certain time, their combined average consumption during a short interval \( \Delta t \), \( a(x) \), is:

\[
a(x) = \min \left( \sum_{P(x)} C(F_c) \Delta t, f(x) - 0.01K \right) .
\]

Note that \( \lim_{\Delta t \to 0} a(x) = \sum_{P(x)} C(F_c) dt \) unless the resource is strictly 0.01K. Thus, in the continuous time limit, the average resource dynamics can be obtained as follows:

\[
\frac{d\langle f \rangle_x}{dt} = \frac{1}{L^2} \sum_x \left[ r f(x) \left( 1 - \frac{f(x)}{K} \right) - \frac{a(x)}{dt} \right],
\]

\[
= \frac{1}{L^2} \sum_x r f(x) \left( 1 - \frac{f(x)}{K} \right) - \frac{N}{L^2} \frac{1}{N} \sum_x \sum_{P(x)} C(F_c),
\]

\[
= r \left[ \langle f \rangle_x \left( 1 - \frac{\langle f \rangle_x}{K} \right) - \frac{\text{var}(f)}{K} \right] - \frac{N}{L^2} \langle C \rangle p .
\]
Figure 6: Autocorrelation obtained from the population power spectrum of Fig. 3. Given the different population sizes for different values of $r$, the autocorrelation is measured on deviations from the population average. We observe an approximately constant self-correlation until $t \approx 100$, where it drops significantly smaller values. Noting that the drop happens at around the same time as the average lifetime found from Fig. 5, we attribute the loss of coherence to the renewal of individuals in the population (and thus the renewal of the population itself).

In Eq. (17) we have assumed that $r$ and $K$ are homogeneous throughout the landscape for simplicity, but of course this does not need to be the case, and the averages should be taken accordingly. If we normalize the resource by its local carrying capacity $K$, defining $n = f/K$, the dynamics becomes:

$$\frac{d\langle n \rangle_x}{dt} = r \langle n \rangle_x (1 - \langle n \rangle_x) - \text{var}(n) - \frac{N}{L^2} \frac{C}{K}.$$  \hspace{1cm} (18)

This equation resembles the prey’s dynamics in a predator-prey system, where the normalized resource takes the role of the prey, and $N/L^2$ is the normalized predator population. The interaction is a consumption of prey at a rate given by the last term in Eq. (18), represented by $C$, since this comes from the consumption rate $a$, which is explicitly related to $f$, the “actual” prey.

Figure 7 shows both sides of Eq. (18) as measured from numerical simulations of the system. Two simulations are shown, identical in all except the initial population. The first one, denoted by $n_1$, had an initial population of only 1 individual. We observe that it starts and stays in equilibrium ($d_t \langle n \rangle_x = 0$). The second one, denoted by $n_{1000}$ has an initial population of 1000 and achieves equilibrium after a transient of $t \approx 1000$ time steps. After that, all curves follow similar paths. In both cases the measured left and right-hand sides of Eq. (18) coincide almost perfectly, supporting the validity of the mean-field dynamics. The initial difference between the $n_{1000}$ curves are a product of different sampling rate when measuring $d_t \langle n \rangle_x$ and the corresponding right-hand side.
To complete the description of the dynamics and calculate the system’s equilibria we need an equation for $d_t N$. The dynamics of $N$ is given by the balance of births and deaths in the population. Since both of these are controlled by the energy of the creatures, which fluctuates because of the resource, births and deaths are stochastic processes. For this reason we must first advance the understanding of the resource dynamics and its fluctuations around its mean value.

4.1. Resource distribution

Figure 8 (left, thick lines) shows the resource distribution in equilibrium for different values of $r$. Since the distributions have sharp spikes at both extremes of the resource range, 0 and 1, we prefer to transform them by taking the logarithm and multiplying by $n(1-n)$ to eliminate any possible divergences. The right panel (thick lines) shows these transformed resource distributions.

Since creatures need to be in a cell to consume its resource, the distribution they perceive is the conditional distribution $\rho(n|o)$ of resource, given that the cell is occupied. Figure 8 (left, thin lines) shows the corresponding perceived distributions as measured from the simulations (for the same values of $r$). Although their concavities are different, as well as their divergences, the perceived distributions display a shape similar to the true ones when transformed accordingly, as shown in Fig. 8 (right, thin lines).

The true and perceived distributions are related by consumption and recovery, and in equilibrium we can derive their relation explicitly. First we decompose the temporal evolution of $\rho$ according to its dependence on the resource $n$:

$$
\frac{d\rho(n)}{dt} = \frac{d\rho(n)}{dn} \frac{dn}{dt}.
$$

(19)
Figure 8: True and perceived resource distributions in the stationary state for several values of recovery rate. Right: Transformed distributions, to remove divergences. Both distributions are practically the same for different values of the resource recovery speed, save for increasing fluctuations as the resource recovery speed decreases. The perceived distribution is always more noisy than the true one because of the reduced sample size. Interestingly, both transformed distributions exhibit the same sinusoidal (or cubic) behaviour, although no analytical explanation of this was found.

Conditional probabilities allow us to write this in terms of occupied (o) and not occupied (¬o) sites, and equal to zero in equilibrium:

\[
\frac{d\rho(n)}{dt} = \frac{d\rho(|o)}{dn} \rho(o) \frac{dn}{dt} |o + \frac{d\rho(|¬o)}{dn} \rho(¬o) \frac{dn}{dt} |¬o = 0. \tag{20}
\]

Now, the dynamics of unoccupied sites is just the logistic recovery, \(d_n|¬o = rn(1 - n)\), while that of occupied sites is a balance of recovery and consumption, \(d_n|o = d_n|¬o - ⟨C⟩ r / K\). Using these in Eq. (20), together with the fact that \(\rho(o)\rho(|o) + ρ(¬o)\rho(|¬o) = ρ(n)\), elementary algebra allows to derive:

\[
rn(1 - n) \frac{d\rho(n)}{dn} = \frac{⟨C⟩ r N}{K} \frac{d\rho(|o)}{dn} \frac{dn}{dt}. \tag{21}
\]

Both sides of Eq. (21) are shown in Fig. 9, calculated from measurements performed on simulations for the three values of \(r\) shown in previous figures. We can see that both bundles of curves coincide, within fluctuations, for a very wide range of values of \(n\), separating when \(n\) approaches unity. It is worth noting that Eq. (21) fails as \(n \to 1\), which indicates that this region never reaches equilibrium fully.

5. Discussion and conclusions

We have proposed and analyzed an agent-based model in which the energy available to living creatures determines the processes that characterize life: birth, reproduction and death. The dynamics of energy itself is defined by genes and indirect competition with other creatures through a common resource. Genes determine body size (i.e., mass and energy reservoir size), movement speed across the domain of the resource and growth.
speed from birth to adulthood. Small, random mutations upon asexual reproduction allow different genetic combinations to compete with one another and therefore natural selection emerges with only a shared resource and random mutations.

It is interesting to relate an intuitive notion of the intensity of natural selection with the system’s capacity to sustain a population. In this sense, natural selection is the force that eliminates individuals while the recovery speed of the resource keeps them alive. As opposed to some artificial selection models in which a quality of individuals is explicitly given and so the best can always be picked, natural selection doesn’t guarantee a surviving population. This is seen for simulations starting with a large genetic diversity and small values of resource recovery speed. For larger recovery speeds, an initial mass extinction is observed, followed by an explosion of the successful genes. Even when the resource recovers instantly (i.e., the resource is limitless), some genes are still unfit to survive and the initial extinction is still observed. We conclude that natural selection decreases with the capacity of a system but is always present to some degree.

After the initial extinction and population explosion, we found that two distinct genetic groups emerge in the stable state of the system. They are separated by their moving speed but indistinguishable in their mass. The growth speed, on the other hand, turned out to be a gene with which fitness increased monotonously, so it never reached a stationary value but didn’t affect the system. Since a faster movement implies more energy consumption, fast populations were found to be smaller. However, only one of these “species” survived in each realization. Species were observed to compete briefly in the initial transient period, as shown in Fig. 10, but one of the groups always eliminated the others in accordance with the competitive exclusion principle [30] and punctuated

Figure 9: Dynamics of the resource, showing both sides of Eq. (21) for three different values of the resource recovery speed. Each side of the equation behaves in the same way for different values of the recovery speed. Despite larger fluctuations in the right hand side (due to a smaller sampling size), the equation holds up to values of $n \approx 0.7$. For values of $n \approx 1$, the curves for the different sides of the equation take very different values (and even have opposite derivatives), indicating that the system is not in equilibrium in this region.
equilibrium.

Simulations starting with a single individual stayed in equilibrium throughout the evolution of the system, even during population growth. They also exhibited stationary states at resource recovery speeds slower than ones at which larger populations completely died out. Additionally, these populations always converged to low energy consumption populations. This suggests that the high energy consumption species found earlier are a result of a non-equilibrium situation, and thus the system is very sensitive to initial conditions. Another noteworthy conclusion is that, according to this model, life beginning from a single organism seems to be more likely than life beginning from a larger group, even at the cost of genetic diversity.

The final result we would like to remark from simulations is that, even though the rules followed by each individual were set to be selfish, the system managed to evolve into an altruistic setting, at least in the limited sense that probability of death was not minimized. This was observed in two instances. The first regards the amount of resource consumed, where creatures were supposed to eat enough to fill their energy reservoirs but ended up keeping them at the minimum necessary to reproduce. The second relates to lifetime, that could end by old age after around ten reproductions, but ended up limited to the time necessary for just one reproduction. This makes perfect sense because the population must remain constant, but it also shows that the probability of producing offspring, usually associated with fitness, is not maximized under natural selection with a shared resource. The population size, used as a fitness measure in works such as [14, 15] does not seem to be maximized either, since for the same conditions
we have found realizations that attain a stationary state with a large population and low consumption, and others with small populations and high consumption. It is worth emphasizing that this single-offspring situation was not programmed in the model, but emerged from constraints and interactions.

Although the model is somewhat artificial (agents are, quite literally, spherical cows), our study demonstrates that all of the phenomena mentioned above are actually parts of a whole. They can be obtained by considering that the genetic code of an individual determines how energy is obtained and employed to harvest more. An open question is whether more detailed models of genetic dependence produce more realistic results in terms of surviving genes. Another is if predator-prey systems are equivalent to resource-population systems where the resource is controlled by genes in the same way as the population.

We have taken an initial step towards answering the first question by adding a rudimentary intelligent behavior to the population. With it, movement speed depends linearly on the amount of energy in a creature’s reservoir instead of being constant. This allows creatures to move more when they are hungry and less when they are full and vice-versa. The first case is more energy efficient but dangerous when there is little food nearby, while the second is less efficient but could be safer in terms of keeping energy levels high. Both cases have their advantages and drawbacks and therefore we cannot predict which will appear. However, it is known that the flatworm *Dugesia tigrina* exhibits this rudimentary intelligence and behaves as in the first case [33]. In our simulations, we found that the creatures in the stationary state were those with genes that made them move more when they were hungry and less when they were not. Since this agrees with the real-life scenario we conclude that, as genetic models get better, the simulation results will also get more realistic.

Finally, we would like to mention the consequences of the mean-field dynamics analysis. Much can be read from Eq. (18), describing the dynamics of the mean normalized resource, which we copy here:

\[ \frac{d\langle n \rangle}{dt} = r \left( \langle n \rangle \left(1 - \langle n \rangle\right) - \text{var}(n) \right) - \frac{N \langle C \rangle P}{KL^2}. \]  

First, Eq. (22) resembles the dynamics of a prey, \( \langle n \rangle \), faced with a population of \( N \) predators consuming them at a rate \( \langle C \rangle/ KL^2 \), as mentioned above. This interpretation suggests that our model is similar to a predator-prey system. It also makes altruistic behavior (keeping energy at the minimum necessary) beneficial not only for the rest of the population (other predators), but also for the resource (the “prey”). Of course, this is a consequence of the population needing the resource (or predators needing prey) to survive. However, since we observed populations with higher energy consumption, we can conclude that this form of altruism isn’t necessary, merely better.

Second, note that the growth term of the resource in Eq. (22) is reduced by the variance of the resource. This means that smaller fluctuations in the resource distribution at the cost of a smaller expected value, sometimes will be more beneficial for the population. This effect is definitely non-negligible since an estimation of the population that does not include this term, \( N \approx rf(1 - f)L^2/C \approx 500 \) (\( r = 0.01, \ P = 0.33, \ f = 0.85 \)), results in an error of around 70% compared to the observed population of \( N \approx 300 \).

Finally, since all real-world food sources have a finite size, their growth is limited. Therefore a logistic equation, as a second-order expansion, is the simplest realistic as-
sumption we can make. The variance term in Eq. (22) stems from the nonlinear term in this expansion and therefore all real-world resources will suffer from it. Moreover, resource distribution is shaped by consumption, so the efficiency of a species can be characterized through that variance. In this sense, more “intelligent” (better distributed) consumption will lead to faster resource recovery and therefore a larger population. This may explain why the first great revolution human beings experienced, the invention of agriculture, was so important [34, 35]. Since we could suddenly control the dynamics of our source of energy, we could minimize variance much more effectively, ensuring that everyone has enough food and allowing the population to grow.

We have set out trying to describe a model that produces natural selection from first principles, namely energy conservation and its flow through living beings. In the process, we have discovered many other known results and characteristics associated with the theory of evolution. We consider this a successful feat, but there is still much work to be done. It is not clear how exactly these phenomena emerge from the underlying process, nor how they are influenced when specifics are changed. We hope to address these questions and any new ones that may appear in future work.

Acknowledgments

We recognize partial support through the following grants: CONICET PIP 112-2017-0100008 CO, UNCUYO SIIP 06/C546, ANPCyT PICT-2018-01181. We also thank S. Bouzat for fruitful discussions.

References

[1] Charles Darwin. *On the origin of species*. John Murray, London, 1859.
[2] Eric Smith and Harold J. Morowitz. *The origin and nature of life on Earth*. Cambridge UP, 2016.
[3] Barbara Drossel. Biological evolution and statistical physics. *Advances in Physics*, 50:209–295, 2001.
[4] Stuart A. Kauffman. *The origins of order: Self-organization and selection in evolution*. Oxford UP, 1993.
[5] Per Bak and Kim Sneppen. Punctuated equilibrium and criticality in a simple model of evolution. *Physical Review Letters*, 71:4083, 1993.
[6] Guillermo Abramson and Damián Zanette. Statistics of extinction and survival in Lotka-Volterra systems. *Physical Review E*, 57:4572, 1998.
[7] Samira El Yacoubi and Sebastien Gourhière. A cellular automata model for adaptive sympatric speciation. In Samira El Yacoubi, Bastien Chopard, and Stefania Bandini, editors, *Cellular Automata*, volume 4173 of *Lecture Notes in Computer Science*, pages 287–296. Springer, 2006.
[8] J. J. Heijnen. Impact of thermodynamic principles in systems biology. In C. Wittmann and R. Krull, editors, *Biosystems Engineering II*, volume 121 of *Advances in Biochemical Engineering/Biotechnology*, pages 139–162. Springer, 2010.
[9] S. A. L. M. Kooijman and T. A. Troost. Quantitative steps in the evolution of metabolic organisation as specified by the Dynamic Energy Budget theory. *Biological Reviews*, 82:113–142, 2007.
[10] A. I. Oparin. *The origin of life*. Dover, 1953.
[11] Carl Woese. The universal ancestor. *Proc. Natl. Acad. Sci. USA*, 95:6854–6859, 1998.
[12] Hinrich Arnoldt, Steven H. Strogatz, and Marc Timme. Toward the Darwinian transition: Switching between distributed and speciated states in a simple model of early life. *Physical Review E*, 92:052909, 2015.
[13] Ayelén Prado. Evolución biológica en el origen de la vida: Transmisión de genes horizontal versus vertical. Master’s thesis, Instituto Balseiro, Universidad Nacional de Cuyo, 2018.
[14] Filippo Menczer and Richard K. Belew. From complex environments to complex behaviors. *Adaptive behavior*, 4(3-4):317–363, 1996.
Filippo Menczer and Richard K. Belew. Latent energy environments. In R. K. Belew and M. Mitchell, editors, *Adaptive individuals in evolving populations: Models and algorithms*. Addison-Wesley, 1996.

Hugh Trenchard and Matjaž Perc. Energy saving mechanisms, collective behavior and the variation range hypothesis in biological systems: A review. *Biosystems*, 147:30–66, 2016.

Melanie Mitchell and Stephanie Forrest. Genetic algorithms and artificial life. *Artificial life*, 1(3):267–289, 1994.

Joel Cracraft. Species concepts in theoretical and applied biology: A systematic debate with consequences. In Quentin D. Wheeler and Rudolf Meier, editors, *Species concepts and phylogenetic theory: A debate*, chapter 1, pages 3–14. Columbia University Press, 2000.

James Mallet. A species definition for the Modern Synthesis. *Trends in Ecology & Evolution*, 10(7):294–299, 1995.

Stephen Jay Gould and Niles Eldredge. Punctuated equilibria: The tempo and mode of evolution reconsidered. *Paleobiology*, 3:115–151, 1977.

Richard Dawkins. *The selfish gene*. Oxford UP, 1976.

Kunihiko Kaneko. *Life: An introduction to complex systems biology*. Springer, 2006.

James D. Murray. *Mathematical Biology*. Springer-Verlag New York, 2002. Third edition, chapter 1.1, pages 3-5.

F. J. Richards. A flexible growth function for empirical use. *Journal of Experimental Botany*, 10(2):290–301, 1959.

Max Kleiber et al. Body size and metabolism. *Hilgardia*, 6(11):315–353, 1932.

D. M. Raup. Biological extinction in Earth history. *Science*, 231:1528–1533, 1986.

Niles Eldredge and Stephen Jay Gould. Punctuated equilibrium prevails. *Nature*, 322:211–212, 1988.

Sodré Gonçalves de Brito Neto, Everton Fernando Alves, and Mariana Cavalcante e Almeida Sá. Speciation in real time and historical-archaeological and its absence in geological time. *Academia Journal of Scientific Research*, 5(7):188–196, 2017.

Benjamin Bonnfeur, Stephen McLoughlin, and Vivi Vajda. Fossilized nuclei and chromosomes reveal 180 million years of genomic stasis in royal ferns. *Science*, 343:1376–1377, 2014.

James D. Murray. *Mathematical biology*. Springer, 2002.

George C. Williams. *Adaptation and natural selection*. Princeton UP, 1966.

N. G. van Kampen. *Stochastic processes in physics and chemistry*. North Holland, 2007.

I. Sheiman, E. Zubina, and Natalia Kreshchenko. Regulation of the feeding behavior of the planarian *Dugesia (Girardia) tigrina*. *Journal of Evolutionary Biochemistry and Physiology*, 38:414–418, 07 2002.

Vered Eshed, Avi Gopher, Timothy B Gage, and Israel Hershkovitz. Has the transition to agriculture reshaped the demographic structure of prehistoric populations? New evidence from the levant. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists*, 124(4):315–329, 2004.

Graeme Barker. *The agricultural revolution in prehistory: Why did foragers become farmers?* Oxford UP, 2009.