An evolutionary advantage of cooperation

Ole Peters[1,2,†] and Alexander Adamou[1,‡]

[1]London Mathematical Laboratory, 8 Margravine Gardens, London W6 8RH, UK
[2]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, 87501 NM, USA
[†]o.peters@lml.org.uk, [‡]a.adamou@lml.org.uk

March 05, 2019

Abstract

Cooperation is a persistent behavioral pattern of entities pooling and sharing resources. Its ubiquity in nature poses a conundrum: whenever two entities cooperate, one must willingly relinquish something of value to the other. Why is this apparent altruism favored in evolution? Classical treatments assume a net fitness gain in a cooperative transaction which, through reciprocity or relatedness, finds its way back from recipient to donor. Our analysis makes no such assumption. It rests on the insight that evolutionary processes are typically multiplicative and noisy. Fluctuations have a net negative effect on the long-time growth rate of resources but no effect on the growth rate of their expectation value. This is an example of non-ergodicity. By reducing the amplitude of fluctuations, pooling and sharing increases the long-time growth rate for cooperating entities, meaning that cooperators outgrow similar non-cooperators. This constitutes a fundamental mechanism for the evolution of cooperation. Its minimal assumptions make it a candidate explanation of cooperation in settings too simple for other explanations, such as emergent function and specialization, to be plausible. One such example is the transition from single cells to early multicellular life.

1

1 Introduction

They give that they may live, for to withhold is to perish.
K. Gibran

Nowak (2006, p. 1563) concludes his review of the mechanisms of cooperation with the words: "Perhaps the most remarkable aspect of evolution is its ability to generate cooperation in a competitive world." Indeed, life is full of cooperative structure. We living beings exist not as minimal self-reproducing chemical units, but as cells, organisms, families, herds, companies, institutions, nations, and so on. Cooperation – the persistent behavioral pattern of entities pooling and sharing their resources – is ubiquitous in nature.

This ubiquity is puzzling because pooling and sharing seem *prima facie* to involve altruism. The temporarily better-off entity in a cooperating pair must willingly relinquish something of value to the worse-off entity to maintain the cooperative pact. If naked altruism is an unsatisfactory explanation of evolved behavior, then we must elucidate the benefit derived by the better-off entity in such an arrangement.

Classical explanations involve two ideas. The first is that a cooperation benefit exists between two entities. Specifically, the fitness gain of the recipient exceeds the fitness cost to the donor. The second is that, over time, this net increase in fitness finds its way back to the donor. This can happen through reciprocity, where past donors become future recipients, or through relatedness, where the recipient carries genetic material that the donor is interested in propagating. Nowak (2006) offers a comprehensive account of this approach, delineating five cooperative arrangements possible when a cooperation benefit exists, and five corresponding conditions under which cooperation is favored in evolution.

Our aim here is to postulate a mechanism through which this cooperation benefit, or net fitness gain, arises. This is important because classical analyses are predicated on its existence.

We start from the basic model of a living entity as something which self-reproduces with temporal fluctuations. Treating biomass as a multiplicative stochastic process yields two exponential growth rates: that achieved in the long-time limit, on which fluctuations have a negative effect; and that achieved in the many-entities limit, unaffected by fluctuations. This is a manifestation of non-ergodicity.

We hypothesize that repeated pooling and sharing is beneficial because, by reducing the net effect of fluctuations, it increases the long-time growth rate at which cooperating entities self-reproduce. Cooperation is observed in nature because those who do it outgrow those who don't. We identify this

increase in the growth rate of biomass with the net fitness gain of classical treatments.

Fitness has many definitions in the biological literature (Orr, 2009). While commonly agreed to refer to the ability of living organisms to survive and reproduce, no precise definition of it has achieved consensus. Of all the definitions proposed, our work relates most closely to geometric mean fitness, on which the effects of fluctuations are recognized (Lewontin and Cohen, 1969; Gillespie, 1977; Orr, 2009). To avoid confusion, we speak here of growth rates rather than fitness, since they are defined unequivocally. Among entities undergoing noisy multiplicative growth, that whose time-average growth rate is highest will, over time, come to dominate its environment.

Motivating this study is a particular evolutionary phenomenon: the transition to multicellularity. This occurs when a species of non-cooperating single cells mutates to a species of multicellular organisms, sharing nutrients through common membranes. Often this is explained by a different type of fitness gain: the emergence of new function. An agglomeration of cells may, for example, develop the ability to swim up a nutrient gradient or funnel resources towards itself (Short et al., 2006; Roper et al., 2013). Implicit in such explanations is a degree of complexity not present in early unicellular life. The performance of specialized tasks requires many cells to be assembled in delicate designs, whose spontaneous development would be extremely improbable. To be a credible theory, evolution must explain not only the rich tapestry of cooperative structure we observe now, but also early cooperative steps, such as from single cell to cell pair, where new function cannot be relied upon. A universal mechanism for the evolutionary advantage of cooperation is needed.

We present our work as follows. In Sec. 2 we introduce a simple mathematical model of noisy multiplicative growth and summarize its relevant properties. In Sec. 3 we describe a cooperation protocol in which entities grow, pool, and share biomass. In Sec. 4 we show that, under certain conditions, entities which cooperate increase the time-average growth rate of their biomass. This, we hypothesize, is a universal explanation of the existence of cooperation. In Sec. 5 we discuss generalizations of our model. We offer concluding remarks in Sec. 6.

2 Noisy multiplicative growth

Let $x_i(t)$ represent the biomass of an entity i at time t. Biomass generates more of itself. For example, a cell collects nutrients and other matter, then splits into two cells, which split into four cells, and so on. This happens

stochastically. Some entities thrive, perhaps in safe and plentiful environments, while others have their growth hampered, sometimes terminally.

While we speak typically of the biomass of an entity, our analysis is agnostic to replacement of 'biomass' by 'population,' 'resources,' 'wealth,' and the like; and of 'entity' by 'cell,' 'organism,' 'colony,' 'herd,' 'society,' and so on. Cooperation in nature occurs in many domains and at many scales.

A common and general model of noisy self-reproduction is geometric Brownian motion. In simple terms, the change in biomass over a short time step is a normally distributed random multiple of the existing biomass. More formally, $x_i(t)$ follows the Itô drift-diffusion process,

$$dx_i = x_i(\mu dt + \sigma dW_i), \tag{1}$$

where μ is the drift and σ is the volatility. The dW_i are independent and identically-distributed random increments of the Wiener process, which are normal with zero mean and variance dt.

Geometric Brownian motion is a universal model because it is an attractor for more complex models that exhibit multiplicative growth (Aitchison and Brown, 1957; Redner, 1990). It is a model of unconstrained growth. Self-reproduction limited by resource or space constraints or by predation would be poorly described by (Eq. 1). The water lily of the famous riddle, recounted in (Meadows et al., 1972), stops growing exponentially once it covers the pond.

Over time T, each biomass experiences a random exponential growth rate, defined as

$$g(x_i, T) \equiv \frac{1}{T} \ln\left(\frac{x_i(T)}{x_i(0)}\right). \tag{2}$$

Imagine starting many cell cultures in separate petri dishes and watching their biomasses evolve according to (Eq. 1) for time T. Assume the dishes are large enough and T short enough that growth does not slow for want of agar. The observed exponential growth rates would be drawn from a normal distribution with mean $\mu - \sigma^2/2$ and variance σ^2/T:

$$g(x_i, T) \sim \mathcal{N}\left(\mu - \frac{\sigma^2}{2}, \frac{\sigma^2}{T}\right). \tag{3}$$

The expectation value or ensemble average of the biomass is defined as

$$\langle x \rangle \equiv \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} x_i. \tag{4}$$

Its evolution is computed by noting that (Eq. 1) implies $d\langle x \rangle = \mu \langle x \rangle \, dt$. This describes exponential growth of $\langle x \rangle$ at a rate equal to the drift. We will call this the *ensemble-average growth rate* and denote it by

$$g(\langle x \rangle) = \mu. \tag{5}$$

The physical meaning of this quantity is worth making explicit. It is the growth rate of the large-sample limit of mean biomass. Since (Eq. 1) is not an ergodic process, this observable is *a priori* uninformative of the temporal behavior of a given trajectory. We don't mean by this the trivial observation that individual trajectories are stochastic, while their expectation value is not. Rather, there are two fundamentally different ways of removing stochasticity from (Eq. 2): we may compute (Eq. 5); or we may consider a single long trajectory, and let time remove randomness from the growth rate.

If only a single system is to be modeled, then (Eq. 5) is, in essence, fiction: the expectation value is an average over imagined parallel universes, where the randomness in (Eq. 1) plays out in all its different possible ways.

One might guess that the growth rate observed in an individual trajectory will converge to (Eq. 5) over time, but that is simply a common conceptual error. Instead, the non-ergodicity of (Eq. 1) manifests itself such that the growth rate observed in a single trajectory converges to a different value, which we call the *time-average growth rate*. This is the almost sure limit of (Eq. 3) as $T \to \infty$,

$$\bar{g}(x_i) \equiv \lim_{T \to \infty} g(x_i, T) = \mu - \frac{\sigma^2}{2}. \tag{6}$$

We note that $\bar{g}(x_i) = \langle g(x_i, T) \rangle$, meaning that $g(x_i, T)$ is an ergodic observable (specifically, one whose expectation value reflects what happens over time in a single trajectory) for the non-ergodic process of noisy multiplicative growth (Peters and Gell-Mann, 2016).

We see in nature what has survived. In our model, the entity with the highest time-average growth rate will, regardless of its ensemble-average growth rate, come to dominate the biomass in the system over time. The ratio of its biomass to that of other entities will grow exponentially. Strategies which maximize $\bar{g}(x_i)$, and not necessarily $g(\langle x \rangle)$, will confer an evolutionary advantage on their adherents. For any entity, $\bar{g}(x_i)$ is less than $g(\langle x \rangle)$ by the fluctuation-sensitive term $\sigma^2/2$. This standard result of Itô calculus (van Kampen, 1992, Ch. 5) has found its way into the biological literature (Gillespie, 1977, Eq. 1). It suggests that strategies which reduce volatility should be favored in evolution and observed in nature.

The inequality of the ensemble-average and time-average growth rates in geometric Brownian motion, discussed in (Peters and Gell-Mann, 2016;

Adamou and Peters, 2016) in the context of economics, is known by evolutionary biologists. A clear description is found in (Lande, 2007, p. 1836). The difference is often labelled as that between arithmetic and geometric mean growth rates (Lewontin and Cohen, 1969) or between arithmetic and geometric mean fitnesses (Orr, 2009). The positive effect of volatility reduction on geometric mean fitness is also recognized, as noted in (Gillespie, 1977) and (Orr, 2009, Box 3). It is precisely this effect in the context of cooperation that we explore here.

3   Cooperation protocol

Having established these properties of (Eq. 1), we now introduce our model of cooperation. Our reference point is a population of N non-cooperating entities, such as single-celled organisms, whose biomasses follow geometric Brownian motions with identical drift and volatility but with independent realizations of the noise. In other words, (Eq. 1) with $i = 1 \ldots N$.

Consider a discretized version of (Eq. 1), such as would be used in a numerical simulation. The non-cooperators grow according to

$$\Delta x_i(t) \;\; = \;\; x_i(t)\left(\mu\Delta t + \sigma\xi_i\sqrt{\Delta t}\right), \qquad (7)$$

$$x_i(t + \Delta t) \;\; = \;\; x_i(t) + \Delta x_i(t), \qquad (8)$$

where ξ_i are independent standard normal variates, $\xi_i \sim \mathcal{N}(0,1)$.

The cooperation mechanism, summarized pictorially for $N = 2$ in Fig. 1, is as follows. We imagine a mutation that hard-wires cooperation into the entities. Think, for example, of a mutation that causes cells to agglomerate by sharing their membranes. Previously independent entities with biomasses $x_i(t)$ now find themselves cooperating. We label their biomasses $y_i(t)$ to distinguish them from the equivalent non-cooperators. For simplicity we assume equal sharing, $y_i(t) = y^{(N)}(t)$ for all i, where $y^{(N)}(t)$ denotes the per-entity biomass for N equal cooperators.

In the discrete-time picture, each time step involves a two-phase process. First there is a growth phase, analogous to (Eq. 7), in which each cooperator increases its resources by

$$\Delta y_i(t) = y_i(t)\left(\mu\Delta t + \sigma\xi_i\sqrt{\Delta t}\right). \qquad (9)$$

This is followed by a cooperation phase, replacing (Eq. 8), in which resources are pooled and shared equally among the cooperators,

$$y_i(t + \Delta t) = \frac{1}{N}\sum_{i=1}^{N}\left(y_i(t) + \Delta y_i(t)\right) = y_i(t) + \frac{1}{N}\sum_{i=1}^{N}\Delta y_i(t), \qquad (10)$$
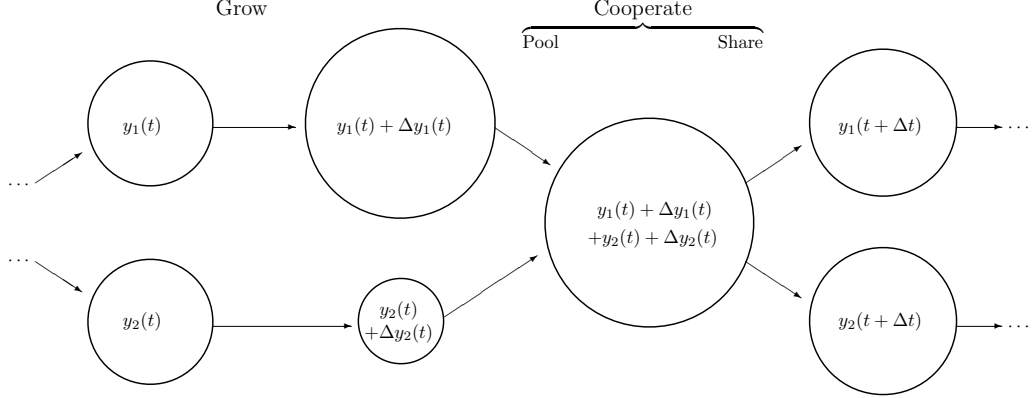
6

Figure 1: Cooperation dynamics. Two cooperators start each time step with equal resources. Then they grow independently according to (Eq. 9). Then they cooperate by pooling resources and sharing them equally, according to (Eq. 10). Then the next time step begins.

where the second equality follows from the equality of the $y_i(t)$. This is equivalent to equal sharing of the total of the individual fluctuations,

$$\Delta y^{(N)}(t) = \frac{1}{N} \sum_{i=1}^{N} \Delta y_i(t). \tag{11}$$

Cooperation has no direct cost in this protocol. In reality, pooling and sharing often require a coordinating mechanism. For example, large organisms have circulatory systems to redistribute nutrients, and human societies have administrative systems to redistribute resources. It is possible for such mechanisms to have costs that make otherwise beneficial cooperation unviable. We do not expect our model to describe well situations where such effects are important. Equally, we ascribe no direct benefit to cooperation. In our basic setup, costs and benefits emerge as the effects of cooperation on time-average growth rates.

Substituting (Eq. 9) into (Eq. 11) yields the dynamic followed by the biomasses of each cooperator,

$$\Delta y^{(N)}(t) = y^{(N)}(t) \left(\mu \Delta t + \frac{\sigma}{N} \sum_{i=1}^{N} \xi_i \sqrt{\Delta t} \right). \tag{12}$$

Sums of independent normal variates are normal. We define a standard normal variate,

$$\xi^{(N)} \equiv \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \xi_i \sim \mathcal{N}(0, 1), \tag{13}$$

which allows us to rewrite (Eq. 12) as

$$\Delta y^{(N)}(t) = y^{(N)}(t) \left(\mu \Delta t + \frac{\sigma}{\sqrt{N}} \xi^{(N)} \sqrt{\Delta t} \right). \tag{14}$$

Thus the net effect of N individual fluctuations with pooling and equal sharing is a single equivalent fluctuation, whose amplitude is $1/\sqrt{N}$ times that of the individual fluctuations. Substituting into (Eq. 12) and letting the time increment $\Delta t \to 0$, we recover a stochastic differential equation of the same form as (Eq. 1), but with the volatility reduced from σ to $\sigma/\sqrt{N}$:

$$dy^{(N)} = y^{(N)} \left(\mu dt + \frac{\sigma}{\sqrt{N}} dW^{(N)} \right). \tag{15}$$

4 Analysis and solution of the cooperation puzzle

The expectation values of the biomasses of a non-cooperator, $\langle x_1(t) \rangle$, and the corresponding cooperator, $\langle y_1(t) \rangle$, with the same initial biomass, $x_1(0)$, are identical: $x_1(0) \exp(\mu t)$. From this perspective there is no incentive for cooperation. Moreover, immediately after the growth phase, (Eq. 9), the better-off entity in a cooperating pair, say $y_1(0) > y_2(0)$, could increase its future expectation value from $[(y_1(0) + y_2(0))/2] \exp(\mu t)$ to $y_1(0) \exp(\mu t)$ by breaking the cooperative pact. Analyzing the growth of the expectation value gives no reason for cooperation to arise and, if it does arise, good reasons for it to end. From this perspective, cooperation looks fragile and its frequent observation in nature seems puzzling.

The solution of the puzzle comes from changing perspectives and considering the time-average growth rate instead of the ensemble-average growth rate. We know from (Eq. 6) that non-cooperating entities grow at $\bar{g}(x_i) = \mu - \sigma^2/2$ over long time. The analogous growth rate for the cooperative dynamic, (Eq. 15), is found by changing σ to $\sigma/\sqrt{N}$ in (Eq. 6) to get

$$\bar{g}\left(y^{(N)}\right) = \mu - \frac{\sigma^2}{2N}. \tag{16}$$

For any non-zero volatility, cooperators have faster time-average growth rates than non-cooperators. The premium increases with the number of cooperators as $1 - 1/N$,

$$\bar{g}\left(y^{(N)}\right) - \bar{g}(x_i) = \frac{\sigma^2}{2} \left(1 - \frac{1}{N} \right), \tag{17}$$

implying that larger cooperatives are favored over smaller ones.

From the perspective of our model, we see that cooperators will eventually dominate the environment and that cooperation will become ubiquitous. The effect is illustrated in Fig. 2 by direct simulation of (Eq. 7)–(Eq. 8) and (Eq. 12). Notice the nature of the Monte Carlo simulation in Fig. 2. No ensemble is constructed. Only individual trajectories are simulated and run for a time that is long enough for statistically significant features to emerge from the noise. This method teases out of the dynamics what happens over time.

Simulating an ensemble and averaging over members to remove noise is a different process which tells a different story. Features emerging from the noise when averaging over an ensemble do not, in general, also emerge over time. From (Eq. 16) we see that the time-average growth rate of a cooperator approaches the expectation-value growth rate of a non-cooperator (or, indeed, that of a member of a finite cooperative) only as the number of cooperators grows large,

$$\lim_{N \to \infty} \bar{g}\left(y^{(N)}\right) = \mu. \tag{18}$$

The pink dashed line in Fig. 2 plots the temporal evolution of the expectation value of the biomass of a non-cooperator, which bears little resemblance to that of the biomasses of individual non-cooperators (green) and members of small cooperatives (blue).

5 Generalizations of the model

5.1 Idiosyncratic entities

Real cooperatives have members of differing abilities as well as differing fortunes. The latter we model already as different realizations of the noise in (Eq. 1). The former we can treat by generalizing (Eq. 1) so that the entities have idiosyncratic drifts and volatilities,

$$dx_i = x_i(\mu_i dt + \sigma_i dW_i), \tag{19}$$

for $i = 1 \dots N$. Some entities will have higher individual time-average growth rates than others. This raises questions. Does it benefit leaders to share with laggards? When should a non-cooperator join a cooperating group? When should the group allow it?

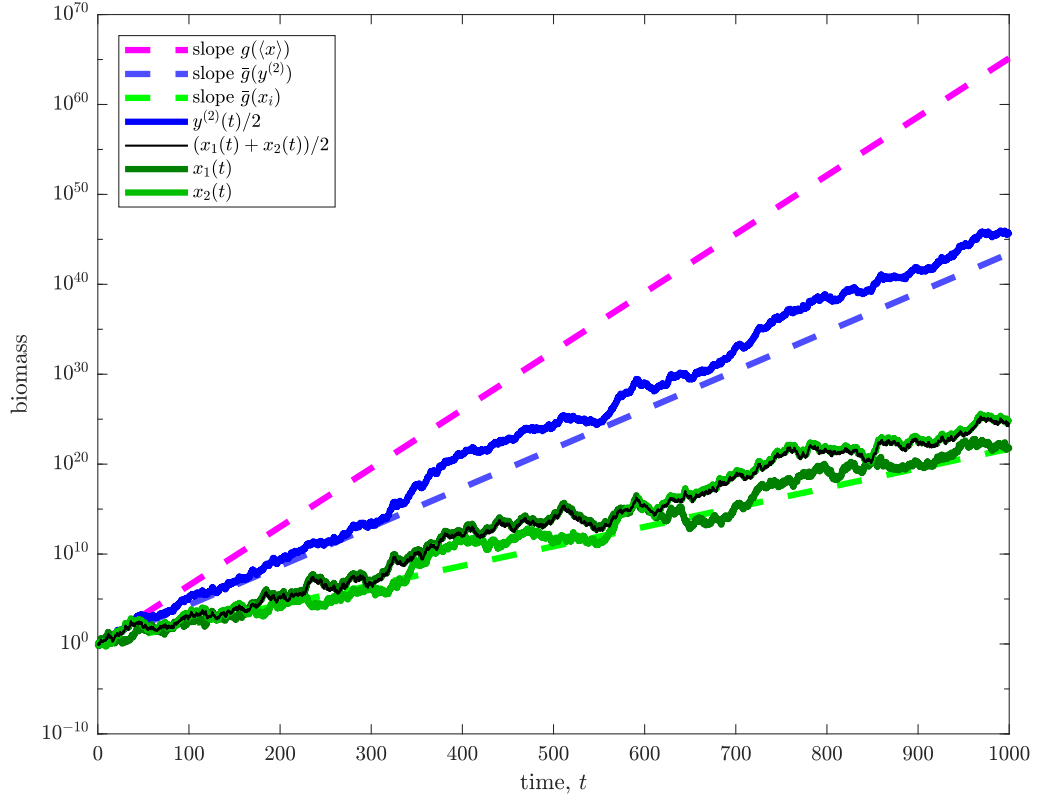Repeating the analysis of growth, pooling, and sharing yields a modified

9

Figure 2: Typical trajectories for two non-cooperating (green) entities and for the corresponding cooperating unit (blue) on a logarithmic vertical scale. Over time, the noise reduction for the cooperator leads to faster growth. Even without effects of specialization or the emergence of new function, cooperation pays in the long run. The thin black line shows the average of the non-cooperating entities, which is far inferior to the cooperating unit. In a very literal mathematical sense the whole, $y_1(t) + y_2(t)$, is more than the sum of its parts, $x_1(t) + x_2(t)$. The algebra of cooperation is not merely that of summation. The expectation-value growth rate (at which the pink dashed line grows) is the growth rate in the limit of infinitely many cooperators.

dynamic,

$$\Delta y^{(N)}(t) = y^{(N)}(t)\left(\frac{1}{N}\sum_{i=1}^{N}\mu_i\Delta t + \frac{1}{N}\sum_{i=1}^{N}\sigma_i\xi_i\sqrt{\Delta t}\right)$$
$$= y^{(N)}(t)\left(\mu^{(N)}\Delta t + \sigma^{(N)}\xi^{(N)}\sqrt{\Delta t}\right). \tag{20}$$

$\xi^{(N)}$ is a standard normal variate, as before, and

$$\mu^{(N)} \equiv \frac{1}{N}\sum_{i=1}^{N}\mu_i, \quad \sigma^{(N)} \equiv \frac{1}{N}\sqrt{\sum_{i=1}^{N}\sigma_i^2} \tag{21}$$

are effective drift and volatility parameters. Therefore, the biomasses of the cooperators evolve according to

$$dy^{(N)} = y^{(N)}\left(\mu^{(N)}dt + \sigma^{(N)}dW^{(N)}\right), \tag{22}$$

with time-average growth rate,

$$\bar{g}\left(y^{(N)}\right) = \mu^{(N)} - \frac{(\sigma^{(N)})^2}{2} = \frac{1}{N}\sum_{i=1}^{N}\left(\mu_i - \frac{\sigma_i^2}{2N}\right). \tag{23}$$

This is the sample mean of the time-average growth rates each entity would achieve by cooperating with like entities, *c.f.* (Eq. 16).

We can now answer the questions. It benefits entity j to join the cooperative if $\bar{g}\left(y^{(N)}\right) > \bar{g}(x_j)$, *i.e.* if

$$\frac{1}{N}\sum_{i=1}^{N}\left(\mu_i - \frac{\sigma_i^2}{2N}\right) > \mu_j - \frac{\sigma_j^2}{2}. \tag{24}$$

Similarly, the cooperative benefits by admitting entity j if

$$\frac{1}{N}\sum_{i=1}^{N}\left(\mu_i - \frac{\sigma_i^2}{2N}\right) > \frac{1}{N-1}\sum_{\substack{i=1 \\ i\neq j}}^{N}\left(\mu_i - \frac{\sigma_i^2}{2(N-1)}\right). \tag{25}$$

Consider two entities, where $\mu_1 > \mu_2$ and $\sigma_1 < \sigma_2$ so that x_1 grows faster over time than x_2. Rearranging (Eq. 24) for $N = 2$, we see that the fast grower should share with the slow grower if $\mu_1 - 3\sigma_1^2/4 < \mu_2 - \sigma_2^2/4$.

5.2 Correlated fluctuations

A second generalization concerns correlations. Fluctuations experienced by different entities are uncorrelated in our model: the dW_i in (Eq. 1) and, consequently, the ξ_i in (Eq. 7) onwards are independent random variables. In reality, cooperators are often spatially localized and experience similar environmental conditions. By allowing correlations between the ξ_i, our model can be adapted to describe such situations.

Suppose the $\xi_i \sim \mathcal{N}(0, 1)$ realized in a given time step are jointly normal and cross-correlated such that $\langle \xi_i \xi_j \rangle = \rho_{ij}$. Assume for simplicity that

$$\rho_{ij} = \begin{cases} 1, & i = j, \\ \rho, & i \neq j, \end{cases} \tag{26}$$

i.e. that the fluctuations for all pairs of different entities have the same covariance, ρ, where $|\rho| \leq 1$. The more general case of a covariance matrix with unequal off-diagonal elements is also tractable, but adds complexity without illumination.

The presence of cross-correlations alters the evaluation of the sum of the normal variates in (Eq. 12). We have now

$$\sum_{i=1}^{N} \xi_i \sim \mathcal{N}(0, N + N(N - 1)\rho). \tag{27}$$

Positive variance requires ρ to be confined to $-1/(N - 1) \leq \rho \leq 1$. Equation (27) suggests defining, analogously to (Eq. 13), a standard normal variate,

$$\xi^{(N)} \equiv \frac{1}{\sqrt{N + N(N - 1)\rho}} \sum_{i=1}^{N} \xi_i \sim \mathcal{N}(0, 1), \tag{28}$$

such that the change in $y^{(N)}$ can be written as

$$\Delta y^{(N)}(t) = y^{(N)}(t) \left(\mu \Delta t + \sigma \sqrt{\frac{1 + (N - 1)\rho}{N}} \xi^{(N)} \sqrt{\Delta t} \right), \tag{29}$$

analogous to (Eq. 14) in the uncorrelated case.

Without correlations, cooperation reduces the amplitude of fluctuations from σ to $\sigma/\sqrt{N}$. With them, the fluctuation amplitude becomes

$$\sigma_\rho \equiv \sigma \sqrt{\frac{1 + (N - 1)\rho}{N}}. \tag{30}$$

The variation of σ_ρ with ρ and N delineates the main features of this generalized model. Firstly, as a consistency check, we note that $\sigma_\rho \to \sigma/\sqrt{N}$ as $\rho \to 0$ for fixed N, recovering the uncorrelated result in the appropriate limit.

For all $N > 1$, we have $0 \leq \sigma_\rho \leq \sigma$, with $\sigma_\rho = \sigma$ if and only if $\rho = 1$. In other words, provided fluctuations are not perfectly correlated, a cooperation benefit always exists. This makes intuitive sense. With perfect correlation, all the ξ_i are identical and pooling and sharing achieves nothing. The cooperative is equivalent to a giant individual following a single trajectory of (Eq. 1). As soon as some variation is introduced between the fluctuations of the entities, the cooperation mechanism can begin to mitigate the negative effects of fluctuations on growth.

Furthermore, we have $\sigma_\rho \to \sigma\sqrt{\rho}$ as $N \to \infty$, with $0 \leq \rho \leq 1$ in this limit. The maximum time-average growth rate achievable by adding cooperators is, therefore,

$$\lim_{N \to \infty} \bar{g}\left(y^{(N)}\right) = \mu - \frac{\sigma^2\rho}{2}, \tag{31}$$

c.f. (Eq. 18). This cannot exceed μ (since ρ is non-negative) and decreases as ρ increases. Again, this is consistent with intuition: as fluctuations become more correlated, the variation between them diminishes and the scope for beneficial cooperation narrows. In our setup, cooperation relies on diversity in individual outcomes.

5.3 Partial cooperation

Another generalization of our model is to partial cooperation. Here entities pool and share only a fraction of their resources, resembling, for example, taxation and redistribution in human societies. We discuss this in a separate manuscript (Berman et al., 2017).

6 Discussion

Our model assumes nothing more than that evolutionary processes are multiplicative and noisy. In this context, the evolutionary advantage of cooperation arises from the nonlinear dependence of growth rates on temporal fluctuations. By reducing the amplitude of fluctuations, pooling and sharing increase the time-average growth rate of cooperating entities. This paints a picture of cooperation driven by self-interest, not altruism, with cooperators outgrowing similar non-cooperators.

Where our model describes well the growth of things in nature, it predicts that cooperation will be prevalent. As an attractor for many processes whose relative changes are independent random variables, we expect the model to resemble many real examples of self-reproduction.

In reality, many effects contribute to the formation of cooperative structure. Members of large cooperatives can coordinate their actions to develop emergent functions, such as the ability to swim towards nutrients. Cells in multicellular organisms differentiate and specialize in particular tasks. Human cooperation works analogously, with firms and individuals becoming proficient in different roles.

By finding a cooperative benefit under minimal assumptions, our analysis may shed light on cases of cooperation without such effects. In particular, it may explain the transition from single cells to bicellular organisms, too small and simple to benefit from new function or specialization.

Classical treatments of cooperation are predicated on the existence of a net fitness gain in a donor-recipient interaction. A universal mechanism by which this benefit arises would strengthen their theoretical foundations. In our model of noisy self-reproduction, we identify the increase in time-average growth rate achieved through cooperation as this net fitness gain. This is consistent with the concept of geometric mean fitness in the biological literature.

The impact of risk reduction on long-time growth suggests that risk management has a rarely recognized significance. Fluctuation reduction, or good risk management, does not merely reduce the likelihood of disaster or the size of up and down swings. It also improves the long-time performance of the structure whose risks are being managed. While the effect of reducing fluctuations depends on the specific setup, it is tantalizing to see that the simple and universal setting of multiplicative growth favors structure in the form of large cooperative units.

The insight that time averages may not be identical to expectation values was only reached in the development of statistical mechanics in the 19th century, where the physical nature of expectation values was identified as the mean over an ensemble of non-interacting systems. The development of ergodic theory in the 20th and 21st centuries provided the concepts that reveal the physically relevant aspect of stochastic processes like (Eq. 1). Today we have at our disposal mathematical tools that allow us to understand nonlinear dynamical effects, such as described here. Applying these tools suggests that our natural tendency to cooperate – manifested in our gut feeling and moral sentiment – is in harmony with a careful formal analysis of the issues involved.

Acknowledgements

References

Adamou, A. and O. Peters
   2016. Dynamics of inequality. *Significance*, 13(3):32–35.

Aitchison, J. and J. A. C. Brown
   1957. *The lognormal distribution*. Cambridge University Press.

Berman, Y., O. Peters, and A. Adamou
   2017. An empirical test of the ergodic hypothesis: Wealth distributions in the United States. Available at SSRN: https://ssrn.com/abstract=2794830.

Gillespie, J. H.
   1977. Natural selection for variances in offspring numbers: A new evolutionary principle. *Am. Nat.*, 111:1010–1014.

Lande, R.
   2007. Expected relative fitness and the adaptive topography of fluctuating selection. *Evolution*, 61:1835–1846.

Lewontin, R. C. and D. Cohen
   1969. On population growth in a randomly varying environment. *Proc. Nat. Ac. Sci.*, 62(4):1056–1060.

Meadows, D. H., D. L. Meadows, J. Randers, and W. W. Behrens III
   1972. *The limits to growth*. University Books.

Nowak, M. A.
   2006. Five rules for the evolution of cooperation. *Science (New York, N.y.)*, 314(5805):1560–1563.

Orr, H. A.
   2009. Fitness and its role in evolutionary genetics. *Nat. Rev. Genet.*, 10:531–539.

Peters, O. and M. Gell-Mann
   2016. Evaluating gambles using dynamics. *Chaos*, 26:23103.

Redner, S.
  1990. Random multiplicative processes: An elementary tutorial. *Am. J. Phys.*, 58(3):267–273.

Roper, M., M. J. Dayel, R. E. Pepper, and M. A. R. Koehl
  2013. Cooperatively generated stresslet flows supply fresh fluid to multi-cellular choanoflagellate collonies. *Phys. Rev. Lett.*, 110:228104.

Short, M. B., C. A. Solari, S. Ganguly, T. R. Powers, J. O. Kessler, and R. E. Goldstein
  2006. Flows driven by flagella of multicellular organisms enhance long-range molecular transport. *Proc. Natl. Acad. Sci. U.S.A.*, 103(22):8315–8319.

van Kampen, N. G.
  1992. *Stochastic Processes in Physics and Chemistry.* Elsevier (Amsterdam).