



**POLITECNICO
DI TORINO**

POLITECNICO DI TORINO

Master Degree course in Physics of Complex Systems

Master Degree Thesis

The Role of Information in the Evolution and Stability of Cooperation in Stochastic Environments

Supervisors

Prof. Luca DALL'ASTA

Prof. Jacopo GRILLI

Candidate

Ricardo GARCIA

ACADEMIC YEAR 2024-2025

Acknowledgements

I would like to express my deepest gratitude to Professor Jacopo Grilli, Dr. Onofrio Mazzarisi, and Dr. Lorenzo Fant for their invaluable guidance and support during my time at the ICTP in Trieste. Beyond their scientific mentorship, they became a true source of inspiration. Their passion and curiosity for research were fundamental in reviving my enthusiasm for science and in shaping the way I now approach it: with fun and wonder.

I am sincerely grateful to my family for their support and trust, and to my friends, both within and outside academia, for their presence throughout this journey.

Abstract

The aim of this thesis is to study the emergence and stability of cooperation in fluctuating environments, in presence of varying levels of available information about the current state of the environment. Cooperation's intrinsic susceptibility to individual deviant behaviours makes its stability a central aspect of this collective behaviour. We consider a model composed of two agents whose resources undergo stochastic multiplicative growth. These growth processes are influenced by the state of a shared fluctuating environment while being guided by the agents' individual strategies of investment in the environment's states. To improve their growth, the agents are allowed to interact through two mechanisms: partitioning the total available information about the fluctuating environment among them, and sharing fractions of their resources as a common good. We analyse the interplay between these mechanisms at varying levels of total information about the environment, thereby exploring the link between uncertainty and cooperation. We show that full cooperation, achieved through complete resource-sharing and equal information partition, is stable and evolvable, provided the initial state exhibits a sufficient level of cooperation. We also find that the size of the basin of attraction of the full cooperation state is determined by the predictability of the environment.

Contents

1	Introduction	5
2	Modeling framework	7
2.1	Model formalisation	7
2.2	The Geometric Brownian Motion	9
2.2.1	Non-ergodicity and growth	10
3	Total information constraint	11
3.1	Mutual information definition and evaluation	11
3.2	Information division and symmetrisation	14
4	Problem setting I: Information Partition	17
4.1	Optimal growth rate evaluation (I)	17
4.2	Growth rate interpretation and its extremal values	20
4.3	Diversification breakdown in the optimal strategy	22
5	Problem setting II: Information Partition and Wealth Sharing	25
5.1	Optimal growth rate evaluation (II)	25
5.1.1	Analytical expression for the growth rate	25
5.1.2	Numerical evaluation of the growth rate	26
5.1.3	Optimal investment strategy	28
5.2	Optimal collective behaviours	28
5.2.1	Equal sharing rates	28
5.2.2	Asymmetric sharing rates	30
6	Conclusion	35
7	Appendix	37
7.1	Continuous limit of the discrete-time GBM process	37
7.2	Itô integration	39
7.3	Information evaluation	41
7.3.1	Direct total information calculation	43
	List of Figures	45

Chapter 1

Introduction

The ability to adapt and thrive under different conditions is essential in biological systems as well as in numerous other situations. In the case of many entities co-existing and interacting, questions about the optimal behaviour to assume towards one another have been a topic of long-standing investigation. The two approaches of competition and cooperation [1] have been widely investigated.

Competition may appear as the most natural and advantageous response, since every individual undergoing evolution experiences competition and should be shaped to maximise its own evolutionary success, potentially at the expense of others. Cooperation, on the contrary, is inherently unstable, as it is susceptible to the tragedy of the commons and to cheating [2], as it typically involves immediate costs with uncertain or delayed benefits. Nevertheless, cooperative behaviours are observed across a wide range of contexts and organisational levels in nature. Therefore, a matter of interest in many different fields, ranging from biology to economics [3, 4], is not only the study of the conditions that are favourable to the existence of cooperation, but also the conditions that can allow its emergence and stability.

Stochasticity plays a crucial role in many systems where cooperation emerges [5]. Therefore, one of the possible approaches when aiming to describe the emergence of cooperation is to adopt a model able to capture the effects of environmental fluctuations. A canonical framework used to represent such contexts [6, 7] is the geometric Brownian motion [8], which describes the dynamics of the resources an agent possesses, when they undergo a stochastic multiplicative growth process. This is the case when the effect of stochasticity scales with the current value of the resources, thereby breaking ergodicity, rather than acting as an additive noise. The universality of the geometric Brownian motion allows the resource variable to be interpreted differently depending on the specific field, such as population abundance in biology or the capital accumulated by an investor in economics.

A straightforward mechanism for cooperation is the exchange of resources. In general, many cooperative structures can arise through this mechanism [9] when a net benefit arises from the cooperative interaction, and it is later redistributed, for example through reciprocity. However, as in [7], this thesis focuses on the case in which the effects typically guaranteeing the net benefit, complementarity of resources and resource thresholds for

target outcomes, are not present. Thus, our approach explores cooperation in the presence of a single qualitative kind of resource for all agents. Resources shared according to the independent choice of every agent are pooled and then equally redistributed. In economics, this may be translated as the collection and redistribution of taxes, whereas in biology it could involve the exchange of a specific nutrient through a common medium. Cooperation through resource pooling in stochastic multiplicative growth, used to reduce environmental uncertainty in an ensemble of agents subject to independent fluctuations, has been studied in [6,7], and shown to be a stable strategy. In these studies, heterogeneity between entities has been explored, being regarded as independent intrinsic differences in growth between agents.

An agent's ability to predict the fluctuations of the environment is fundamental in determining the stochastic dynamics of its resources [10,11]. In realistic settings, agents often have means to obtain some partial knowledge about the environment's state, quantified by the mutual information between the signals they individually receive and the actual state of the environment. Information in stochastic multiplicative growth systems has been studied from the perspective of cooperative behaviours. As with resources, a straightforward cooperative behaviour involving information would be the sharing of information between individuals. Information pooling, as a means to maximise collective growth, has been investigated as an alternative or complementary mechanism to resource pooling in [12]. In the latter work, information pooling and resource sharing are considered together and described as complementary strategies to enhance growth.

In this thesis, we introduce a fixed level of total information received by the totality of the agents to serve as a classification for systems, allowing us to compare diverse systems that share the same level of environmental predictability. This constraint can push agents to interact in an attempt to assess their optimal share of information. Therefore, we focus on the inverse approach to information pooling, the information partition mechanism, which is the process by which agents determine the subdivision of the total information (that subsequently remains private to each individual).

In this context, minimal models that incorporate both resource pooling and information can provide a good quantitative framework to study how cooperation emerges and persists under uncertainty. The model presented in this thesis follows this approach, bridging individual decision-making under limited information and resource sharing in fluctuating environments.

Chapter 2

Modeling framework

We consider a simple system of two agents whose resources undergo stochastic multiplicative growth. We introduce limited information regarding a fluctuating binary environment as a means to improve the agents' growth. We allow agents to interact by two mechanisms: (i) partitioning the available information between them, and (ii) sharing fractions of their resources as a common good. The interplay between these two mechanisms allows us to explore the relation linking uncertainty and cooperation in a paradigmatic model with applications in economics and population biology.

2.1 Model formalisation

The system we study consists of two primary components: the environment and the agents. Agents grow individual quantities, which we call resources or wealth, through a stochastic process. The environment assumes various states which affect the growth dynamics of these resources. The agents' goal is to maximise the growth of their own resources in the long term. Agents can do so by investing their resources: they choose how to partition and bet the totality of their wealth on the possible states the environment can assume. Agents are informed about the current state of the environment through noisy signals. Each agent receives its own private and independent signal that matches the state of the environment with a probability that, in principle, could be different for each private signal. In the present section, we formalise this system.

We assume the environment E can have two states

$$E = \{+1, -1\}, \quad (2.1)$$

with equal probability $P(E) = 1/2$. Agent i receives a signal

$$S_i = \begin{cases} E & \text{with prob. } P(E = S_i|S_i) = p_i \\ -E & \text{with prob. } P(E \neq S_i|S_i) = 1 - p_i \end{cases} \quad (2.2)$$

where p_i represents signal S_i 's reliability.

We will focus our study on the case of two agents and two states of the environment, and parameterise distributions as

$$P(E) = 1/2 \quad (2.3)$$

$$P(S_i) = 1/2 \quad (2.4)$$

$$P(E|S_i) = P(S_i|E) = p_i \delta_{E,S_i} + (1 - p_i) \delta_{E,-S_i} \quad (2.5)$$

where the first equality in Equation 2.5 follows from Bayes and the choice of equiprobable environment's states.

Agents use the information they get from their signals to invest all their resources in the states of the environment. At each time step, each agent independently invests a constant fraction

$$f_i^+ \in [0,1] \quad (2.6)$$

of their resources on the environment state $E = +1$, and the remaining part $f_i^- = 1 - f_i^+$ on the opposite state $E = -1$. Hence, they receive returns from their investments, depending on the actual state of the environment. We define the investments' returns as follows:

$$\begin{cases} b_{\uparrow} \cdot f_i^e x_i(t) & \text{if } e = E \\ b_{\downarrow} \cdot f_i^e x_i(t) & \text{if } e \neq E \end{cases} \quad (2.7)$$

where $x_i(t)$ represents the total amount of resources belonging to agent i at time t and the reward factors

$$b_{\uparrow}, b_{\downarrow} \geq 0, \quad (2.8)$$

are constants that apply uniformly to all agents, reflecting a property of the system (of the environment the individuals live in) rather than a specificity of single agents. The arrows notation suggests the context of a system in which a higher reward is given to an investment matching the environment, while the resources invested in the wrong one have a lower return, or even are completely lost if $b_{\downarrow} = 0$. In the following, we will therefore assume, without loss of generality,

$$b_{\uparrow} \geq b_{\downarrow} \geq 0. \quad (2.9)$$

We can now assume a discrete-time process and describe the resource update at each time step as a stochastic process. The relation with the continuous-time process can be found in section 7.1. We will also always assume delta correlation in time. The *investment strategy*, that is, the fractions of wealth an agent invests in every state of the environment, will depend on its signal. This means we can define signal-dependent fractions as

$$f_{S_i}^+ + f_{S_i}^- = 1, \quad (2.10)$$

where $f_{S_i}^+$ is the fraction invested in state $E = +1$ once received a signal S_i . In these terms, the resources $x_i(t)$ update reads, for a general signal S_i , as

$$x_i(t+1) = \begin{cases} f_{S_i}^+ b_{\uparrow} x_i(t) + f_{S_i}^- b_{\downarrow} x_i(t) & \text{with probability } P(E = +1|S_i) \\ f_{S_i}^+ b_{\downarrow} x_i(t) + f_{S_i}^- b_{\uparrow} x_i(t) & \text{with probability } P(E = -1|S_i) \end{cases} \quad (2.11)$$

We notice that the investment fractions feature some symmetries. Indeed, it is easy to see that, since the environment's states are equiprobable, the fraction of resources invested in the state of the environment matching (or mismatching) the signal should only depend on the signal reliability, and not on the specific value $+1$ or -1 , so that we can write

$$\begin{cases} f_{S_i}^+ + f_{S_i}^- = 1 \\ f_+^+ = f_-^- \\ f_-^+ = f_+^- \end{cases} \quad (2.12)$$

This reduces the degrees of freedom of the choice of each agent to only one variable, which we choose to be $f = f_+^+$. This will prove to be useful in the calculation of the relevant growth rate value.

2.2 The Geometric Brownian Motion

The Geometric Brownian Motion (GBM) constitutes a natural description of the resource growth process. The relevant characteristics of this process [8] are therefore briefly described here. A first idea of what a GBM process is can be grasped from the fact that it can be regarded in the logarithms' space as a simple Brownian Motion. The GBM is a multiplicative White Noise process described by the equation

$$dx = x\mu dt + x\sigma dW(t), \quad (2.13)$$

where $W(t)$ is a Wiener process and $dW(t)$ represents the noise term; x stands for the agent's wealth in our case. We also define the new variable

$$q = \log x. \quad (2.14)$$

This new variable q constitutes a better quantity to handle, and it is the variable we will mainly treat in the following. Performing the change of variable by resorting to Itô calculus 7.2 we can write the Brownian Motion equation Equation 7.38

$$\dot{q}(t) = \left(\mu - \frac{\sigma^2}{2}\right) + \sigma\epsilon(t), \quad (2.15)$$

where $\epsilon(t)$ is a Gaussian white noise, and

$$q(t) = g_0 t + \sigma \int_0^t dt' \epsilon(t'), \quad (2.16)$$

with the new drift term $g_0 := \mu - \frac{\sigma^2}{2}$. This allows us to calculate the solution for $x(t)$ (7.46), and its average:

$$\langle x(t) \rangle \approx x(0) \exp \left(\left(g_0 + \frac{\sigma^2}{2} \right) t \right) \quad (2.17)$$

$$\approx x(0) e^{\mu t}. \quad (2.18)$$

We conclude by noticing the fact that the average value grows much faster than the typical (median of the lognormal distribution) value

$$\langle x(t) \rangle \sim e^{\mu t} \quad (2.19)$$

$$e^{\langle q(t) \rangle} \sim e^{g_0 t} = e^{(\mu - \frac{\sigma^2}{2})t}. \quad (2.20)$$

2.2.1 Non-ergodicity and growth

The last equations (2.19) and (2.20) highlight a fundamental property of the GBM: the lack of ergodicity. The solution (7.46) shows how the process $x(t)$ is not stationary. On one hand, the time average can either be 0 or diverge, depending on the sign of the Itô exponent $\mu - \frac{\sigma^2}{2}$. On the other hand, the ensemble average always grows exponentially in time.

The time-average growth rate, which we called g_0 , is the result of the removal of stochasticity using time [13]

$$g_0 \doteq \lim_{t \rightarrow \infty} \frac{1}{t} \langle \log x(t) \rangle = \mu - \frac{\sigma^2}{2}, \quad (2.21)$$

and is able to capture the effect of fluctuations, contrary to the ensemble average growth rate

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \langle x(t) \rangle = \mu, \quad (2.22)$$

which is instead the result of the removal of stochasticity by means of an increasing ensemble size

In a multiplicative growth picture, the difference in the time and ensemble average growth rates can be better understood by noticing the parallelism of these quantities with, respectively, the geometric and arithmetic mean. From our perspective, the relevant quantity is the typical growth an individual agent can expect to experience in the long term, rather than the growth achieved in the large-population limit by the population average. Therefore, the time average g_0 is the most natural quantity that agents can look at to maximise their growth.

Resorting to our discrete-time model, we can define agents' objective function as follows

$$\langle g_0^{(i)} \rangle_{S_{i,E}} = \left\langle \log \frac{x_i(t+1)}{x_i(t)} \right\rangle_{S_{i,E}}. \quad (2.23)$$

Chapter 3

Total information constraint

Motivation

In the study of stochastic growth under environmental uncertainty, information can play a role comparable to that of resources under the circumstances that it is a limited asset that agents must exploit to improve their growth.

By constraining the total information available to the agents, we compare systems under the same "budget of knowledge" about the environment, and we can thus explore how different partitions of this budget affect the agents' behaviour. This perspective is particularly relevant in fluctuating environments, where asymmetry in information can alter the balance between cooperative and selfish strategies. The definition of an appropriate measure of information division is therefore essential to link the analysis of cooperation with the informational constraints naturally present in real systems.

3.1 Mutual information definition and evaluation

The agents receive information about the fluctuating environment through their private, independent, noisy signals. We set a constraint on the value of the total information that agents receive. We can define the total mutual information $\mathcal{I}$ between the state of the environment E and the signals $\vec{S}$ [12, 14]

$$\mathcal{I} \doteq I(E; \vec{S}) = \sum_{i=1}^N I(E; S_i) - R(E; \vec{S}), \quad (3.1)$$

where N is the total number of agents and R is the coefficient of redundancy. The latter measures the difference $I(S_1; \dots S_N) - I(S_1; \dots S_N|E)$. To have a better interpretation of this quantity, we can consider the relation between mutual information and entropy. Mutual information can be written as the difference in Shannon entropy

$$I(E; S_i) = H(E) - H(E|S_i) =: -\frac{\Delta H(E)}{\Delta S_i}. \quad (3.2)$$

This allows us to generalise the information on multiple signals as

$$I(E; \vec{S}) = \sum_{i=1}^N I(E; S_i) - \sum_{i>j=1}^N \frac{\Delta^2 H(E)}{\Delta S_i \Delta S_j} - \dots - \frac{\Delta^N H(E)}{\Delta S_1 \dots \Delta S_N}, \quad (3.3)$$

and using the identity $H(X, Y) = H(X|Y) + H(Y)$ we can write

$$\frac{\Delta^2 H(E)}{\Delta S_i \Delta S_j} = \frac{\Delta}{\Delta S_j} \left(H(E|S_i) - H(E) \right) \quad (3.4)$$

$$= I(S_i; S_j) - I(S_i; S_j|E) \quad (3.5)$$

$$= R(E; S_i; S_j). \quad (3.6)$$

Then, the coefficient of redundancy can be interpreted as the part of information that would appear in every term of the sum of single agent informations $I(S_i; E)$, and its subtraction corresponds to taking care of the overcounting of the shared (redundant) information. Hence, the total information constraint in our two-agent system reads

$$\mathcal{I} = I(E; S_1) + I(E; S_2) - I(S_1; S_2) + I(S_1; S_2|E), \quad (3.7)$$

which allows for the straightforward graphical representation shown in [Figure 3.1](#). Using

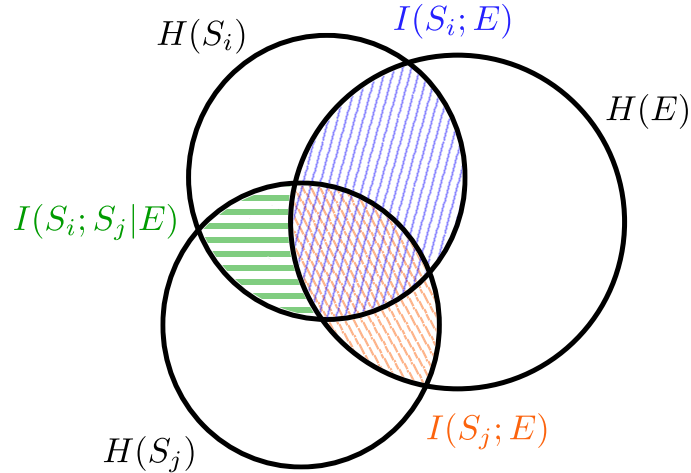


Figure 3.1. Entropy representation of information. Mutual information between two events corresponds to the intersection of their entropy shape. The information $I(S_i; S_j)$ corresponds to the union of the green and bi-coloured areas. Overcounting of the latter area explains the redundancy term.

the Kullback-Leibler definition of information, each term gives ([Appendix 7.3](#)):

$$I(S_i; E) = \log 2 + p_i \log p_i + (1 - p_i) \log(1 - p_i) \quad (3.8)$$

$$I(S_1; S_2) = \log 2 + \pi(p_1, p_2) \log(\pi(p_1, p_2)) + (1 - \pi(p_1, p_2)) \log(1 - \pi(p_1, p_2)) \quad (3.9)$$

$$I(S_1; S_2|E) = 0, \quad (3.10)$$

where $\pi(x, y) = x + y - 2xy$ is related to the mutual exclusive probability $\mathbb{P}(X \oplus Y)$. Putting all terms together, or equivalently, using the mutual information definition directly

$$I(E; \vec{S}) = \sum_{\{E\}, \{S_1\}, \{S_2\}} P(E, S_1, S_2) \log \frac{P(E, S_1, S_2)}{P(E) \cdot P(S_1, S_2)}, \quad (3.11)$$

the condition on total information reads (Appendix 7.3.1):

$$\begin{aligned} \mathcal{I} = I(E; \vec{S}) &= \log 2 + \\ &+ p_1 \log p_1 + (1 - p_1) \log(1 - p_1) + \\ &+ p_2 \log p_2 + (1 - p_2) \log(1 - p_2) + \\ &- \pi(p_1, p_2) \log \pi(p_1, p_2) - (1 - \pi(p_1, p_2)) \log(1 - \pi(p_1, p_2)) \\ &= \log 2 - h(p_1) - h(p_2) + h(\pi(p_1, p_2)), \end{aligned} \quad (3.12)$$

where we defined the entropy function $h(x) = -x \log x - (1 - x) \log(1 - x)$. The shape of the total information is shown in Figure 3.2.

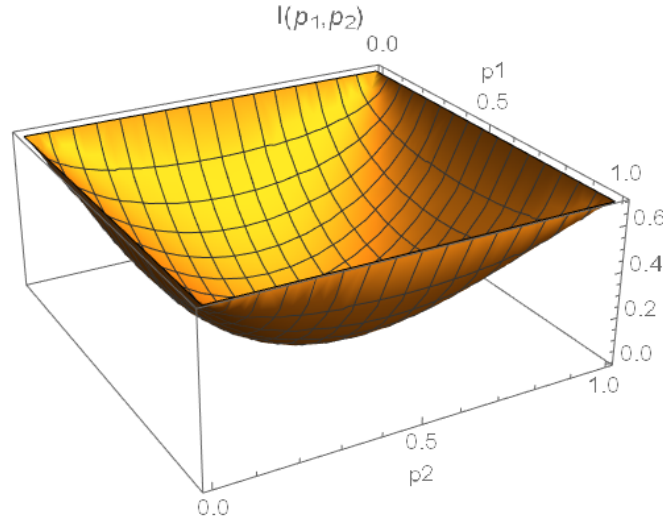


Figure 3.2. Shape of the information function on the $p_1 - p_2$ space.

The imposition of this constraint corresponds to identifying sets of couples (p_1, p_2) , i.e. sets of systems, that receive the same total information in varying partitions between the two agents. This means we are classifying systems according to different contour lines at values $\mathcal{I}$ on the total information. The shape of some of these lines of constant total information, restricted to the first quarter of the $(p_1 - p_2)$ -plane thanks to the symmetry around the $p_i = 1/2$ value, is shown in Figure 3.3. Additionally, symmetry between the two agents upon exchange $p_1 \leftrightarrow p_2$ allows us to further restrict to half the first quarter. We observe that, while for lower $\mathcal{I}$ values the contour lines can be

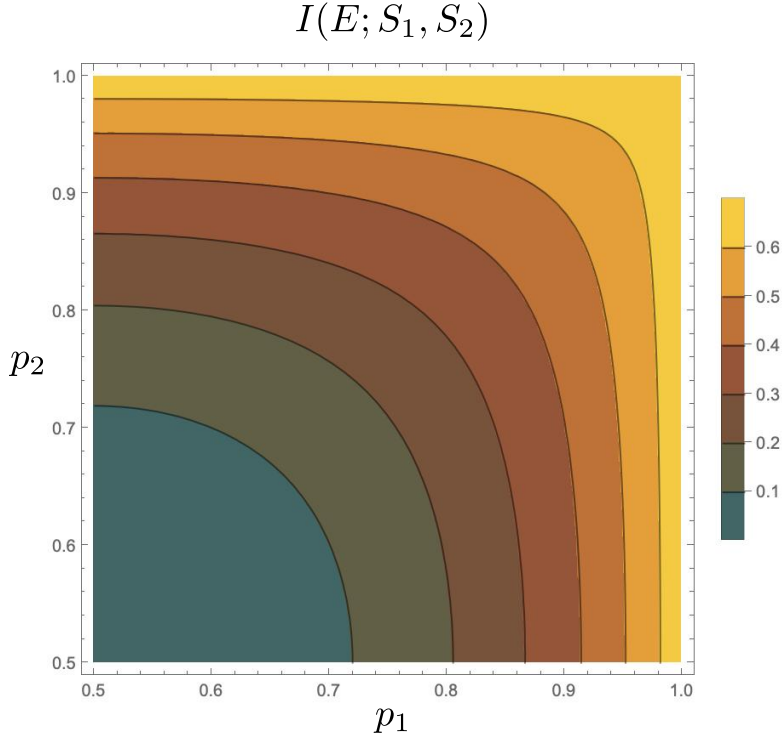


Figure 3.3. Contour plot of the total information function on the first quarter of the (p_1, p_2) -plane centred in $(0.5, 0.5)$.

approximated by circumferences, for higher constant values of total information, the contour lines progressively transition to the rectangular shape of the maximum $\mathcal{I} = \log 2$. For $\mathcal{I}$ tending to $\log 2$, the two signal reliabilities p_1 and p_2 approach the limit curve corresponding to perfect signals $p_1 = p_2 = 1$. In other words, the values $p_{\text{mid}}(\mathcal{I})$ of symmetry between the agents' signals, lying on the intersection between the contour line of value $\mathcal{I}$ and the diagonal $p_1 = p_2$, approach the maximum value p_{max} for increasing $\mathcal{I}$.

3.2 Information division and symmetrisation

The total information constraint bounds the signal reliability values as $(p_1, p_2(p_1, \mathcal{I}))$ and allows us to choose $p_1 \in [0, p_{\text{max}}(\mathcal{I})]$, restricted to $p_1 \in [1/2, p_{\text{max}}(\mathcal{I})]$, as a natural variable along which to evaluate the growth rate values at fixed $\mathcal{I}$. However, we notice that, due to the contour level shapes, the position of $p_{\text{mid}}(\mathcal{I}) := p_1 = p_2$ is shifted towards p_{max} . The choice of p_1 is, therefore, natural, but cannot display the symmetry around the point of equal information partition, which is required to disclose some of the features of this model. We define here, as an alternative to p_1 , a new measure that exhibits this symmetry. For this purpose, we resort to the entropy representation of information.

Exploiting the previous calculation of the terms composing the total information

$I(E; S_1, S_2)$, we can improve the representation given in Figure 3.1. We notice that, due

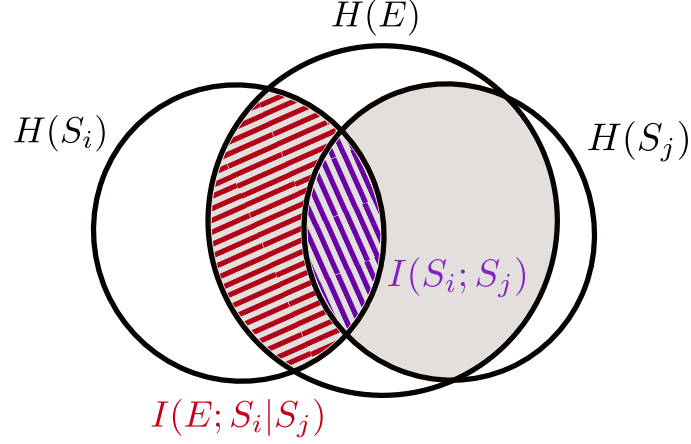


Figure 3.4. Entropy representation of information, in the case $P(S_1, S_2|E) = P(S_1|E)P(S_2|E)$. The area shaded in grey corresponds to $\mathcal{I} = I(E; S_1, S_2)$, while the union of the hatched areas corresponds to $I(S_1; E)$.

to $P(S_1, S_2|E) = P(S_1|E)P(S_2|E)$, the term $I(S_1, S_2|E) = 0$, and therefore the diagrams corresponding to $H(S_1)$ and $H(S_2)$ can have intersections only within $H(E)$, as shown in Figure 3.4. Hence, we can write

$$I(S_1; E) = I(S_1; E|S_2) + I(S_1; S_2) \quad (3.13)$$

$$I(S_2; E) = I(S_2; E|S_1) + I(S_1; S_2). \quad (3.14)$$

This means that the surface covered by $H(S_1)$ and $H(S_2)$ within $H(E)$, the pictorial element corresponding to $\mathcal{I}$, must remain constant along a contour line, while the different values of p_1 and p_2 , determining $I(S_1; E)$ and $I(S_2; E)$, increase with the area of the intersection between their respective signal entropy and $H(E)$.

Our objective here is to find a measure that could represent the partition of total information between the two agents, translating our representation in terms of signal reliabilities into another representation closer to information quantities. We observe that the single-agent information of agent 1, $I(S_1; E)$, is composed by a *unique* and a *redundant* part, respectively corresponding to $I(S_1; E|S_2)$ and $I(S_1; S_2)$, while for the *synergistic* part $I(S_1, S_2|E) = 0$ [15]. The *unique* part of information, which is not shared with the partner, is what actually determines the difference between the two agents, while the shared part plays the same role for both agents. The new measure must therefore be of the form

$$I_1(p_1) = I(S_1; E|S_2) + f_1(I(S_1; S_2)) \quad (3.15)$$

$$I_2(p_1) = I(S_2; E|S_1) + f_2(I(S_1; S_2)), \quad (3.16)$$

where f_1, f_2 are some functions of the redundant part of information. In order to identify suitable functions f_1 and f_2 , we require the new measure to satisfy the properties:

- The sum of the measures of the two agents is the total information

$$I_1(p_1) + I_2(p_1) = \mathcal{I} \quad \forall p_1. \quad (3.17)$$

- The measures of the two agents are equal when the information partition is fair

$$I_1(p_{\text{mid}}) = I_2(p_{\text{mid}}). \quad (3.18)$$

- No signal is preferred

$$I_1(p_2) = I_2(p_1). \quad (3.19)$$

- The measure is non-negative

$$I_1(p_1), I_2(p_1) \geq 0 \quad \forall p_1. \quad (3.20)$$

- The measure is monotonic in p_1 .

A simple solution consists in equally dividing the redundant information between the agents:

$$I_1(p_1) = I(S_1; E|S_2) + \frac{1}{2}I(S_1; S_2) = I(S_1; E) - \frac{1}{2}I(S_1; S_2) \quad (3.21)$$

$$I_2(p_1) = I(S_2; E|S_1) + \frac{1}{2}I(S_1; S_2) = I(S_2; E) - \frac{1}{2}I(S_1; S_2). \quad (3.22)$$

This choice satisfies all the previous properties, as shown in [Figure 3.5](#), and can be written explicitly as

$$I_1(p_1) = h(p_1) - \frac{1}{2}h\left(\pi(p_1, p_2(p_1))\right) - \frac{1}{2}\log 2. \quad (3.23)$$

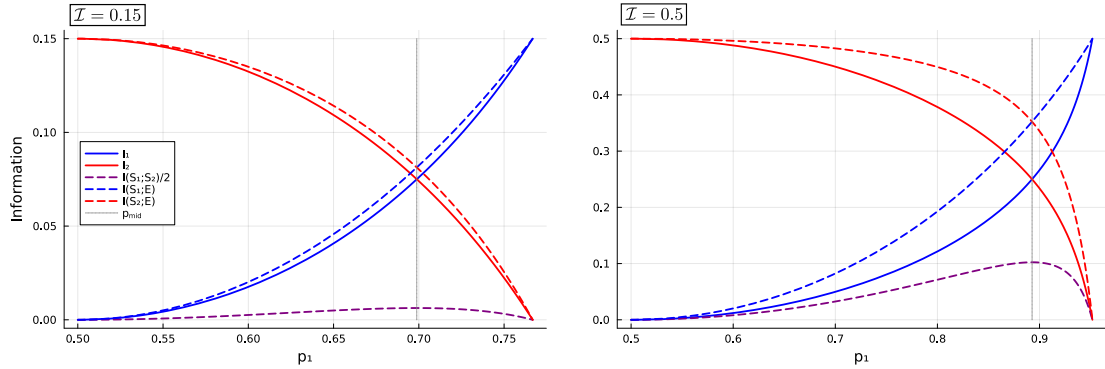


Figure 3.5. Single agent information measure for two values of total information $\mathcal{I}$. Both agents' measures and their components are shown.

Chapter 4

Problem setting I: Information Partition

Motivation

Before introducing interaction through the mechanism of wealth sharing, we discuss the fundamental properties of systems composed of agents that receive a partitioned information $\mathcal{I}$. Our focus is on the long-term growth rate, which is the quantity that agents aim at maximising. We therefore analyse its dependence on varying partitions of the total information.

4.1 Optimal growth rate evaluation (I)

As discussed in [section 2.1](#), the stochastic multiplicative growth process consists of agents gambling on the state of the environment. We can define, for every agent, returns from investments in discrete time. The investment amounts consist of continuous fractions of the resources $x_i(t)$ owned by agent i at time t . We assume that every agent always invests the totality of its resources in the two possible states of the environment, so that at the next time step $t + 1$, the wealth becomes:

$$x_i(t + 1) = \begin{cases} f_{S_i}^+ b_{\uparrow} x_i(t) + f_{S_i}^- b_{\downarrow} x_i(t) & \text{with probability } P(E = +1|S_i) \\ f_{S_i}^+ b_{\downarrow} x_i(t) + f_{S_i}^- b_{\uparrow} x_i(t) & \text{with probability } P(E = -1|S_i) \end{cases} . \quad (4.1)$$

The long-term growth rate for agent i , defined in [Equation 2.21](#), reads

$$g_0^{(i)} = \log \frac{x_i(t + 1)}{x_i(t)} = \begin{cases} f_{S_i}^+ b_{\uparrow} + (1 - f_{S_i}^+) b_{\downarrow} & \text{with probability } P(E = +1|S_i) \\ f_{S_i}^+ b_{\downarrow} + (1 - f_{S_i}^+) b_{\uparrow} & \text{with probability } P(E = -1|S_i) \end{cases} , \quad (4.2)$$

and making use of the symmetries of the investment fractions ([2.12](#)), we can express the growth rate with only one degree of freedom, which we choose to be $f_+^{+(i)}$.

With these simplifications, and omitting the agent index in $f_{S_i}^{\pm(i)} = f_{S_i}^{\pm}$ for clarity, the

average of the growth rate becomes

$$\langle g_0^{(i)} \rangle = \left\langle \log \frac{x_i(t+1)}{x_i(t)} \right\rangle_{S_i, E} \quad (4.3)$$

$$= \sum_{\{S_i\}, \{E\}} P(S_i, E) \log \frac{x_i(t+1)}{x_i(t)} \quad (4.4)$$

$$= \sum_{\{S_i\}} P(S_i) \left[P(E = +1|S_i) \log \left(f_{S_i}^+ b_\uparrow + (1 - f_{S_i}^+) b_\downarrow \right) + \right. \quad (4.5)$$

$$\left. + P(E = -1|S_i) \log \left((1 - f_{S_i}^+) b_\uparrow + f_{S_i}^+ b_\downarrow \right) \right]$$

$$= \frac{1}{2} \left[P(E = +1|S_i = +1) \log \left(f_+^+ b_\uparrow + (1 - f_+^+) b_\downarrow \right) + \right. \quad (4.6)$$

$$+ P(E = -1|S_i = +1) \log \left((1 - f_+^+) b_\uparrow + f_+^+ b_\downarrow \right) +$$

$$+ P(E = +1|S_i = -1) \log \left((1 - f_+^+) b_\uparrow + f_+^+ b_\downarrow \right) +$$

$$\left. + P(E = -1|S_i = -1) \log \left(f_+^+ b_\uparrow + (1 - f_+^+) b_\downarrow \right) \right]$$

$$= P(E = +1|S_i = +1) \log \left(f_+^+ b_\uparrow + (1 - f_+^+) b_\downarrow \right) + \quad (4.7)$$

$$+ P(E = -1|S_i = -1) \log \left((1 - f_+^+) b_\uparrow + f_+^+ b_\downarrow \right)$$

$$= p_i \log \left[f_+^+ b_\uparrow + (1 - f_+^+) b_\downarrow \right] + (1 - p_i) \log \left[(1 - f_+^+) b_\uparrow + f_+^+ b_\downarrow \right]. \quad (4.8)$$

In accordance with the Kelly criterion [16], to maximise the long-term growth rate, here we derive with respect to the only degree of freedom $f_i := f_+^{+(i)}$, and find the stationary point:

$$0 = \partial_{f_i} \langle g_0^{(i)} \rangle \quad (4.9)$$

$$= p_i \frac{b_\uparrow - b_\downarrow}{f_i b_\uparrow + (1 - f_i) b_\downarrow} + (1 - p_i) \frac{b_\downarrow - b_\uparrow}{(1 - f_i) b_\uparrow + f_i b_\downarrow}, \quad (4.10)$$

leading to

$$p_i \frac{b_\uparrow - b_\downarrow}{f_i b_\uparrow + (1 - f_i) b_\downarrow} = (1 - p_i) \frac{b_\uparrow - b_\downarrow}{(1 - f_i) b_\uparrow + f_i b_\downarrow} \quad (4.11)$$

$$p_i (b_\uparrow + f_i (b_\downarrow - b_\uparrow)) = (1 - p_i) (b_\downarrow + f_i (b_\uparrow - b_\downarrow)) \quad (4.12)$$

$$f_i (b_\downarrow - b_\uparrow) (p_i + 1 - p_i) = (1 - p_i) b_\downarrow - p_i b_\uparrow, \quad (4.13)$$

and finally

$$f_i^* = \frac{b_\uparrow}{b_\uparrow - b_\downarrow} p_i - \frac{b_\downarrow}{b_\uparrow - b_\downarrow} (1 - p_i). \quad (4.14)$$

The fraction of resources invested in a state can only take values in the range $[0, 1]$, hence we need to bound the optimal investment fraction $f_i^* = \arg \max_{f_i} \langle g_0^{(i)} \rangle$ as a piecewise function:

$$\hat{f}_i(p_i) = \begin{cases} 1 & \text{if } f_i^* > 1 \iff p_i > \frac{b_\uparrow}{b_\uparrow + b_\downarrow} \\ 0 & \text{if } f_i^* < 0 \iff p_i < \frac{b_\downarrow}{b_\uparrow + b_\downarrow} \\ \frac{b_\uparrow}{b_\uparrow - b_\downarrow} p_i - \frac{b_\downarrow}{b_\uparrow - b_\downarrow} (1 - p_i) & \text{if } f_i^* \in (0, 1) \end{cases}, \quad (4.15)$$

An example of the analytical behaviour of $\hat{f}_i$ for the case $b_\uparrow = 10$, $b_\downarrow \in [0, 1]$ shown in [Figure 4.1](#).

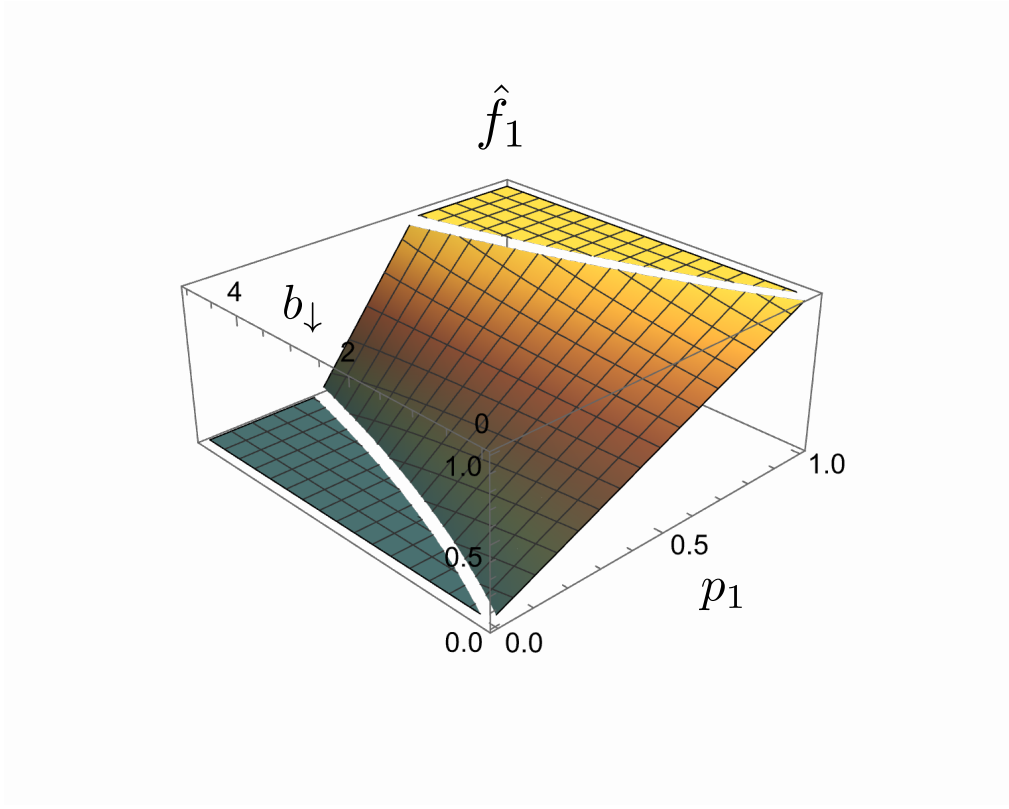


Figure 4.1. Kelly optimal investment as function of p_1 and $b_\downarrow$, given $b_\uparrow = 10$. For $b_\downarrow = 0$, $\hat{f}_1(p_1) = p_1$. For non-zero values of $b_\downarrow$, the linear function acquires a shift and an angular coefficient determined by the combination of the two reward factors.

By substituting the optimal unbounded strategy f_i^* into $g_0^{(i)}$, and consequently expressing it as a function of p_i

$$\begin{aligned}
 \langle g_0^{(i)} \rangle^* &= p_i \log [f_i^* b_\uparrow + (1 - f_i^*) b_\downarrow] + (1 - p_i) \log [(1 - f_i^*) b_\uparrow + f_i^* b_\downarrow] \\
 &= p_i \log [f_i^* (b_\uparrow - b_\downarrow) + b_\downarrow] + (1 - p_i) \log [-f_i^* (b_\uparrow - b_\downarrow) + b_\uparrow] \\
 &= p_i \log [b_\uparrow p_i - b_\downarrow (1 - p_i) + b_\downarrow] + (1 - p_i) \log [-b_\uparrow p_i + b_\downarrow (1 - p_i) + b_\uparrow] \\
 &= p_i \log [p_i (b_\uparrow + b_\downarrow)] + (1 - p_i) \log [(1 - p_i) (b_\uparrow + b_\downarrow)] \\
 &= \log(b_\uparrow + b_\downarrow) - h(p_i),
 \end{aligned} \tag{4.16}$$

so that we can now define the piecewise function $\hat{g}_0^{(i)}(p_i) := \langle g_0^{(i)}(\hat{f}_i(p_i), p_i) \rangle$, reading:

$$\hat{g}_0^{(i)} = \begin{cases} p_i \log b_\uparrow + (1 - p_i) \log b_\downarrow & \text{if } f^* = 1 \iff p_i > \frac{b_\uparrow}{b_\uparrow + b_\downarrow} \\ p_i \log b_\downarrow + (1 - p_i) \log b_\uparrow & \text{if } f^* = 0 \iff p_i < \frac{b_\downarrow}{b_\uparrow + b_\downarrow} \\ \log(b_\uparrow + b_\downarrow) - h(p_i) & \text{if } f^* \in (0, 1) \end{cases} \tag{4.18a}$$

$$\tag{4.18b}$$

$$\tag{4.18c}$$

where the cases (4.18a) and (4.18b) straightforwardly follow from Equation 4.8 by substituting, respectively, $f_+^+ = 1$ and $f_+^+ = 0$.

4.2 Growth rate interpretation and its extremal values

We begin by focusing on the analytical expression of the average long-term growth rate in the case where the investment fractions remain within their physical limits (4.18c), i.e. without saturating to their extremal values, as illustrated in Figure 4.2. We observe that its expression is composed of two terms, allowing for the following interpretation:

- An *environment generosity* term $\log(b_\uparrow + b_\downarrow)$, independent of the signal reliability p_i . Depending only on the sum of the two reward factors, this term accounts for the stress of the environment, regardless of the agents' ability to predict its state. It is therefore a positive shift of the entropic term.
- An *uncertainty* term, minus the entropy $h(p_i)$, dependent on the signal reliability p_1 . Recalling Equation 3.8, $h(p_i)$ can be rewritten as $\log 2 - I(S_i; E)$, that is, the information agent i would need to have perfect knowledge about the two-state environment.

Then, the average long-term growth rate for diversified investments can be interpreted as the baseline contribution of the environment generosity, corrected by the penalty due to uncertainty. Equivalently, it is determined by how far the agent remains from the ideal reference of perfect knowledge of the environment.

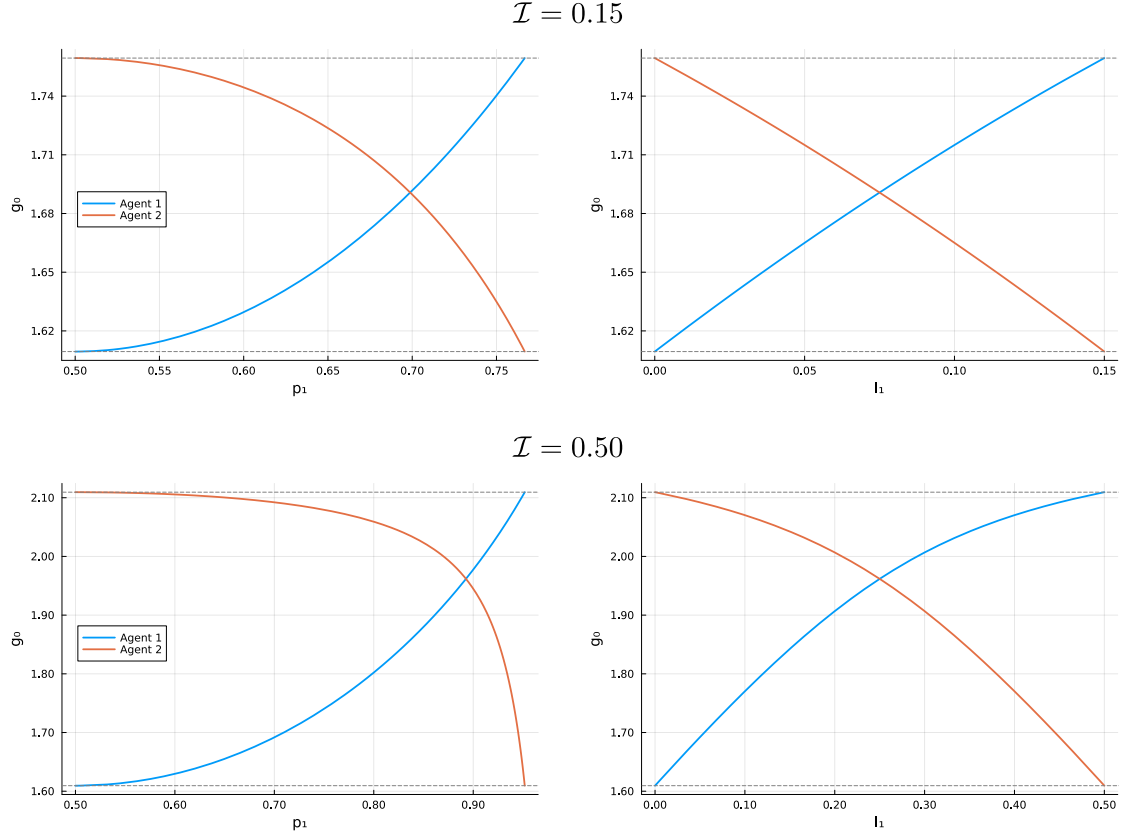


Figure 4.2. Average long-term growth rate in the case $f^* \in (0,1)$, shown for different values of the total information constraint, as a function of the two variables p_1 and I_1 , for parameters $b_\uparrow = 10$, $b_\downarrow = 0$. The minimum and maximum theoretical values, respectively $\log(b_\uparrow + b_\downarrow) - \log 2$ and $\log(b_\uparrow + b_\downarrow) - (\log 2 - \mathcal{I})$, are shown in grey. The I_1 variable representation emphasises symmetry between the two agents' growth rates around the equal information partition point, where they coincide. Higher information leads to faster growth, and brings the equal growth rate value closer to the upper bound, reflecting a higher redundancy $I(S_1; S_2)$ in information.

Thanks to the direct relation of the average long-term growth rate with information, by means of the previous reformulation of the entropic term, we can explore the implications of the maximum total information constraint on agents' growth.

We can rewrite Equation 4.18c in terms of the information quantity $I(S_i; E)$ as

$$\hat{g}_0^{(i)} \Big|_{f^* \in (0,1)} = \log(b_\uparrow + b_\downarrow) - [\log 2 - I(S_i; E)]. \quad (4.19)$$

Now, recalling the total information decomposition in Equation 3.7, and considering the constraint $\mathcal{I} = I(S_1, S_2; E)$, follows that the single agent information must be bounded as

$$0 \leq I(S_i; E) \leq \mathcal{I} \leq \log 2, \quad (4.20)$$

hence, the analytical average long-term growth rate

$$\log(b_{\uparrow} + b_{\downarrow}) - \log 2 \leq \hat{g}_0^{(i)} \Big|_{f^* \in (0,1)} \leq \log(b_{\uparrow} + b_{\downarrow}) - (\log 2 - \mathcal{I}) \leq \log(b_{\uparrow} + b_{\downarrow}). \quad (4.21)$$

Therefore, we conclude that the constraint on total information sets the limit on the maximum growth rate.

4.3 Diversification breakdown in the optimal strategy

In the previous [section 4.2](#), we characterised the growth rate in the case where agents optimally diversify without saturating their allocation. However, depending on the system parameters, the optimal investment strategy can drive the allocation fractions to their boundaries, corresponding to bets on a single state of the environment. In this section, we take a step back and turn our attention to the regimes $f^* = 1$ and $f^* = 0$, where diversification breaks down.

We define a system as saturating when, for some partition of the available information, the optimal investment fractions reach their extremal values. Therefore, we identify the distinctive characteristic of these systems as the combination of the values of the set of parameters $(\mathcal{I}, b_{\uparrow}, b_{\downarrow})$.

To illustrate the role these parameters play in deciding whether investment fraction saturation occurs, we can show in the $p_1 - p_2$ space the total information contour line at information $\mathcal{I}$ together with the function $\hat{f}_1$, as in [Figure 4.3](#). The shape of $\hat{f}_1$ as a func-

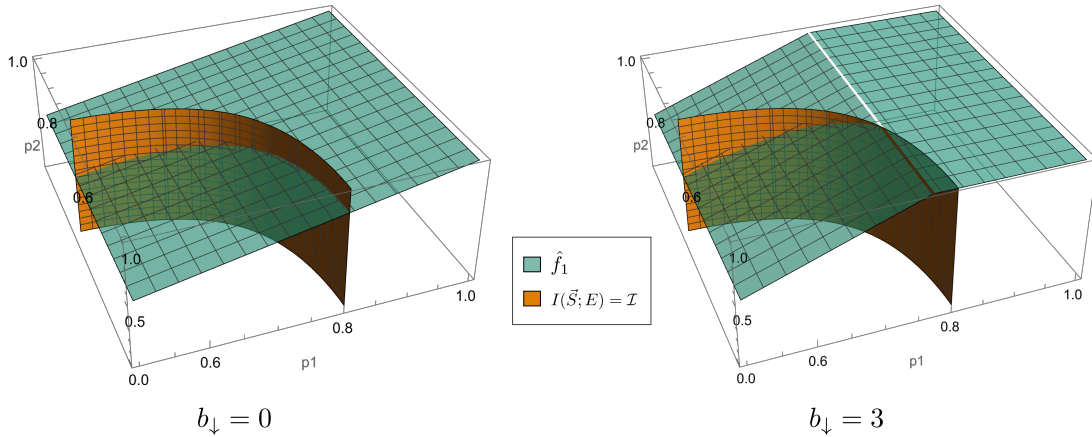


Figure 4.3. Investment fraction $\hat{f}_1$ for two values of the parameter $b_{\downarrow}$ and $\mathcal{I} = 0.2$, $b_{\uparrow} = 10$. The contour line of constant information $\mathcal{I}$ is shown to highlight the intersection between the two curves.

tion of p_1 at fixed $\mathcal{I}$ is given by the projection of the intersection between the contour line and the surface, which results in a piecewise linear profile. By symmetry, the surface

for $\hat{f}_2$ in the p_1-p_2 plane coincides with that of $\hat{f}_1$, but oriented to be increasing along the p_2 direction. Hence, the projection of the intersection as a function of p_1 results in a piecewise curve. The investment fractions for both agents are shown together as functions of p_1 in Figure 4.4, for the same set of parameters as in Figure 4.3.

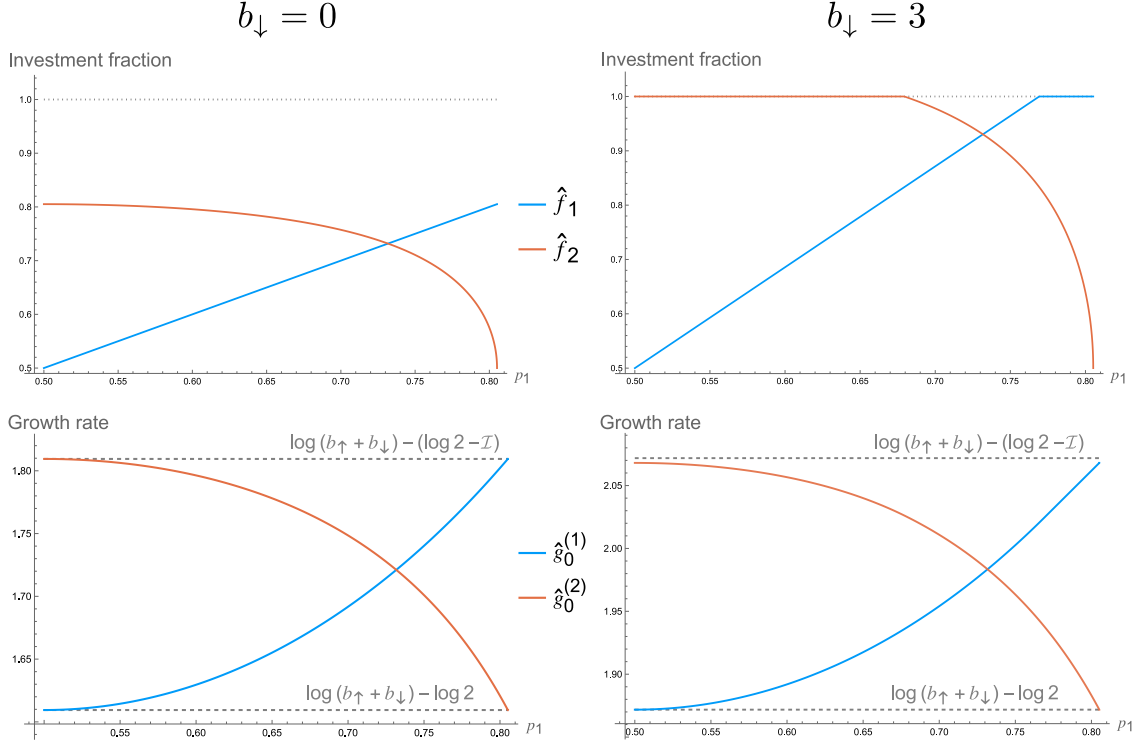


Figure 4.4. Investment fractions and corresponding growth rates, for both agents, with parameters $b_{\uparrow} = 10$, $\mathcal{I} = 0.2$.

The onset of saturation is determined by the intersection of the contour line with the $\hat{f}_1 = 1$ or $\hat{f}_1 = 0$ regions. This means that, restricting the discussion to $p_1, p_2 > 0.5$ by symmetry, and recalling the assumption $b_{\uparrow} > b_{\downarrow}$, we can analyse the role of each parameter as follows:

- b_↓)** For $b_{\downarrow} = 0$, saturation is inhibited. For $b_{\downarrow} > 0$, $\hat{f}_1$ can saturate for sufficiently high p_1 .
- b_↑)** Once $b_{\downarrow} > 0$, the difference between the two reward factors $b_{\uparrow} - b_{\downarrow}$ determines the size, in terms of p_1 values, of the saturating region. For lower differences in reward factors, gambling on the two states of the environment has a more similar outcome; investing in the mismatched state becomes less "dangerous", or less disadvantageous, and this allows for more extreme strategies, giving up on the safety ensured by diversification.

The condition $b_{\downarrow} = 0$, $b_{\uparrow} > 0$ makes bets on the matching state *infinitely* more convenient, requiring conservative investments.

$\mathcal{I}$) Higher values of total information select a contour line closer to the saturation regions: highly reliable signals allow agents to invest with confidence.

In Figure 4.4 we show, for both agents, the investment fractions corresponding to the cases illustrated in Figure 4.3, and the resulting average long-term growth rates. In this figure, one can notice that $\hat{g}_0^{(1)}$ and $\hat{g}_0^{(2)}$ do not reach the maximum theoretical value $\log(b_\uparrow + b_\downarrow) - (\log 2 - \mathcal{I})$. This is due to the saturating investment fractions: to reach the latter value, agents would be required to invest more than the totality of their resources, exceeding the upper limit of their investment fractions. Indeed, the distance from the analytical maximum disappears for the unbounded growth rate $\langle g_0^{(i)} \rangle^*$ of Equation 4.17, while it increases for fractions saturating in greater regions of the $p_1 - p_2$ plane, as shown in Figure 4.5.

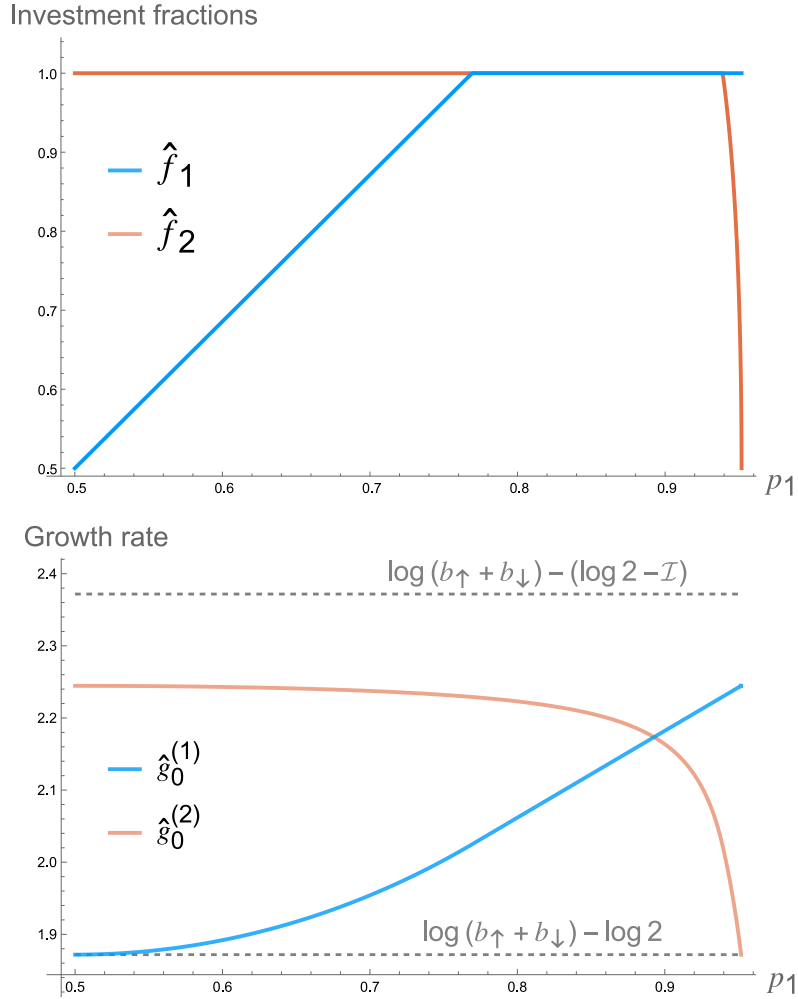


Figure 4.5. Investment fractions and growth rates for $b_\uparrow = 10$, $b_\downarrow = 3$, $\mathcal{I} = 0.5$.

Chapter 5

Problem setting II: Information Partition and Wealth Sharing

Motivation

In [chapter 4](#), the main properties of systems of agents partitioning a fixed total information have been discussed. Now, we introduce the second mechanism of interaction: wealth sharing.

Sharing resources is a behaviour that emerges in both biological [\[17, 18\]](#) and socio-economic systems [\[19\]](#), where individuals may pool part of their resources for collective benefit. In our framework, this mechanism allows agents to cooperate by contributing a fraction of their wealth to a common pool, which is then redistributed equally. This addition allows agents to altruistically and asymmetrically cooperate with the aim of increasing their own average long-term growth rate. Reciprocity is therefore required for the agents to balance the cost of sharing and improve their growth. In the present chapter, we study this mechanism of interaction at interplay with the previously analysed information partition, expanding on the setting described in [\[5–7\]](#)

5.1 Optimal growth rate evaluation (II)

5.1.1 Analytical expression for the growth rate

This [subsection 5.1.1](#) mainly follows the contents from [\[6\]](#).

In this new wealth-sharing setting, the interaction mechanism can be formalised in discrete time. We define the value a_i as the fraction of resources $x_i(t)$ that is shared by agent i at time t . An important remark is that while the amount of resources evolves in time, the fraction a_i is a constant characterising agent i 's strategy in the long-term growth. The totality of the pooled resources is then redistributed equally between all the agents, which in our case are two, independently of the fraction of resources they shared. The discrete-time dynamics can therefore be considered as a two-step process (with the order of the steps being irrelevant in the continuous time limit). The first step consists

in the sharing process

$$x_1(t + 1/2) = x_1(t) + \frac{-a_1x_1(t) + a_2x_2(t)}{2}, \quad (5.1)$$

and the second step consists in the individual growth of resources

$$x_1(t + 1) = x_1(t + 1/2)\zeta_1(t), \quad (5.2)$$

where the stochastic multiplicative noise $\zeta_1(t)$ corresponds to the discrete-time resources update described in [Equation 2.11](#).

In continuous time, agents' dynamics $\dot{x}_i(t) = \mu_i x_i + \sigma_i x_i \xi_i(t)$, where $\xi_i(t)$ is a delta-correlated white noise, become coupled:

$$\dot{x}_i(t) = \mu_i x_i + \sigma_i x_i \xi_i(t) + \frac{1}{2}[\alpha_j x_j(t) - \alpha_i x_i(t)], \quad (5.3)$$

where $j \neq i$ and $\alpha_i = a_i/\Delta t$ in the continuous time limit $\Delta t \rightarrow 0$ is the sharing rate. The $q_i(t) = \log(x_i(t))$ is again convenient to analytically describe the average long-term growth rate

$$g_i = \lim_{t \rightarrow \infty} \frac{1}{t} \langle q_i(t) \rangle, \quad (5.4)$$

and using Itô calculus it's possible to show that, for agent 1

$$\langle \dot{q}_1 \rangle = g_{0|0}^{(1)} - \frac{\alpha_1}{2} + \frac{\alpha_2}{2} \langle e^{q_2 - q_1} \rangle(t), \quad (5.5)$$

where, with a slight change of notation, $g_{0|0}^{(i)}$ is the uncoupled average long-term growth rate for some choice of investment strategy. The dynamics of $\langle e^{q_2 - q_1} \rangle(t)$ can be analytically derived, and it is ergodic with a stationary distribution $\langle e^{q_2 - q_1} \rangle_{\text{eq}}$. Then, it is possible to find an analytical expression for $g_{\alpha_1|\alpha_2}^{(i)}$, the growth rate of agent i in the presence of wealth sharing with sharing rates α_1, α_2 :

$$g_{\alpha_1|\alpha_2}^{(1)} = g_{0|0}^{(1)} - \frac{\alpha_1}{2} + \frac{\sqrt{\alpha_1\alpha_2}}{2} \frac{K_{-1+\gamma+\frac{\alpha_2-\alpha_1}{2\sigma^2}}(\frac{\sqrt{\alpha_1\alpha_2}}{\sigma^2})}{K_{-\gamma+\frac{\alpha_2-\alpha_1}{2\sigma^2}}(\frac{\sqrt{\alpha_1\alpha_2}}{\sigma^2})}, \quad (5.6)$$

where $K_\nu(z)$ are modified Bessel functions of the second kind, $\sigma^2 := \frac{\sigma_{(1)}^2 + \sigma_{(2)}^2}{2}(1 - \rho)$ is the effective magnitude of stochasticity, ρ is the correlation across agents, and $\gamma := \frac{g_{0|0}^{(1)} - g_{0|0}^{(2)}}{\sigma^2}$ is the difference between the uncoupled growth rates with respect to stochasticity.

5.1.2 Numerical evaluation of the growth rate

The analytical expressions discussed in [subsection 5.1.1](#) have been implemented in Julia and Wolfram Mathematica programming languages. The code numerically evaluates the average long-term growth rate of the agents in the presence of wealth sharing, for

given parameters $(p_1, \mathcal{I}, \alpha_1, \alpha_2, b_\uparrow, b_\downarrow)$, relying on the `SpecialFunctions` package for the evaluation of the modified Bessel functions of the second kind.

Evaluating numerically the average long-term growth rate, for arbitrary investment strategies, as a function of p_1 (Figure 5.1), a first property emerges:

$$\alpha_1 > 0 \wedge \alpha_2 > 0 \implies g_{\alpha_1|\alpha_2}^{(1)} = g_{\alpha_1|\alpha_2}^{(2)}, \quad (5.7)$$

for all sets of physical values of the remaining parameters $(p_1, \mathcal{I}, b_\uparrow, b_\downarrow)$.

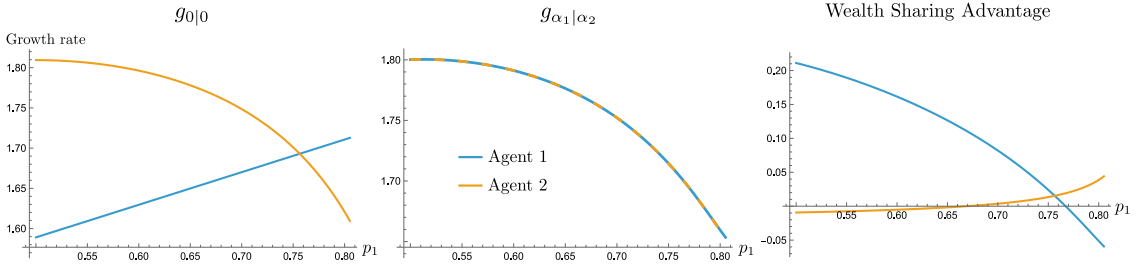


Figure 5.1. Growth rates for zero and non-zero wealth sharing, and the individual wealth sharing advantage $g_{\alpha_1|\alpha_2}^{(i)} - g_{0|0}^{(i)}$, for investment strategies $f_1 = 0.6$, $f_2 = \hat{f}_2(p_1)$, sharing rates $\alpha_1 = 0.2$, $\alpha_2 = 0.05$, system parameters $b_\uparrow = 10$, $b_\downarrow = 0$, $\mathcal{I} = 0.2$.

The coincidence of the two growth rate values for arbitrarily small sharing rates can be understood considering the fact that any difference in growth rate between the two agents is exponentially amplified in terms of accumulated resources over the infinite time horizon we are considering. Consequently, the agent experiencing the lower multiplicative noise dominates the long-term dynamics, effectively taking the lead by setting the pace of collective growth. The other agent, whose fluctuations are larger, has a growth rate dictated by the resources inflow received through sharing, and therefore reflects the partner's growth rate, according to the stationarity of the distribution of $\langle \exp(q_2 - q_1) \rangle_{eq}$.

Furthermore, a deeper understanding of this cooperative coupling between the agents' dynamics can be obtained by comparing the wealth-sharing growth rate with the uncoupled one. The difference between the two, $g_{\alpha_1|\alpha_2}^{(i)} - g_{0|0}^{(i)}$, quantifies the net advantage of cooperation.

As illustrated in Figure 5.1, we can see how the wealth sharing advantage for agent 2 is negative when its uncoupled growth rate significantly exceeds the partner's, and it is positive when it benefits from the other agent's faster growth. Interestingly, when the two uncoupled growth rates are equal, the advantage in sharing remains positive. The collective benefit originates from the threefold diversification of investments: each agent effectively allocates its resources not only into the two states of the environment, but also into the partner's wealth. The latter acts as a reinvesting channel, by gambling (approximately with the same proficiency) with the received resources and sharing at the next time step, thereby damping fluctuations.

This effect is therefore a first hallmark of an advantageous cooperative behaviour, and we will investigate it in the following.

5.1.3 Optimal investment strategy

The agents' objective is to maximise the long-term growth of their resources. Their investment strategies will therefore be designed to accomplish this task. In general, for agents able to share their wealth, such strategies would correspond to $\arg \max_{f_i} g_{\alpha_1|\alpha_2}^{(i)}$, where the optimal fraction depends on the signal reliability, as in [Equation 4.15](#), but also on the coupled dynamics through the sharing rates. The mutual dependence of the two agents' strategies in the three-dimensional parameter space $(p_1, \alpha_1, \alpha_2)$ makes the analytical treatments of this optimisation non-trivial.

For this reason, in the following analysis here, we will adopt a simplified but significant approach: we will assume that each agent invests according to the uncoupled optimal fractions, i.e., the values $\hat{f}_i$ that maximise the growth rate in the absence of wealth sharing $g_{0|0}^{(i)}$. This choice allows us to isolate and characterise the sole effect of wealth sharing on the growth dynamics, without introducing additional dependencies due to reoptimisation of the investment strategy. In other words, agents are considered to follow their individual Kelly strategies while optimising the wealth sharing rates, so that any observed change in growth rate can be attributed exclusively to the cooperative behaviour rather than to a redefinition of their investment strategy. Moreover, this assumption can be interpreted as a realistic constraint: in biological, social, or economic systems, agents may not continuously reoptimise their strategies due, for example, to adaptation timescales or inertia in their response. The investment rule that is optimal in isolation thus serves as a baseline behaviour, with respect to which the effects of cooperation can be evaluated.

5.2 Optimal collective behaviours

The study of the emergence of cooperation requires analysing how cooperation behaves under small perturbations of the agents' strategies. In this context, an evolutionary stable state (ESS) corresponds to a configuration of the agents' parameters that cannot be invaded by small deviations in the sharing rate or in the information partition.

For our analysis, we focus on the analysis of the average long-term growth rate $\hat{g}_{\alpha_1|\alpha_2}$, already optimised over the agents' investment strategies. As a first step, we consider the simplified case where both agents adopt the same sharing rate. This allows us to characterise the collective behaviours before introducing asymmetrical strategic choices.

5.2.1 Equal sharing rates

A first step in the study of the average long-term growth rate optimised as specified in [subsection 5.1.3](#), denoted as $\hat{g}_{\alpha_1|\alpha_2}$, is to consider the special case in which the two agents adopt the same sharing rate

$$\alpha := \alpha_1 = \alpha_2. \quad (5.8)$$

We can numerically evaluate $\hat{g}_{\alpha|\alpha}(p_1)$ for different values of the common sharing rate α , as shown in [Figure 5.2](#). It is possible to identify a threshold $\alpha_c \approx 0.35$ corresponding

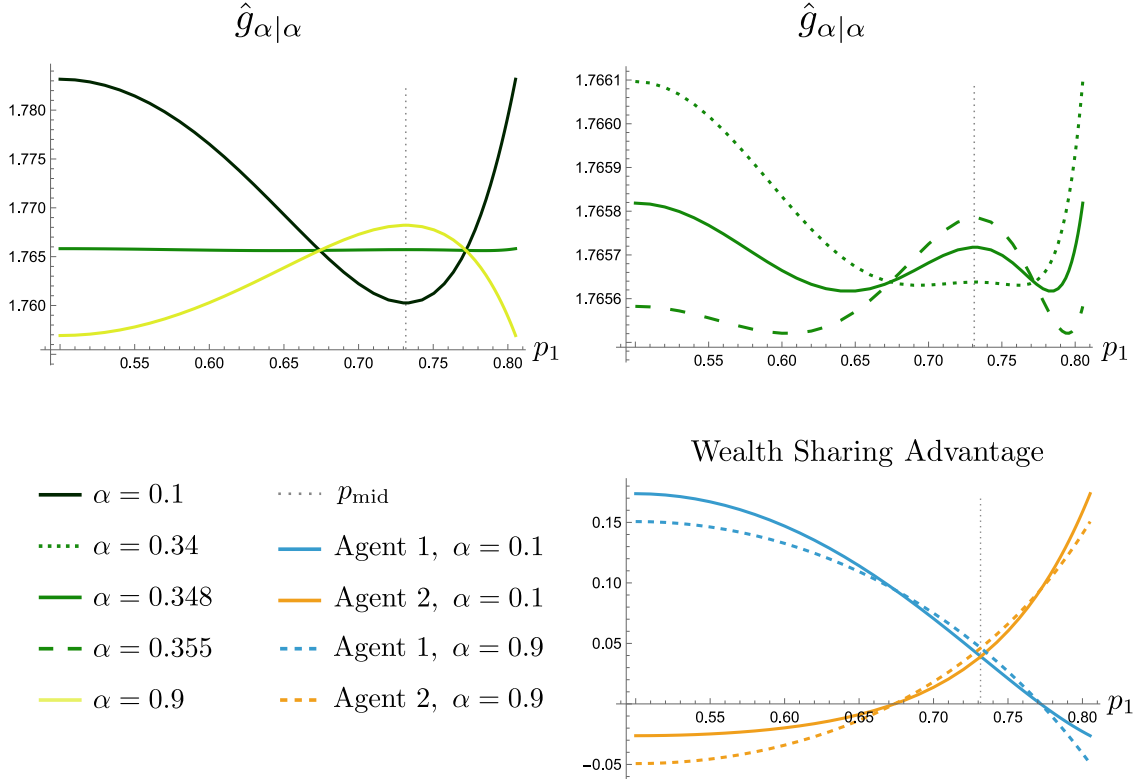


Figure 5.2. Average long-term growth rate $\hat{g}_{\alpha|\alpha}$ and wealth sharing advantage $\hat{g}_{\alpha|\alpha} - \hat{g}_{0|0}^{(i)}$, for various values of $\alpha = \alpha_1 = \alpha_2$, with parameters $\mathcal{I} = 0.2$, $b_{\uparrow} = 10$, $b_{\downarrow} = 0$.

to the critical point of a supercritical pitchfork bifurcation [20].

By examining the wealth sharing advantage, $\hat{g}_{\alpha|\alpha} - \hat{g}_{0|0}^{(i)}$, we observe that the difference between the curves above and below α_c lies in the fact that at equal information partition higher sharing rates enhance growth more than lower ones, while for unbalanced information partition the situation is the opposite.

A broader picture can be obtained by numerically evaluating the growth rate $\hat{g}_{\alpha|\alpha}$ in the bidimensional parameter space (p_1, α) . As shown in Figure 5.3, the transition along the α direction gives rise to the coexistence of multiple local maxima. These can be regarded as attraction points of the system's evolutive dynamics in this restricted parameter space.

We can identify two qualitatively distinct evolutionary outcomes:

Asymmetric attractors: They are characterised by one agent that effectively drives the growth by employing all the information, while the other agent follows. A wealth sharing rate close to zero reduces the cost of maintaining the slower agent, while any value of α strictly higher than zero is yet sufficient for the follower agent to share the leader's growth rate. On the other hand, both agents' advantage comes from supplying the leading agent with as much information as possible, thus pushing the system to the extremes of the p_1 axis.

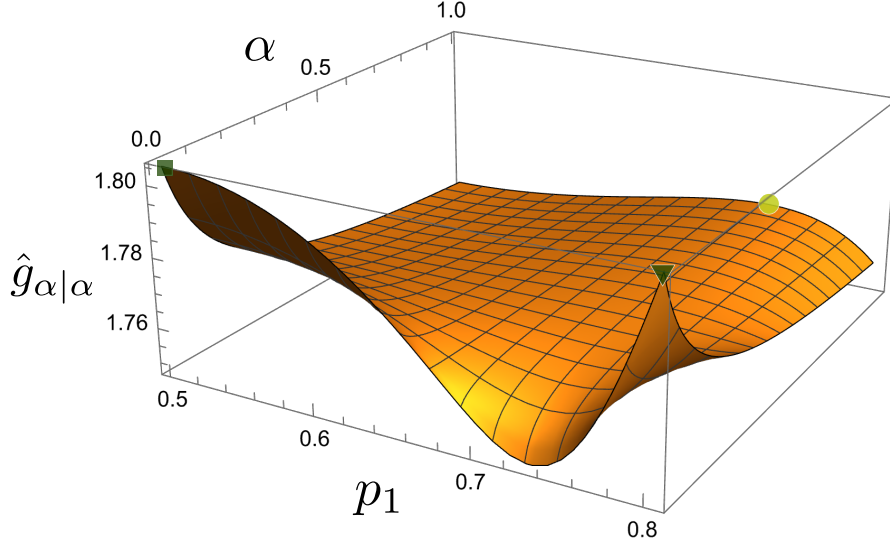


Figure 5.3. Average long-term growth rate $\hat{g}_{\alpha|\alpha}$ as function of p_1 and $\alpha = \alpha_1 = \alpha_2$, for $\alpha \in [0.01, 0.99]$, with parameters $\mathcal{I} = 0.2$, $b_{\uparrow} = 10$, $b_{\downarrow} = 0$. The two kinds of attraction points have been highlighted with different markers: (light green circle) cooperation attraction point $(p_{\text{mid}}, 1)$; (dark green triangle and square) asymmetric attraction points, respectively $(p_{\text{max}}, 0.01)$ and $(0.5, 0.01)$.

Because of the source of growth of the follower agent, this point of attraction can describe "parasitic" systems, or, taking the opposite point of view, "greedy" behaviours.

In our case, restricted to $\alpha > 0$, we observe two symmetric attractors of this kind, namely $(p_{\text{max}}, 0.01)$ and $(0.5, 0.01)$, highlighted with dark green markers in Figure 5.3.

Full cooperation attractors: These attraction points represent evolutionary outcomes in which both agents equally participate to cooperate. Along the information partition axis, stability is ensured by the fact that the equal partition point p_{mid} corresponds to the maximum position. As shown in [6], along the wealth sharing rate axis, agents with equal uncoupled growth rates take advantage in sharing more than the partner, leading to the ESS $\alpha_{\text{ess}} \rightarrow \infty$. Therefore, if we restrict our system to the one represented in Figure 5.3 with values of $\alpha = \alpha_1 = \alpha_2 < 1$, we can identify the point $(p_{\text{mid}}, \alpha_{\text{max}})$ as an ESS (highlighted with a light green marker in figure); more in general, systems with $\alpha_1 = \alpha_2$ exhibit the evolutionary stable state $(p_{\text{mid}}, \alpha \rightarrow \infty)$.

5.2.2 Asymmetric sharing rates

We are now ready to generalise the framework investigated in subsection 5.2.1 by relaxing the constraint of equal wealth sharing rates. We can still numerically evaluate the average long-term growth rate, but the now three-dimensional parameter space $(p_1, \alpha_1, \alpha_2)$

requires different methods for a visual representation. Therefore, we take a step further in the direction of addressing the questions concerning not only the location of the attractors, but also the size of the basins of attraction, i.e. the sets of points in the parameter space leading to each of the points of attraction. With this in mind, we numerically evaluate the gradient of the average long-term growth rate by making use of the Julia package `FiniteDifferences`. Additionally, we perform the gradient evaluation in the rescaled parameter space $(p_1, \alpha_1, \alpha_2) \rightarrow (I_1, \alpha_1, \alpha_2)$, using the single agent information measure [Equation 3.23](#). The latter ensures symmetry along the information partition direction, thus paving the way for a more significant study of the basins of attraction by properly rescaling the finite differences employed in the gradient evaluation. Furthermore, it allows us to restrict our discussion only to the first half of the information partition axis $I_1 \in [0, \mathcal{I}/2]$, being the gradient such that

$$\nabla \hat{g}_{\alpha_2|\alpha_1}(\mathcal{I} - I_1) = \left(-\frac{\partial}{\partial I_1} \hat{g}_{\alpha_1|\alpha_2}(I_1), \quad \frac{\partial}{\partial \alpha_2} \hat{g}_{\alpha_1|\alpha_2}(I_1), \quad \frac{\partial}{\partial \alpha_1} \hat{g}_{\alpha_1|\alpha_2}(I_1) \right), \quad (5.9)$$

where $\nabla = (\partial_{I_1}, \partial_{\alpha_1}, \partial_{\alpha_2})$.

For the analysis of the gradient, we implemented a Julia code that, using the `Distributed` package for parallelisation, evaluates the gradient on a three-dimensional lattice; generates walkers on every lattice point; tracks their walk in the continuous parameter space, following at every step the previously evaluated gradient for the currently occupied lattice cell. In this way, we are able to characterise the range of parameters considered in the previous [subsection 5.2.1](#) $\alpha_1, \alpha_2 \in (0,1)$.

We find that the restricted dynamics converge to three attraction points, shown in [Figure 5.4](#), which fall into the same two categories exposed in [subsection 5.2.1](#), at coordinates:

$$C = (\mathcal{I}/2, \alpha_{\max}, \alpha_{\max}), \quad (5.10)$$

$$G = (0, \alpha_{\max}, \alpha_{\min}), \quad (5.11)$$

$$P = (\mathcal{I}, \alpha_{\min}, \alpha_{\max}). \quad (5.12)$$

The full cooperation attraction point C shares the same coordinates as the one encountered in the symmetric sharing case. Indeed, the surface shown in [Figure 5.3](#) lies on the diagonal surface of the new "cubic" parameter space that contains the I_1 axis and the point C . The two asymmetric attractors, instead, have different sharing rates with respect to the previous ones. The increase in the sharing rate of the follower agent can be explained by the fact that, when gambling without any information, the best effective investment is represented by the other agent. Hence, these two attractors can represent a slightly different, more exploitative type of parasitic relation.

To further characterise these attractors, we can now focus on their basins of attraction. In [Figure 5.5](#), we colour every point of the parameter space lattice, up to the equal information-partition surface, with the colour of the final point of the corresponding dynamics. We rescale each coordinate to be in the interval $[0, 255]$ and use the codification

$$\text{Red} = I_1 \quad (5.13)$$

$$\text{Green} = \alpha_1 \quad (5.14)$$

$$\text{Blue} = \alpha_2, \quad (5.15)$$

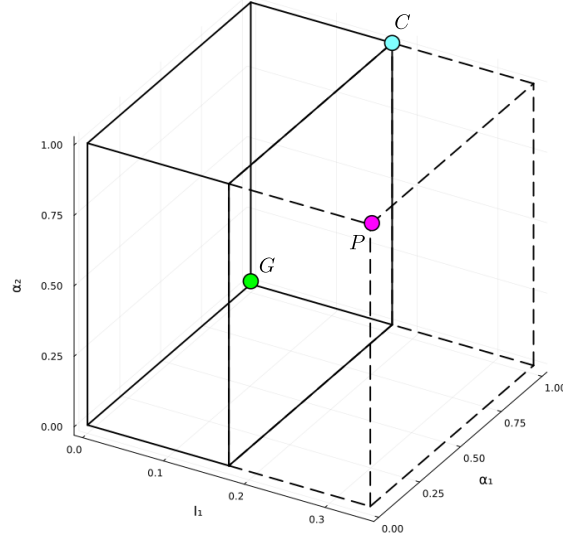


Figure 5.4. Attraction points in the parameters space $(I_1, \alpha_1, \alpha_2)$, for $\alpha_1, \alpha_2 \in (0,1)$, with $\mathcal{I} = 0.35$, $b_{\uparrow} = 10$, $b_{\downarrow} = 0$. The colours of the markers correspond to their coordinates codified in RGB. The dashed-line part of the cube is redundant by symmetry with the solid-line part.

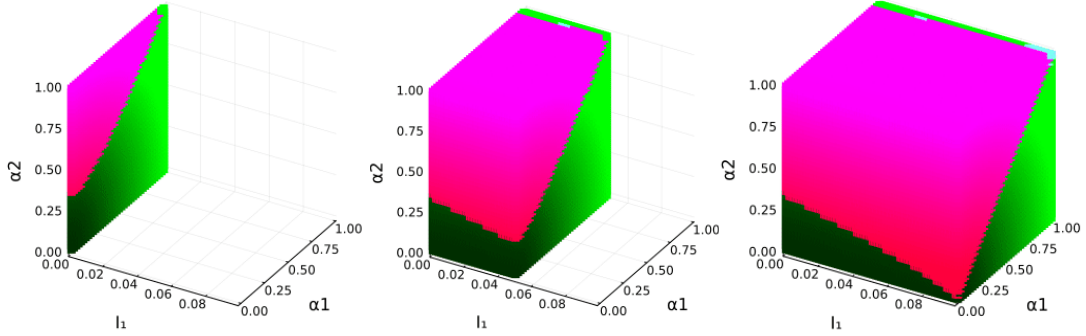


Figure 5.5. Basins of attraction over the first half of the information partition axis, for $\alpha_1, \alpha_2 \in (0,1)$, with $\mathcal{I} = 0.2$, $b_{\uparrow} = 10$, $b_{\downarrow} = 0$. Each colored cell corresponds in position to the starting point of a single dynamics, and in colour to the coordinates, codified in RGB, of the final point of that dynamics.

so that, for example, the full cooperation attraction point C is coloured in (R: 128, G: 255, B:255), while the colour shading is due to the slow optimisation dynamics of some points.

The structure of the basins of attraction strongly depends on the total information $\mathcal{I}$ available to the agents. By varying this parameter, we can explore how environmental predictability shapes the stability and prevalence of cooperative behaviours.

For low values of $\mathcal{I}$, the agents have limited predictive ability, and stochastic fluctuations dominate the dynamics. In this regime, the cooperative attractor is smaller and confined to initial conditions with already high levels of sharing. As $\mathcal{I}$ increases, the agents' capacity to anticipate environmental states improves, and the basin of full cooperation expands, progressively incorporating configurations with weaker initial co-operation.

This behaviour reflects the stabilising role of information: higher environmental predictability not only enhances the average growth rate but also makes cooperation more resistant to deviations. The qualitative change in the shape and size of the basins as $\mathcal{I}$ grows can be observed by the comparison of Figure 5.5, at $\mathcal{I} = 0.2$, with Figure 5.6.

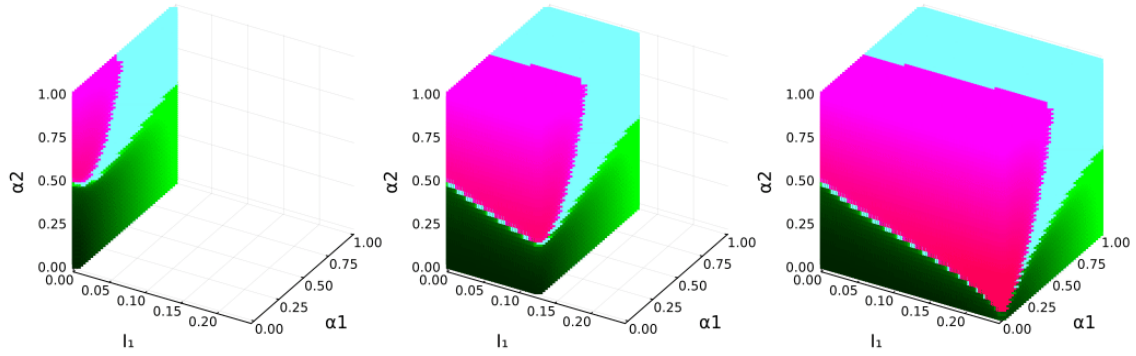


Figure 5.6. Basins of attraction over the first half of the information partition axis, for $\alpha_1, \alpha_2 \in (0,1)$, with $\mathcal{I} = 0.5$, $b_{\uparrow} = 10$, $b_{\downarrow} = 0$.

Chapter 6

Conclusion

This thesis has investigated the emergence and stability of cooperation in stochastic environments through the lens of information partitioning and wealth-sharing mechanisms. By integrating concepts from information theory, stochastic processes, and evolutionary game theory, we have provided new insights into how environmental uncertainty shapes cooperative behaviours in multiplicative growth systems.

Results

Our first contribution is introducing a total information constraint in a system of agents with partial knowledge about a fluctuating environment. This serves as a framework for comparing systems with equivalent predictive capabilities and introduces the additional dimension of information partition for the optimisation of growth.

In [chapter 3](#), we formally defined the total information constraint, presented the relevant properties of mutual information and proposed an information measure able to capture the symmetries of the system under study. In [chapter 4](#), we explored the implications of the total information constraint on the growth of two-agent systems. We discussed how limited knowledge about a binary stochastic environment can influence the decision-making process and shape the optimal average long-term growth rate. We showed how, depending on the combination of the values of investment return factors and individual knowledge, the growth of an agent's resources can have two kinds of regimes. The first regime involves non-diversified investments. Whenever the two potential returns from investing resources become sufficiently similar, a condition that allows the individual information to adequately compensate for the risk of loss, leading to the diversification breakdown. The second regime requires diversified investments to handle the amplitude of variations. In this case, the growth rate is given by the sum of two terms: the first, dependent on the reward factors, accounts for the generosity of the system; the second term is dependent on the signal reliability and expresses, through entropy, the knowledge about the environment.

A further substantial contribution of this thesis lies in the explicit analysis of the interplay between information partitioning and wealth sharing.

In [chapter 5](#), we integrated the mechanism of information partition into the framework

of [6], which allows for the sharing of resources. This approach enabled us to quantitatively investigate the combination of these two distinct forms of interaction, both operating under the constraint on the value of total information about a stochastic environment. We characterised the optimal strategies for maximising the collective long-term growth rate, and found two kinds of attraction points, respectively corresponding to asymmetric leader-follower agent roles and equitable full cooperation. The former can describe exploitation relations, relying on the (minimum) generosity of an agent that employs the totality of the available information to improve its growth, and consequently the partner's. The latter kind of attraction point, full cooperation, describes a completely symmetric strategy. It shows that an equal information partition can be stable in sufficiently cooperative resource-sharing contexts. This implies that in the corresponding basin of attraction, agents with superior initial knowledge of the environment are willing to sacrifice part of it to reach equality, the advantage of this informational cost coming from the enhancement of the performance of the diversification channel provided by the partner.

Finally, we numerically illustrated how higher values of the total available information constraint lead to larger sizes of the basin of attraction of cooperation, therefore expanding the range of initial strategies that converge to an egalitarian cooperative behaviour. Contrarily, systems with more unforeseeable fluctuations are more inclined to produce social disintegration.

Model limitations and future directions

Several simplifying assumptions of our model suggest directions for future research. Our analysis focused on the two-agent case for analytical tractability. Extending to N-agent systems would reveal how cooperation scales with group size and whether new qualitative behaviours emerge in larger populations, eventually integrating network structures. The choice of a binary environment with equiprobable states, while making the mathematical analysis more treatable, limits the generalisability to more complex environments. Future work could explore, for example, the introduction of diverse probability distributions for the environment's states. The optimisation of the wealth-sharing average long-term growth rate (currently in progress) could provide new insights into collective behaviours. The introduction of explicit costs for cooperative behaviours would allow for a generalisation of the model towards more realistic situations. Finally, we are currently investigating the possibility of agents sharing the private signals with the partners through noisy channels. This will introduce the mechanism of information sharing in the current framework, requiring the development of more sophisticated collective strategies.

Chapter 7

Appendix

7.1 Continuous limit of the discrete-time GBM process

Our discrete-time GBM process can be written as

$$x(t+1) = x(t)\zeta(t) \quad (7.1)$$

$$q(t+1) = q + \log \zeta(t), \quad (7.2)$$

where $\log \zeta(t)$ is a random variable with $\mathbb{E}[\log \zeta(t)] := g$ and $\text{Var}[\log \zeta(t)] := s^2$. We can rewrite [Equation 7.2](#) as

$$q(t + M\Delta t) = q(t) + \sum_{m=1}^M \log \tilde{\zeta}(t - m\Delta t), \quad (7.3)$$

assuming the discrete-time noise term to be, in the limit $\Delta t \rightarrow 0$, the result of a continuous process or, in other words, a Gaussian noise

$$\log \zeta(t) = \sum_{m=1}^M \log \tilde{\zeta}(t - m\Delta t) \sim \mathcal{N}(g \cdot M\Delta t, s^2 \cdot M\Delta t) \quad (7.4)$$

$$\mathbb{E}[\log \tilde{\zeta}(t + \Delta t)] = g\Delta t \quad (7.5)$$

$$\text{Var}[\log \tilde{\zeta}(t + \Delta t)] = s^2\Delta t, \quad (7.6)$$

implying

$$\zeta(t) = \prod_{m=1}^M \tilde{\zeta}(t - m\Delta t) \quad (7.7)$$

$$\sim \text{Lognormal}(g \cdot M\Delta t, s^2 \cdot M\Delta t) \quad (7.8)$$

$$\sim \exp(gM\Delta t + s\sqrt{M\Delta t}\eta(t)), \quad (7.9)$$

where $\eta(t) \sim \mathcal{N}(0, 1)$.

Assuming continuous time as the limit $dt \rightarrow 0$ where $\Delta t = N dt$ in such a way that

$\frac{\Delta t}{dt} = N$, we can write

$$x(t + \Delta t) = x(t) + \log [\exp (g\Delta t + s\sqrt{\Delta t N} d\xi(t))] \quad (7.10)$$

$$= x(t) + g\Delta t + s\Delta t \xi(t), \quad (7.11)$$

where $\eta(t)$ has been interpreted as a Wiener process, since it is the result of the sum of “microscopic” noises $\xi(t) \sim \mathcal{N}(0, 1)$. Finally, the limit for $\Delta t \rightarrow 0$ gives

$$\dot{x}(t) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} (g\Delta t + s\Delta t \xi(t)) = g + s\xi(t), \quad (7.12)$$

which features *approximately* the same values of mean and variance as the discrete process.

An alternative way to derive the continuous process from the discrete one is the following. We write the noise $\zeta(t)$ as

$$\zeta(t) = \mathbb{E}[\zeta(t)] + \text{Std}[\zeta(t)] \tilde{\zeta}(t) \quad (7.13)$$

$$= 1 + \mu + \sigma \tilde{\zeta}(t), \quad (7.14)$$

where $\zeta(t) \sim \text{Lognormal}(a, b)$. The equation for $x(t)$ becomes

$$\begin{aligned} x(t + M\Delta t) &= x(t + (M-1)\Delta t) [1 + \mu\Delta t + \sigma\sqrt{\Delta t} \tilde{\zeta}(t + (M-1)\Delta t)] \\ &= x(t) [1 + M\Delta t\mu + \sigma\sqrt{\Delta t} \sum_{m=1}^M \tilde{\zeta}(t - m\Delta t) + \\ &\quad + \sigma^2\Delta t \sum_{m=1}^{M-1} \sum_{l=m+1}^M \tilde{\zeta}(t - m\Delta t) \tilde{\zeta}(t - l\Delta t) + O((\Delta t)^2)] \\ &= x(t) [1 + M\Delta t\mu + \sigma\sqrt{\Delta t} \sqrt{M} \eta(t) + \sigma^2\Delta t \sqrt{\frac{M(M-1)}{2}} \eta'(t) + O((\Delta t)^2)] \\ &\simeq x(t) [1 + M\Delta t\mu + \sigma\sqrt{M\Delta t} \eta(t) + \frac{1}{\sqrt{2}} M\Delta t \eta'(t)] \end{aligned} \quad (7.15)$$

$$\doteq x(t) \zeta^*(t), \quad (7.16)$$

where η, η' are standard Gaussian noise terms, and the equation for the logarithms for a single Δt step reads

$$\begin{aligned} \log \zeta^*(t) &= \log [1 + \Delta t\mu + \sigma\sqrt{\Delta t} \eta(t) + \frac{1}{\sqrt{2}} M\Delta t \eta'(t)] \\ &\simeq \Delta t\mu + \sigma\sqrt{\Delta t} \eta(t) + \frac{1}{\sqrt{2}} \sigma^2 \Delta t \eta'(t) - \frac{1}{2} \Delta t \eta^2(t) + O((\Delta t)^{3/2}) \\ &\simeq \Delta t\mu + \sigma\sqrt{\Delta t N} d\xi(t) - \frac{1}{2} \sigma^2 \Delta t - \frac{1}{2} \sigma^2 \Delta t \sqrt{N} d\xi_\chi(t) + O((\Delta t)^{3/2}), \end{aligned} \quad (7.17)$$

where ξ, ξ_χ are other standard Gaussian noise terms and taking again the limit, we get

$$\dot{q}(t) = \lim_{\substack{\Delta t \rightarrow 0 \\ \Delta t/dt \rightarrow N}} \frac{1}{\Delta t} \log \zeta^*(t) \simeq \left(\mu - \frac{\sigma^2}{2} \right) + \sigma \xi(t). \quad (7.18)$$

7.2 Itô integration

The amount of wealth an agent has at a certain time $x(t)$ is statistically independent of the noise at time $t' \geq t$ due to the causality of the process. The Stochastic Differential Equations (SDE) described by Equation 2.13 can be exactly solved using Itô Integration. A stochastic quantity $x(t)$ obeys the general Itô SDE

$$dx(t) = a(x(t), t) dt + b(x(t), t) dW(t), \quad (7.19)$$

if $\forall t, t_0$

$$x(t) = x_0 + \int_{t_0}^t a(x(t'), t') dt' + \int_{t_0}^t b(x(t'), t') dW(t'), \quad (7.20)$$

and for the solution to exist and to be unique, the following are required

- Lipschitz condition

$$\exists K : |a(x, t) - a(y, t)| + |b(x, t) - b(y, t)| \leq K|x - y|. \quad (7.21)$$

- Growth condition

$$\exists K : \forall t \in [t_0, T] \quad |a(x, t)|^2 + |b(x, t)|^2 \leq K^2(1 + |x|^2). \quad (7.22)$$

In the following, we will assume that these conditions are satisfied due to the finiteness of the process.

In GBM case coefficients are products of $x(t)$ and do not explicitly depend on t :

$$a(x(t), t) = a(x(t)) = \mu x(t) \quad (7.23)$$

$$b(x(t), t) = b(x(t)) = \sigma x(t). \quad (7.24)$$

Then, the SDE becomes

$$dx(t) = \mu x(t) dt + \sigma x(t) dW(t) \quad (7.25)$$

$$= \mu x(t) dt + \sigma x(t) \epsilon(t) dt, \quad (7.26)$$

where we used the Wiener process definition as the integral of a Gaussian white noise $\epsilon(t)$

$$\int_0^t \epsilon(t') dt' = W(t) \quad (7.27)$$

$$dW(t) = W(t + dt) - W(t) = \epsilon(t) dt, \quad (7.28)$$

and we renamed the coefficients to match their meaning of mean (μ) and variance (σ^2) in the GBM case.

We can now recall the following properties of the Wiener process in non-anticipating functions' Itô integrals

$$dW(t)^2 \rightarrow dt \quad (7.29)$$

$$dW(t)^{2+N} \rightarrow 0 \quad \forall N > 0 \quad (7.30)$$

$$dW(t) dt \rightarrow 0. \quad (7.31)$$

These imply that when differentiating a function $f(W(t), t)$ we need to keep only terms up to $dW(t)^2$:

$$\begin{aligned} df(W(t), t) &= \frac{\partial f}{\partial t} dt + \frac{1}{2} \frac{\partial^2 f}{\partial t^2} (dt)^2 + \frac{\partial f}{\partial W} dW(t) + \frac{1}{2} \frac{\partial^2 f}{\partial W^2} [dW(t)]^2 + \\ &\quad + \frac{\partial^2 f}{\partial W \partial t} dt dW(t) + \dots \\ &= \left(\frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial W^2} \right) dt + \frac{\partial f}{\partial W} dW(t). \end{aligned} \quad (7.32)$$

We can now consider the case of an arbitrary function of $x(t)$: $f[x(t)]$. It will obey the differential equation

$$\begin{aligned} df[x(t)] &= f[x(t) + dx(t)] - f[x(t)] \\ &= f'[x(t)] dx(t) + \frac{1}{2} f''[x(t)] dx(t)^2 + \dots \\ &= f'[x(t)] \cdot \left\{ a[x(t), t] dt + b[x(t), t] dW(t) \right\} + \\ &\quad + \frac{1}{2} f''[x(t)] \cdot b[x(t), t]^2 dt + \dots \\ &= \left\{ a[x(t), t] f'[x(t)] + \frac{1}{2} b[x(t), t]^2 f''[x(t)] \right\} dt + b[x(t), t] f'[x(t)] dW(t), \end{aligned} \quad (7.33)$$

where terms of higher order have been discarded. Equation 7.33 is known as *Itô's formula*. It shows that a variable change such as (2.14) is not given by ordinary calculus since $f[x(t)]$ is not linear in $x(t)$.

In the specific case of our variable change (2.14), in the GBM frame, we get

$$\begin{aligned} dq(t) &= \left\{ \mu x(t) \frac{1}{x(t)} + \frac{1}{2} \sigma^2 x(t)^2 \left(-\frac{1}{x(t)^2} \right) \right\} dt + \sigma x(t) \frac{1}{x(t)} dW(t) \\ &= \left\{ \mu - \frac{1}{2} \sigma^2 \right\} dt + \sigma dW(t) \\ &= \left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma dW(t) \end{aligned} \quad (7.34)$$

$$= \left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma \epsilon(t) dt. \quad (7.35)$$

That is, we obtain a Brownian Motion for q

$$a[x(t), t] = a = \mu - \frac{\sigma^2}{2} \quad (7.36)$$

$$b[x(t), t] = b = \sigma \quad (7.37)$$

$$\dot{q}(t) = \left(\mu - \frac{\sigma^2}{2} \right) + \sigma \epsilon(t), \quad (7.38)$$

where we define

$$g_0 := \mu - \frac{\sigma^2}{2}, \quad (7.39)$$

so that Equation 7.38 simply reads

$$\dot{q}(t) = g_0 + \sigma \epsilon(t). \quad (7.40)$$

Finally, we can solve Equation 7.40 integrating as [6]

$$q(t) = g_0 t + \sigma \int_0^t dt' \epsilon(t'). \quad (7.41)$$

We can rewrite the last integral as

$$\int_0^t dt' \epsilon(t') = Q(t) \eta(t), \quad (7.42)$$

where

$$Q^2(t) = \int_0^t ds' \int_0^t ds'' \langle \epsilon(s') \epsilon(s'') \rangle, \quad (7.43)$$

and using in the general case the noise correlation to have the form

$$\langle \epsilon(t) \epsilon(t') \rangle = \exp(-|t - t'|/\tau)/2\tau, \quad (7.44)$$

we get that in the case of delta-correlated noise ($t \gg \tau$) $Q(t) \sim \sqrt{t}$; η is a Gaussian random variable with zero mean and variance 1. Then we can write the equation for $x(t)$ as

$$x(t) \approx x(0) \exp \left(g_0 t + \sigma Q(t) \eta(t) \right) \quad (7.45)$$

$$\approx x(0) \exp \left(g_0 t + \sigma \sqrt{t} \eta(t) \right), \quad (7.46)$$

$$\langle x(t) \rangle \approx x(0) \exp \left(\left(g_0 + \frac{\sigma^2}{2} \right) t \right) \quad (7.47)$$

$$\approx x(0) e^{\mu t}. \quad (7.48)$$

7.3 Information evaluation

We can evaluate each term of the total information constraint

$$\mathcal{I} = I(E; S_1) + I(E; S_2) - I(S_1; S_2) + I(S_1; S_2|E), \quad (7.49)$$

using the definition

$$I(X; Y) = \sum_{\{X\}} \sum_{\{Y\}} P(X, Y) \log \left(\frac{P(X, Y)}{P(X)P(Y)} \right), \quad (7.50)$$

as follows:

- Information carried by a single signal S_i

$$I(E; S_i) = \sum_{\{E\}} \sum_{\{S_i\}} P(E, S_i) \log \left(\frac{P(E, S_i)}{P(E)P(S_i)} \right) \quad (7.51)$$

$$= \sum_{\{E\}} \sum_{\{S_i\}} P(S_i|E)P(E) \log \frac{P(S_i|E)}{P(S_i)} \quad (7.52)$$

$$= \frac{1}{2} \sum_{\{E\}} \sum_{\{S_i\}} [p_i \delta_{S_i, E} + (1 - p_i) \delta_{S_i, -E}] \cdot [\log 2 + \log(p_i \delta_{S_i, E} + (1 - p_i) \delta_{S_i, -E})] \quad (7.53)$$

$$= \log 2 + p_i \log(p_i) + (1 - p_i) \log(1 - p_i), \quad (7.54)$$

where we used $P(E) = 1/2$ and $P(S_i|E) = p_i \delta_{S_i, E} + (1 - p_i) \delta_{S_i, -E}$.

- Mutual information between the two signals

$$I(S_1; S_2) = \sum_{\{S_1\}} \sum_{\{S_2\}} P(S_1, S_2) \log \left(\frac{P(S_1, S_2)}{P(S_1)P(S_2)} \right) \quad (7.55)$$

$$= \sum_{\{S_1\}} \sum_{\{S_2\}} \sum_{\{E'\}} P(S_1, S_2, E') \log \left(\frac{P(S_1, S_2)}{P(S_1)P(S_2)} \right) \quad (7.56)$$

$$= \sum_{\{S_1\}} \sum_{\{S_2\}} \sum_{\{E'\}} P(S_1|E')P(S_2|E')P(E') \log \left(\frac{P(S_1, S_2)}{P(S_1)P(S_2)} \right) \quad (7.57)$$

$$= \frac{1}{2} \sum_{\{S_1\}} \sum_{\{S_2\}} \sum_{\{E'\}} P(S_1|E')P(S_2|E') \log P(S_1, S_2) + \quad (7.58)$$

$$- \frac{1}{2} \sum_{\{E'\}} \left(\sum_{\{S_1\}} P(S_1|E') \right) \left(\sum_{\{S_2\}} P(S_2|E') \right) \log P(S_2) +$$

$$- \frac{1}{2} \sum_{\{E'\}} \left(\sum_{\{S_1\}} P(S_1|E') \right) \left(\sum_{\{S_2\}} P(S_2|E') \right) \log P(S_1)$$

$$= \frac{1}{2} \sum_{\{S_1\}} \sum_{\{S_2\}} \sum_{\{E'\}} P(S_1|E')P(S_2|E') \log P(S_1, S_2) - 2 \log \frac{1}{2}. \quad (7.59)$$

Now we can use the parameterisations of $P(S_1|E)$ and $P(S_2|E)$ in the log in the

first term to write

$$\log P(S_1, S_2) = \log \left[\sum_{\{E''\}} P(S_1, S_2|E'')P(E'') \right] \quad (7.60)$$

$$= \log \left\{ \sum_{\{E''\}} \left[p_1 \delta_{S_1, E''} + (1 - p_1) \delta_{S_1, -E''} \right] \cdot \right. \\ \left. \cdot \left[p_1 \delta_{S_1, E''} + (1 - p_1) \delta_{S_1, -E''} \right] \right\} - \log 2 \quad (7.61)$$

$$= \log \left\{ (1 - \pi(p_1, p_2)) \delta_{S_1, S_2} + \pi(p_1, p_2) \delta_{S_1, -S_2} \right\} - \log 2, \quad (7.62)$$

where $\pi(x, y) := x + y - 2xy$, and therefore we find

$$I(S_1; S_2) = \pi(p_1, p_2) \log(\pi(p_1, p_2)) + \\ + (1 - \pi(p_1, p_2)) \log(1 - \pi(p_1, p_2)) + \log 2, \quad (7.63)$$

- Mutual information between the two signals, given the state of the environment

$$I(S_1; S_2|E) = \sum_{\{S_1\}} \sum_{\{S_2\}} \sum_{\{E\}} P(S_1, S_2, E) \log \frac{P(S_1, S_2|E)}{P(S_1|E)P(S_2|E)} \quad (7.64)$$

$$= \sum_{\{S_1\}} \sum_{\{S_2\}} \sum_{\{E\}} P(S_1, S_2, E) \log \frac{P(S_1|E)P(S_2|E)}{P(S_1|E)P(S_2|E)} \quad (7.65)$$

$$= 0. \quad (7.66)$$

7.3.1 Direct total information calculation

An alternative direct calculation for the total mutual information is the following:

$$I(E; \vec{S}) = \sum_{\{E\}\{\vec{S}\}} P(E, \vec{S}) \log \frac{P(E, \vec{S})}{P(E)P(\vec{S})} \quad (7.67)$$

$$= \sum_{\{E\}\{S_1\}\{S_2\}} P(E)P(S_1|E)P(S_2|E) \log \frac{P(E)P(S_1|E)P(S_2|E)}{P(E)P(S_1, S_2)} \quad (7.68)$$

$$= \frac{1}{2} \sum_{\{E\}\{S_1\}} P(S_1|E) \log P(S_1|E) + \frac{1}{2} \sum_{\{E\}\{S_2\}} P(S_2|E) \log P(S_2|E) + \quad (7.69)$$

$$- \frac{1}{2} \sum_{\{E\}\{S_1\}\{S_2\}} P(S_1|E)P(S_2|E) \log \left(\frac{1}{2} \sum_{\{E'\}} P(S_1|E')P(S_2|E') \right) \\ = \log 2 - h(p_1) - h(p_2) + h(\pi(p_1, p_2)). \quad (7.70)$$

List of Figures

3.1	Entropy representation of information. Mutual information between two events corresponds to the intersection of their entropy shape. The information $I(S_i; S_j)$ corresponds to the union of the green and bi-coloured areas. Overcounting of the latter area explains the redundancy term.	12
3.2	Shape of the information function on the $p_1 - p_2$ space.	13
3.3	Contour plot of the total information function on the first quarter of the (p_1, p_2) -plane centred in $(0.5, 0.5)$	14
3.4	Entropy representation of information, in the case $P(S_1, S_2 E) = P(S_1 E)P(S_2 E)$. The area shaded in grey corresponds to $\mathcal{I} = I(E; S_1, S_2)$, while the union of the hatched areas corresponds to $I(S_1; E)$	15
3.5	Single agent information measure for two values of total information $\mathcal{I}$. Both agents' measures and their components are shown.	16
4.1	Kelly optimal investment as function of p_1 and $b_\downarrow$, given $b_\uparrow = 10$. For $b_\downarrow = 0$, $\hat{f}_1(p_1) = p_1$. For non-zero values of $b_\downarrow$, the linear function acquires a shift and an angular coefficient determined by the combination of the two reward factors.	19
4.2	Average long-term growth rate in the case $f^* \in (0,1)$, shown for different values of the total information constraint, as a function of the two variables p_1 and I_1 , for parameters $b_\uparrow = 10, b_\downarrow = 0$. The minimum and maximum theoretical values, respectively $\log(b_\uparrow + b_\downarrow) - \log 2$ and $\log(b_\uparrow + b_\downarrow) - (\log 2 - \mathcal{I})$, are shown in grey. The I_1 variable representation emphasises symmetry between the two agents' growth rates around the equal information partition point, where they coincide. Higher information leads to faster growth, and brings the equal growth rate value closer to the upper bound, reflecting a higher redundancy $I(S_1; S_2)$ in information.	21
4.3	Investment fraction $\hat{f}_1$ for two values of the parameter $b_\downarrow$ and $\mathcal{I} = 0.2, b_\uparrow = 10$. The contour line of constant information $\mathcal{I}$ is shown to highlight the intersection between the two curves.	22
4.4	Investment fractions and corresponding growth rates, for both agents, with parameters $b_\uparrow = 10, \mathcal{I} = 0.2$	23
4.5	Investment fractions and growth rates for $b_\uparrow = 10, b_\downarrow = 3, \mathcal{I} = 0.5$	24

5.1	Growth rates for zero and non-zero wealth sharing, and the individual wealth sharing advantage $g_{\alpha_1 \alpha_2}^{(i)} - g_{0 0}^{(i)}$, for investment strategies $f_1 = 0.6$, $f_2 = \hat{f}_2(p_1)$, sharing rates $\alpha_1 = 0.2$, $\alpha_2 = 0.05$, system parameters $b_\uparrow = 10$, $b_\downarrow = 0$, $\mathcal{I} = 0.2$	27
5.2	Average long-term growth rate $\hat{g}_{\alpha \alpha}$ and wealth sharing advantage $\hat{g}_{\alpha \alpha} - \hat{g}_{0 0}^{(i)}$, for various values of $\alpha = \alpha_1 = \alpha_2$, with parameters $\mathcal{I} = 0.2$, $b_\uparrow = 10$, $b_\downarrow = 0$	29
5.3	Average long-term growth rate $\hat{g}_{\alpha \alpha}$ as function of p_1 and $\alpha = \alpha_1 = \alpha_2$, for $\alpha \in [0.01, 0.99]$, with parameters $\mathcal{I} = 0.2$, $b_\uparrow = 10$, $b_\downarrow = 0$. The two kinds of attraction points have been highlighted with different markers: (light green circle) cooperation attraction point $(p_{\text{mid}}, 1)$; (dark green triangle and square) asymmetric attraction points, respectively $(p_{\text{max}}, 0.01)$ and $(0.5, 0.01)$	30
5.4	Attraction points in the parameters space $(I_1, \alpha_1, \alpha_2)$, for $\alpha_1, \alpha_2 \in (0, 1)$, with $\mathcal{I} = 0.35$, $b_\uparrow = 10$, $b_\downarrow = 0$. The colours of the markers correspond to their coordinates codified in RGB. The dashed-line part of the cube is redundant by symmetry with the solid-line part.	32
5.5	Basins of attraction over the first half of the information partition axis, for $\alpha_1, \alpha_2 \in (0, 1)$, with $\mathcal{I} = 0.2$, $b_\uparrow = 10$, $b_\downarrow = 0$. Each colored cell corresponds in position to the starting point of a single dynamics, and in colour to the coordinates, codified in RGB, of the final point of that dynamics.	32
5.6	Basins of attraction over the first half of the information partition axis, for $\alpha_1, \alpha_2 \in (0, 1)$, with $\mathcal{I} = 0.5$, $b_\uparrow = 10$, $b_\downarrow = 0$	33

Bibliography

- [1] P. A. Kropotkin. *Mutual Aid: A Factor of Evolution*. McClure Phillips, mcclure phillips, new york edition, 1902.
- [2] Garrett Hardin and Garrett Hardin. The tragedy of the commons. *Science*, 1968.
- [3] W. D. Hamilton. The genetical evolution of social behaviour. II. *Journal of Theoretical Biology*, 7(1):17–52, July 1964.
- [4] Jonathan Bendor, Jonathan Bendor, Piotr Świstak, and Piotr Swistak. The evolutionary stability of cooperation. *American Political Science Review*, 1997.
- [5] Gur Yaari and Sorin Solomon. Cooperation Evolution in Random Multiplicative Environments, January 2010.
- [6] Lorenzo Fant, Onofrio Mazzarisi, Emanuele Panizon, and Jacopo Grilli. Stable cooperation emerges in stochastic multiplicative growth. *Physical Review E*, 108(1):L012401, July 2023.
- [7] Ole Peters and Alexander Adamou. The ergodicity solution of the cooperation puzzle. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 380(2227):20200425, May 2022.
- [8] Crispin W. Gardiner. *Handbook of Stochastic Methods: For Physics, Chemistry and the Natural Sciences*. Number 13 in Springer Series in Synergetics. Springer, Berlin Heidelberg, study ed., 2. ed., 6. print edition, 2002.
- [9] Martin A. Nowak. Five Rules for the Evolution of Cooperation. *Science*, 314(5805):1560–1563, December 2006.
- [10] Jordan T Kemp and Luís M A Bettencourt. Learning increases growth and reduces inequality in shared noisy environments. *PNAS Nexus*, 2(4):pgad093, April 2023.
- [11] R. C. Lewontin and D. Cohen. On population growth in a randomly varying environment. *Proceedings of the National Academy of Sciences*, 62(4):1056–1060, April 1969.
- [12] Jordan T Kemp, Adam G Kline, and Luís M A Bettencourt. Information synergy maximizes the growth rate of heterogeneous groups. *PNAS Nexus*, 3(2):pgae072, February 2024.
- [13] O. Peters and W. Klein. Ergodicity Breaking in Geometric Brownian Motion. *Physical Review Letters*, 110(10):100603, March 2013.
- [14] Luís M. A. Bettencourt, Vadas Gintautas, and Michael I. Ham. Identification of Functional Information Subgraphs in Complex Networks. *Physical Review Letters*, 100(23):238701, June 2008.
- [15] Paul L. Williams and Randall D. Beer. Nonnegative Decomposition of Multivariate Information, April 2010.

- [16] J. L. Jr. Kelly. *The Kelly Capital Growth Investment Criterion: Theory and Practice*. World Scietific, Singapore, 2011.
- [17] Elisa Angulo-Cánovas, Ana Bartual, Rocío López-Igual, Ignacio Luque, Nikolai P. Radzinski, Irina Shilova, Maya Anjur-Dietrich, Gema García-Jurado, Bárbara Úbeda, José Antonio González-Reyes, Jesús Díez, Sallie W. Chisholm, José Manuel García-Fernández, and María del Carmen Muñoz-Marín. Direct interaction between marine cyanobacteria mediated by nanotubes. *Science Advances*, 10(21):eadj1539, May 2024.
- [18] Gyanendra P. Dubey and Sigal Ben-Yehuda. Intercellular Nanotubes Mediate Bacterial Communication. *Cell*, 144(4):590–600, February 2011.
- [19] Jean-Philippe Bouchaud. On growth-optimal tax rates and the issue of wealth inequalities. *Journal of Statistical Mechanics: Theory and Experiment*, 2015(11):P11011, November 2015.
- [20] Steven H Strogatz. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. Chapman and Hall/CRC, Boca Raton, 3 edition, January 2024.