On the fitness of informative cues in complex environments

Fabrizio Mafessoni[1]*, Michael Lachmann[2] and Chaitanya S. Gokhale[3]
[1]Department of Evolutionary Genetics,
Max-Planck-Institute for Evolutionary Anthropology
Deutscher Platz 6, 04103 Leipzig, Germany
[2]Santa Fe Institute, Santa Fe, NM, USA
[3]Research Group for Theoretical Models of Eco-evolutionary Dynamics,
Department of Evolutionary Theory, Max Planck Institute for Evolutionary Biology,
August Thienemann Str. 2, 24306, Plön, Germany,
*fabrizio_mafessoni@eva.mpg.de

Abstract

To be able to deal with uncertainty is of primary importance to all organisms. When cues provide information about the state of the environment, organisms can use them to respond flexibly. Thus information can provide fitness advantages. Without environmental cues, an organism can reduce the risks of environmental uncertainty by hedging its bets across different scenarios. Risk mitigation is then possible by adopting a life-history of bet-hedging, either randomly switching between phenotypes (diversifying bet-hedging) or adopting intermediate phenotypes (conservative bet-hedging). Hence, understanding patterns of bet-hedging is necessary in order to quantify the fitness benefit of environmental cues, since it provides a baseline fitness in the absence of informative cues. Quantifying fitness benefits in terms of mutual information reveals deep connections between Darwinian evolution and information theory. However, physiological constraints or complex ecological scenarios often lead to the number of environmental states to exceed that of potential phenotypes, or a single intermediate phenotype is adopted, as in the case of conservative bet-hedging. Incorporating these biological complexities, we generalise the relationship between information theory and Darwinian fitness. Sophisticated bet-hedging strategies - combining diversifying and conservative bet-hedging - can then evolve. We show that, counterintuitively, environmental complexity can reduce, rather than increase, the number of phenotypes that an organism can adopt. In conclusion, we develop an information-theoretic extensible approach for investigating and quantifying fitness in ecological studies.

Keywords: Shannon Information, nonlinear fitness, stochastic switching, inaccessible phenotypes

1 Introduction

Change is a significant constant in the natural world. Organisms have to be plastic enough to ride out the variability in the environment. A potential strategy to cope with this variability is to be plastic by relying on information provided by environmental cues. An informative cue is "a feature of the world, animate or inanimate, which can be used by an animal to guide future actions" [1, 2], or development. Many organisms demonstrate a plastic response to the environment, using environmental cues to modulate their phenotype [3].

However, some of the most basic adaptive mechanisms do not require informative cues. For instance, a wide genetic diversity results in a better response of populations to random environmental fluctuations [4]. Another way can be for an organism to adopt a life history of "bet-hedging". This is possible via adopting a generalist phenotype (conservative bet-hedging) or by randomly switching between different phenotypes, within or between generations (diversifying bet-hedging) [5, 6, 7, 8]. Hereafter we use bet-hedging as a synonym for diversifying bet-hedging.

Even without genetic variation in the population, bet-hedging enables coping with uncertain environments [9, 10, 11]: when all individuals in a population experience the same environmental state, developmental variability in the phenotypes of an individual's offspring will ensure that at least part of its progeny will develop a proper phenotype for the current environmental state. The theory of bet-hedging has had a long history [12, 13, 14, 15] and instances of this phenomenon can be observed in nature [16, 17, 18]. The bet-hedging theory hinges on the trade-off between maximizing short term fitness and reducing the adverse long term effects of its variability due to the stochasticity of the environment. Interestingly, even when cues are present but do not provide full information, bet-hedging can still occur as a consequence of their uncertainty [19].

Previous studies quantified the fitness values of informative cues by comparing the fitness of plastic strategies relying on cues and of bet-hedging strategies, not relying on any cue. These studies pointed to a tight relationship between evolution and information. In simple cases where the number of phenotypes equals that of environmental states, bits can quantify the evolutionary benefit of an informative cue [20]. While information theory provides a promising framework to study evolution, relatively few studies used this framework to study ecological and evolutionary phenomena [21]. In this study, we extend the information theory approach previously described by Bergstrom and Lachmann [20, 22] to examine the effects of more complex scenarios. In nature, often environmental complexity exceeds physiological and evolutionary flexibility, and the number of possible environmental states might exceed that of phenotypes. For example, the phenotypes might lie far from each other in the fitness landscape to be accessible, or physiological constrains in an organism might limit the variety of phenotypes it can develop. Besides, one of the available phenotypes may perform better in more than one environment. In these cases, as environmental complexity increases, the optimal strategy is more likely the one in which the same phenotypes are adopted in response to different environmental states, i.e. conservative bet-hedging occurs along with diversifying bet-hedging. These instances, in which the number of environmental states exceeds the number of phenotypes, are the focus of this study. We show that in these cases, vibrant patterns of bet-hedging emerge and that, counterintuitively, the potential fitness value of informative cues can decrease.

We first begin by reviewing the connection between Shannon and Gould information as per [20, 22], by following the calculations from Donaldson-Matasci *et al.* [22] where an organism can have as many phenotypes as the number of environmental states (represented by a square matrix). Then we generalize this analysis to any number of phenotypes and environmental states. Inequality between the number of phenotypes and environmental states (represented by a non-square matrix) is also possible. We explore representative cases of asymmetric scenarios, characterized by simple probability distributions describing the occurrence of different environmental states. In these cases, the best bet-hedging strategy depends on the probability of the environmental states in a non-linear fashion. We discuss some examples, among which those of organisms adapted only to a range of environmental conditions, for which often increased environmental uncertainty is present at the borders of their distribution. We show that relatively more complex scenarios, the fitness value of the informative cues is less than the mutual information between the cue and the environment. We generalize this observation to any asymmetric scenarios, showing that under the more natural asymmetric conditions of complex environments, a lower boundary to the fitness benefit of an informative cue is expected.

In biology, a direct method to quantify the impact of a cue is to compare the fitness when knowing and not knowing the cue. The difference in the fitness of the strategy adopted when knowing the cue and not knowing the cue is the value of the cue measured in units of fitness gained [23, 24]. Examples where such cues are essential range from the germination in seeds and diapause in insects, to offspring clutch size control by parents. If a cue provides complete information about the predators' presence, then the animal can opt more likely for a proper course of action. This measure is **Gould information**.

A classical method to quantify the value of an environmental cue is **Shannon Information**. Consider a fair coin. The two possible outcomes "heads" (1) and "tails" (0) are both likely with equal probabilities ($p_0 = p_1 = 0.5$). Hence the amount of surprise which we have when we know the result of a certain coin toss event E is simply, $H(E) = -p_0 \log_2 p_0 - p_1 \log_2 p_1 = 1$. This is the Shannon entropy as defined in classical information theory [25, 26]. Traditional approaches in statistical physics, communication, engineering and related fields make use of the concept of 'mutual information'. The amount by which a cue (C) reduces the uncertainty about the environment (E) is defined as mutual information ($I(E;C)$). It is measured in terms of entropy as,

$$I(E;C) = H(E) - H(E|C) \tag{1}$$

where $H(E) = -\sum_e p_e \log p_e$ is the 'entropy' of the random variable E denoting the environment. Similarly $H(E|C) = -\sum_c p_c \sum_e p(e|c) \log p(e|c)$ is the entropy when the cue has been received. Here p_c is the probability of observing the cue c and the probability that the environment is in state e when cue c is observed is given by $p(e|c)$. If a cue is not related to the environment, then the entropy remains unchanged ($H(E|C) = H(E)$) and the mutual information between the cue and the environment is zero. However for a perfect cue this probability is $p(e|c) = 1$ and hence $H(E|C) = 0$. Thus the cue reveals the environment entirely, and the mutual information is precisely equal to the entropy of the system $I(E;C) = H(E)$.

Strong links exist between the two measures of the value of informative cues [20, 22]. Studies show that if the environmental cue is flawlessly informative then for an intermediate probability of an adverse event occurring, the best approach is to bet-hedge. This region of probability space in which bet-hedging occurs is a function of the Shannon entropy, and the fitness value of information is bounded above by the Shannon entropy [20]. Thus an intimate connection exists between the classical information-theoretic approach and the biologically intuitive Gould information approach.

Throughout the analyses we adopt a geometric mean approach [27, 28, 22]. Whereas the arithmetic-mean does not capture the effects of environmental variance on fitness, the multiplicative nature of geometric-mean efficiently describes the effects on the growth rate as long as there are no interactions between the phenotypes, i.e. where the average fitness is frequency independent [12, 29, 30, 31]. Indeed, in a variable environment the allele having the higher geometric mean takes over the population [32, 33, 34]. Therefore short term fitness (usually considered in the frequency-dependent analysis, e.g. in evolutionary game theory) might not be the most useful statistic in case of environmental variability. Instead, a geometric mean approach can be fruitfully adopted in these cases, elucidating the role of bet-hedging strategies in minimizing fitness variability in the absence of informative cues [20, 22, 8].

2 Model and Results

2.1 Symmetric case: two environmental states with two phenotypes

Traditionally bet-hedging models have focused on the example of annual plants. We use it here as well, presenting a representative case using simple fictitious interaction matrices. A detailed description of possible ecological interpretations follows in the discussion.

A year could be wet e_1 or dry e_2. Seeds of desert annuals are better off staying dormant over dry years and do better by germinating in wet years. This relationship can be represented in the form of an interaction matrix as follows,

$$\begin{array}{cc} & \begin{array}{cc} e_1 & e_2 \end{array} \\ \begin{array}{l} \text{Germinate} \\ \text{Dormant} \end{array} & \begin{pmatrix} 5 & 1 \\ 2 & 2 \end{pmatrix}, \end{array} \tag{2}$$

The probability that a flash flood occurs and a year is wet is p, while the year is dry with the probability $(1-p)$. The gain of the two phenotypes averaged over the two environments are thus given by,

$$\begin{aligned} f_G &= 5\,p + 1\,(1-p) \\ f_D &= 2\,p + 2\,(1-p) \end{aligned} \tag{3}$$

These are the average fitnesses of the two phenotypes. The internal equilibrium of this system is at $p^* = 1/4$, where the fitnesses of the two strategies are equal. In the short run organisms will maximize their expected fitness by employing a strategy that maximizes its single generation expected fitness, $F(E) = max[f_D, f_G]$ where E is a random variable representing the state of the environment (Fig. 1). Hence if the probability of a flash flood is greater than p^* it is better to germinate else staying dormant is a safe bet.

$$F(E) = \begin{cases} f_G & \text{for } p > p^* \text{ always Germinate} \\ f_D & \text{for } p < p^* \text{ always Dormant.} \end{cases} \tag{4}$$

In order to quantify the reduction in fitness due to environmental uncertainty, we can now suppose that the organisms might receive a cue C, which indicates the environmental state precisely. Therefore, then there is no confusion over choosing to germinate or not regardless of p. In this case the single generational expected fitness will be, $F(E|C) = 5p + 2(1-p)$. The value of the cue is then is the difference between the fitness with the cue and without the cue,

$$\Delta F(E; C) = F(E|C) - F(E) \tag{5}$$

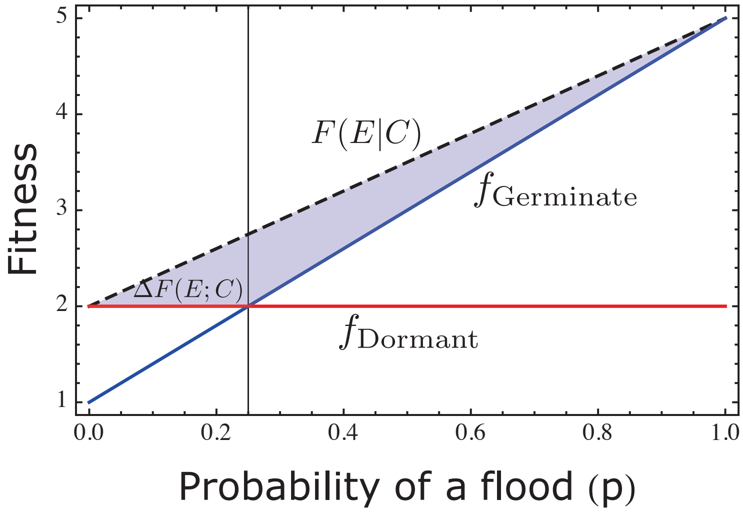

Figure 1: **Average fitnesses of the Germinate and remain Dormant strategies as a function of the probability of a flood p.** The internal equilibrium is given by the vertical line at $p^* = 1/4$. For $p < p^*$ it pays to remain dormant while for $p > p^*$ it is better to germinate. Knowing the state of the environment exactly from the cue a seed can obtain $F(E|C)$. Hence the fitness benefit due to the cue is given by the shaded area.

This is the result which we expect if the organism maximizes the single generational expected fitness. However, in a temporally varying environment, the phenotype that is most likely to fix over the long term is the one with the highest expected long-term growth rate. To approach this question, we define the frequencies with which each phenotype is adopted in a given season as x (Germinate) and $1 - x$ (Dormant) respectively. The mean log fitness can estimate the long-term growth rate of a given strategy, or, equivalently, the log of the geometric mean fitness [12, 35, 36, 19]. The expected long term growth rate is then given by,

$$g(x) = p \log[env(e_1, x)] + (1 - p) \log[env(e_2, x)] \tag{6}$$

where the fitnesses in the two environments are given by,

$$\begin{aligned} env(e_1, x) &= 5x + 2(1 - x) \\ env(e_2, x) &= x + 2(1 - x). \end{aligned} \tag{7}$$

Maximizing the expected long term growth rate results in,

$$x^*(p) = \begin{cases} 0 & \text{for } p \leq 1/4 \text{ (Dormant)} \\ \frac{2}{3}(4p-1) & \text{for } 1/4 < p < 5/8 \text{ (Bet Hedging)} \\ 1 & \text{for } 5/8 < p \text{ (Germinate).} \end{cases} \tag{8}$$

We can therefore identify a bet-hedging region, namely a region in probability space where probabilistic switching between the phenotypes occurs. Substituting this result back in the growth rates we get,

$$G(E) = \begin{cases} \log(2) & \text{for } p \leq 1/4 \text{ Dormant} \\ (1-p)\log(8/3) + p\log(8) - H(E) & \text{for } 1/4 < p < 5/8 \text{ Bet Hedging} \\ (1-p)\log(5) & \text{for } 5/8 < p \text{ Germinate.} \end{cases} \tag{9}$$

where $H(E) = -p\log(p) - (1-p)\log(1-p)$ is the entropy of the random variable E. Making use of a perfectly informative cue the growth rate can be given by, $G(E|C) = p\log(5) + (1-p)\log(2)$. Hence now the value of the cue is the difference between the growth rate with the cue and the one without i.e. $G(E|C) - G(E)$ illustrated in Fig. 2,

$$\Delta G(E;C) = \begin{cases} p\log(5/2) & \text{for } p \leq 1/4 \text{ Dormant} \\ H(E) + (1-p)\log(6/8) + p\log(5/8) & \text{for } 1/4 < p < 5/8 \text{ Bet Hedging} \\ (1-p)\log(2) & \text{for } 5/8 < p \text{ Germinate.} \end{cases} \tag{10}$$

Previous studies show that $\Delta G(E;C)$ peaks within the region of bet-hedging and is bounded by the mutual information between the cue and the environment [**?**]. As in this case where the cue is perfect this mutual information is simply the Shannon entropy of the environment $H(E)$.

2.2 Examples with asymmetry in strategies and environments

The environment can vary in time and space. Generally, an organism is adapted only to a limited range within an environmental spectrum. Thus including intermediate environmental states can better represent relevant environmental variability. For example, the risk of meeting a predator or the occurrence of certain climatic events can vary along both large or short ecological gradients. Drought or extreme flooding generally limit the growth of the seedlings while a moderate amount of rain is preferable, and precipitations might vary along with geographic distances.

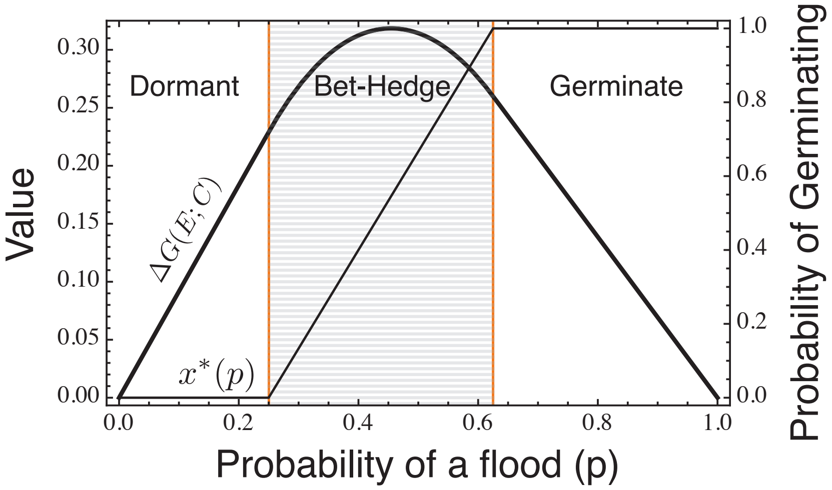

Figure 2: **The value of the cue and the probability to germinate as function of the probability of environmental state e_1.** The value of the cue $\Delta G(E;C)$ increases until it peaks within the area of bet-hedging. The bounds of the area of bet-hedging can be calculated analytically Eq. (8). These bounds delineate the transitions in the effective probability of germinating $x^*(p)$ which shifts from 0 (Dormancy) to 1 (Germination) via a linear increase within the area of bet-hedging.

Here we consider the different possible fitness effects of multiple environmental states for the example of seed dormancy. For simplicity we take into account the possibility that two consecutive storms occur within one season (e_1), only one (e_2), or none, leading to a dry year (e_3). These events occur respectively with probabilities p^2, $2(1-p)p$ and $(1-p)^2$. In principle, we can assume a complicated function with two variables instead of only p. However, as we show below, even this simple parameterisation, is capable of resulting in intricate bet-hedging patterns, which is sufficient to make our point.

2.2.1 Adaptation to intermediates

Assuming that a single flood might provide the necessary humidity but two consecutive ones might damage the seedlings we can write down the following interaction matrix,

$$\begin{array}{c c} & \begin{array}{ccc} e_1 & e_2 & e_3 \end{array} \\ \begin{array}{l} \text{Germinate} \\ \text{Dormant} \end{array} & \begin{pmatrix} 1 & 4 & 1 \\ 2 & 2 & 2 \end{pmatrix}, \end{array} \tag{11}$$

We have a non-monotonic behaviour of the value of the cue (Appendix, Fig. 3 (a)). However, the satisfactory probability of playing the 'Always Germinate' strategy never materialises. Instead, the seeds do best remaining dormant close to the extreme conditions and only germinating about half the time at most when hedging their bets. The bet-hedging region in this particular case explores the mixed phenotype space between the pure phenotypes.

2.2.2 Adaptation to extremes: Multiplicity of bet-hedging.

Contrary to the previous example, it might also be possible that some plants do better in extremes rather than in common environmental conditions. For example, annual pioneer plants can be easily outcompeted by others in intermediate environments (e_2). However at environmental extremes (e_1 and e_2) their seeds have an advantage.

$$\begin{array}{c c} & \begin{array}{ccc} e_1 & e_2 & e_3 \end{array} \\ \begin{array}{l} \text{Germinate} \\ \text{Dormant} \end{array} & \begin{pmatrix} 4 & 1 & 6 \\ 3 & 3 & 3 \end{pmatrix}, \end{array} \tag{12}$$

In such a case, we see that the probability of germinating decreases after a specific threshold value of p Fig. 3 (b). The non-linear decrease is up to the pure strategy of dormancy. After a specific threshold value of p, it is better to hedge bets with a non-linear increasing probability of germination reaching the 'Always Germinate' extreme. The value of the cue peaks locally in the bet-hedging regions (Appendix and Fig. 3 (b) bold curve).

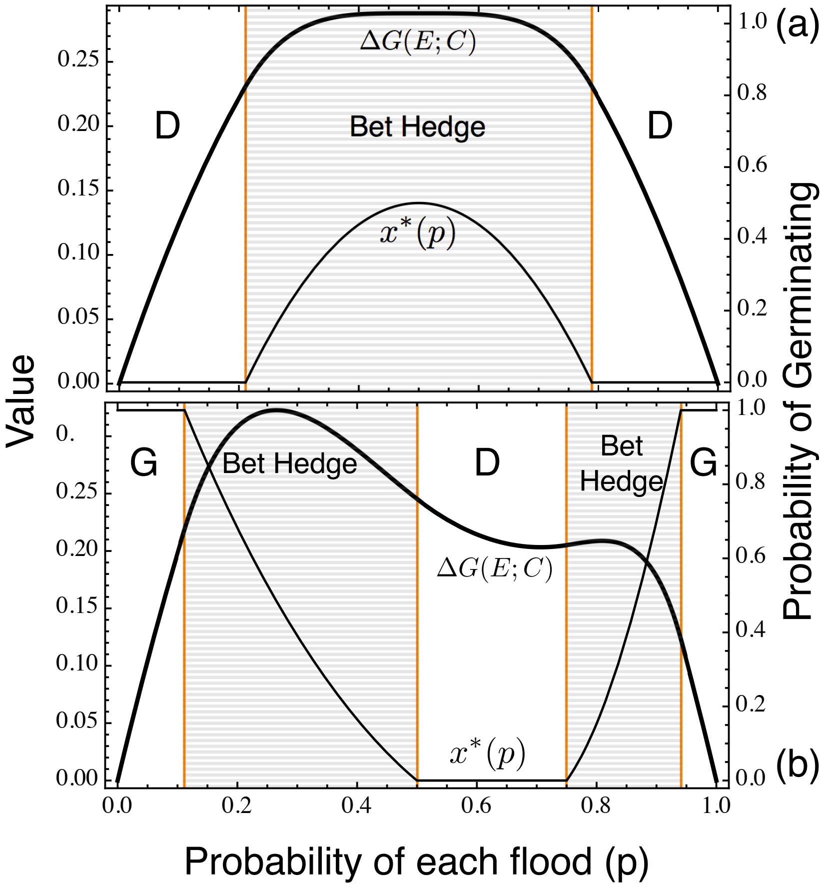

Figure 3: **The value of the cue and the probability of germinating as function of the probability of a single flash flood p.** Panel (a) is for the example where the intermediate environmental state is favourable for germination. In this case the value of the cue peaks in the area of bet-hedging. The probability of germinating ($x^*(p)$) increases in frequency but just explores the intermediate regime without ever reaching pure Germination. Panel (b) explores the case where germinating in extreme environments is favourable. This case results in two regions of bet-hedging where again the value of the cue ($\Delta G(E;C)$) peaks locally. The probability of germinating goes from 0 Always Dormant) all the way to the other pure phenotype 1 (Always Germinating) non-linearly hedging its bets along. It changes again in the second region of bet-hedging decreasing to 0 (Always Dormant) where it remains until p reaches 1. For both the cases, the bounds of the area of bet-hedging can be calculated analytically (ESM). These bounds delineate the transitions in the effective probability of germinating $x^*(p)$.

Figure 4: **Bet hedging regions in the probability space for the illustrative examples.** For general probabilities of the three environmental states obeying $e_1+e_2+e_3 = 1$, we can plot the effective probability of dormancy x^* for the three illustrative examples. The solid parabolas are the quadratic binomial curves used to describe the examples. For (a) we see a hump in the probability of germinating but the pure Germination strategy is not reached. In (b) the probability of germinating decreases and reaches the pure Dormancy state but then it reduces to pure Germination as the system comes closer to the e_1 vertex.

2.3 Environmental gradients in the simplex

The probability distribution considered in the previous examples traces a quadratic binomial cline. Such parabolic clines are typical of those distributions in which e_2 is an intermediate environmental state between the others. Thus the distribution of the environments has a direct impact on the bet-hedging regions. For skewed non-linear distributions, it is possible to have multiple regions of bet-hedging. We can extend this case to an arbitrary probability distribution of the environmental states. Assuming that the three states occur with probability p_{e_1}, p_{e_2} and p_{e_3} where they all sum up to unity ($p_{e_1} + p_{e_2} + p_{e_3} = 1$). Hence now bet-hedging can occur in a simplex defined by these three probabilities as shown in Figure 4.

We explore the evolution of bet-hedging in the simplex of environmental probabilities by using stochastic simulations for finite populations, in order to relax the assumption of populations as unbounded and growing at an exponential rate used for the geometric mean approach. We consider the interaction matrix (11), and performed Wright-Fisher simulations considering a population size of 1000 individuals, evolved for 1000 generations. Analyzing 100 such realizations in Fig. 5 (a) we show the probability of expressing a phenotype, germination, for general environmental distributions. The gradient of the probability of germination was obtained by interpolating a grid of 10000 points for the whole simplex.

In Fig. 5 (right panel) we show further simulations performed for the points marked on the clines in Fig. 5 (simplex). The clines represent the particular probability distributions for the environmental states. For instance, the previous examples considered a second-order binomial indicated by the grey dashed cline, where e_1 occurs with probability p^2 and e_3 with $(1-p)^2$. The second black cline indicates the environmental states distributed according to a fourth-order binomial. Thus e_1 occurs with probability p^4 while e_3 with probability $(1-p)^4$. The state e_2 occurs with the complementary probability. Depending on the eccentricity of the clines we can get one or two regions of bet-hedging. We see the hump shape of the probability of germinating as predicted by the infinite population size case in Fig. 3 (a) also recovered via

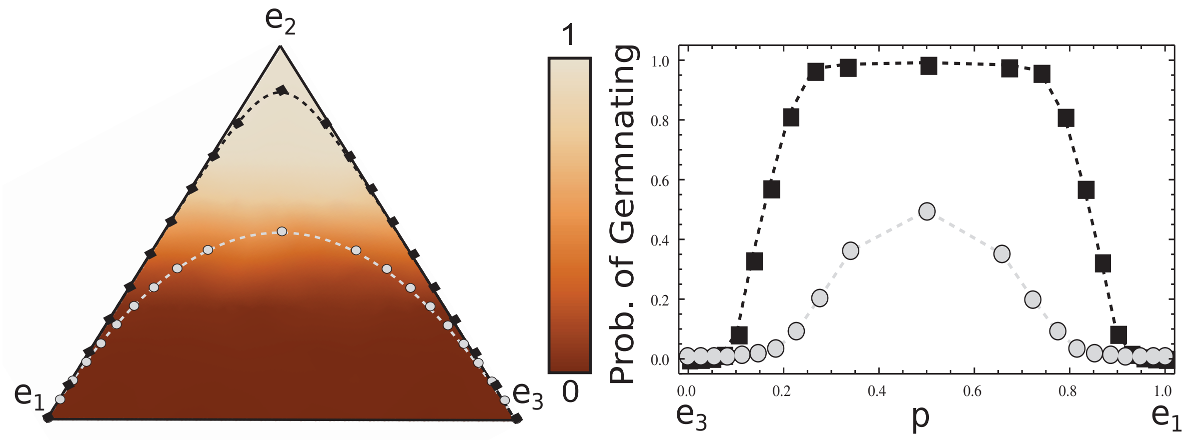

Figure 5: **Stochastic simulations of the mean probability of germinating as a function of particular distribution s of environmental states.** Wright-Fisher simulations were performed with a population size, N of 1000. After N generations the mean probability of germination was calculated. The results were then interpolated to generate the density ternary plot (a). The marks along the clines represent the probability distributions in the simplex space used for the simulations in (b), shown along the clines. Grey dots indicate the second order binomial clines also used for the examples. Black squares indicates instead a fourth order binomial, where the mixed term are assigned to the intermediate state e_2.

stochastic simulations. For the higher-order cline, the simulation results in the ‘Germinating’ being almost fixed as is the predicted case on the eccentricity of the cline (Fig. 5 simplex). Hence not only the interaction matrix but also the particular distribution of the environments influences the observed patterns of bet-hedging.

Note that we chose to explore parabolic clines for simplicity and to represent environmental gradient on a single axis. However, exploring indipendently the probabilities of different environments in the simplex can be relevant for more complex environments or when different events interact with each other. We provide some examples in an interactive R notebook on Github

2.4 The fitness value of information in presence of constraints

In all previous examples, the fitness value of a cue, $\Delta G(E, C)$, is calculated with an information theory approach, despite the non-linear interactions between environment and phenotypes. Donaldson-Matasci et al. [22] have shown that in the symmetric case, its value is bounded by the mutual information between the distribution of an informative cue C and that of the environment E, $I(E;C)$. Self-information, denoted as $I(E;E)$, is equal to $H(E)$. Hence, for an entirely informative cue, the fitness value is bound by the entropy $H(E)$. Does this boundary always hold even in the asymmetric case? If yes, is it possible to draw a lower boundary? Let us consider environmental state e_i with i in $1 \dots n$ and phenotypes ϕ_j with $j = 1, \dots, m$. Previous models [20, 22] explored cases in which the optimal response to each different environmental state e_i is a phenotype ϕ_i. In this case, a square n x n interaction matrix defines the payoff for each phenotype ϕ_i in each different environmental state e_i [22], occurring with probability p_i. The optimal matches between the environmental state and phenotype appear along the diagonal, the corresponding gains being $a_{i,i}$ for every $i = \{1, 2, \dots n\}$.

$$\begin{array}{c|ccccc} & e_1 & \dots & e_i & \dots & e_n \\ \hline \phi_1 & a_{1,1} & \dots & a_{1,i} & \dots & a_{1,n} \\ \hline \vdots & \dots & \dots & \dots & \dots & \dots \\ \hline \phi_n & a_{n,1} & \dots & a_{n,i} & \dots & a_{n,n} \end{array} \tag{13}$$

However, the assumption of an optimal, accessible phenotype for each environmental state might not hold, particularly in the context of the evolution to multiple environments, where a single phenotype might represent the best option in response to multiple environmental states.

Here we extend the information-theoretic approach developed in Bergstrom et al. [20] to estimate the fitness value of a cue but applied to such cases.

To do this, we start by the proportional betting case, in which the best phenotype response for environmental state e_i provides a payoff $a_{i,i} = a_i$ while all the other phenotypes are fatal ($a_{i,j} = 0$ for $i \neq j$)[22], resulting in a diagonal payoff matrix. In general, the best bet-hedging strategy x, employing a phenotype ϕ_j with probability x_j, can be determined by solving the Lagrangian [26] of the long term growth rate with constraint $(\sum_i x_i = 1)$:

$$\sum_i p_i \,\log \sum_j (a_{j,i}\, x_j) - \lambda(\sum_i x_i - 1). \tag{14}$$

In the proportional betting case, each phenotype is adopted with a probability equal to the occurrence of the environmental state where the payoff is non-zero, denoted as a_i. Hence, in the symmetric case, when for each environmental state, a different optimal phenotype exists:

$$G(E)_n = \sum_{i=1}^{n} p_i \log p_i a_i = \sum_{i=1}^{n} p_i \log a_i - H(E). \tag{15}$$

We study the asymmetric case with the number of environments n larger than the number of phenotypes m. To this aim, we can extend a diagonal payoff matrix m x m to include additional columns representing the payoffs of the different phenotypes to environments $j >$

m. Hence, we investigate payoff matrices for which $a_{i,j} \neq 0$ only when $i = j$ or for $a_{m,j}$ and $j \geq m$. We write $a_{i,i} = a_i$ for $i < m$, and $a_{m,i} = b_i$ for $i \geq m$. We can the write:

$$\begin{aligned}
G(E)_m &= \sum_{i=1}^{n} p_i \log\left(\sum_{j=1}^{m} a_{j,i} x_j\right) = \\
&= \sum_{i=1}^{m-1} p_i \log\left(\sum_{j=1}^{m} a_{j,i} x_j\right) + \sum_{i=m}^{n} p_i \log\left(\sum_{j} a_{j,i} x_j\right) = \\
&= \sum_{i=1}^{m-1} p_i \log\left(a_i x_i\right) + \sum_{i=m}^{n} p_i \log\left(a_{m,i} x_m\right) = \\
&= \sum_{i=1}^{m-1} p_i \log x_i + \sum_{i=m}^{n} p_i \log x_m + \sum_{i=1}^{n} p_i \log a_i \qquad (16)
\end{aligned}$$

Which gives a maximum at

$$\begin{aligned}
G(E)_m &= \sum_{i=1}^{m-1} p_i \log p_i + \sum_{i=m}^{n} p_i \log \sum_{j=m}^{n} p_i + \qquad (17) \\
&+ \sum_{i=1}^{m-1} p_i \log a_i + \sum_{i=1}^{m-1} p_i \log b_i \qquad (18)
\end{aligned}$$

Therefore:

$$\begin{aligned}
&G(E)_m - G(E)_n = \\
&= \sum_{i}^{m-1} p_i \log p_i + \sum_{i=m}^{n} p_i \log\left(\sum_{j=m}^{n} p_j\right) - \sum_{i}^{n} p_i \log p_i - \sum_{i=m}^{n} p_i \log a_i/b_i = \\
&= \sum_{i}^{m-1} p_i \log \frac{p_i}{p_i} + \sum_{i=m}^{n} p_i \log \frac{\sum_{j=m}^{n} p_j}{p_i} - \sum_{i=m}^{n} p_i \log a_i/b_i
\end{aligned} \tag{19}$$

When the cue C provides full information about the environment, the right phenotype will be chosen in each case, with payoff a_i, except for environmental states $i > m$, when one strategy cannot develop phenotype ϕ_i. Therefore just as in the symmetric diagonal case, $G(E|C)_n = \sum_{i=1}^{n} p_i \log a_i$, while $G(E|C)_m = G(E|C)_n - \sum_{i=m}^{n} p_i \log a_i/b_i$, and

$$\Delta G(E;C)_n - \Delta G(E;C)_m = \sum_{i=m}^{n} p_i \log \frac{\sum_{j=m}^{n} p_j}{p_i} \tag{20}$$

This is the Kulback-Leibler divergence between the optimal strategy when all phenotypes $1 \ldots n$ are available, and used with frequency p_i, and when no optimal specific types for environments $m \ldots n$ are available, and for all those phenotype m is used with the sum frequency $\sum_{m}^{n} p_i$. The expression is independent of the payoffs and always positive, confirming that the mutual information bound for the fitness value of an informative cue exists for the asymmetric case as well. Furthermore, this quantity is part of the conditional entropy of the environment given a fully informative cue, reflecting the residual uncertainty of a cue unable to discriminate between environmental states for which a single phenotype is optimal. Since $I(E_n|C) - H(E|C_{i\geq m}) = I(E_m|C)$, this indicates that in the presence of a generalist phenotype, the fitness of an informative cue is further decreased by the amount of information necessary to differentiate between the equally paying environmental states. The decrease is equal to the mutual information of the environment with a cue informative only on environmental states for which different phenotypes are optimal.

These results, shown for the proportional bet-hedging case for simplicity, are also valid in the general case of non-completely specific phenotypes, in which the payoff of the least advantageous phenotype-environmental state combinations may provide non-zero payoffs. In fact, following Donaldson-Matasci et al.,2008 [37], note that the fitness profile of these phenotypes in different environments can be represented as a mixture of specialized phenotypes with zero-payoffs for disadvantageous phenotype-environmental state combinations, i.e. as in the proportional bet-hedging case. Thus, also bet-hedging strategies employing these non-completely specialized phenotypes can be seen in turn as mixtures of completely specialized phenotypes. Hence, even in the more complex cases, we see that the value of the cue is bound by the mutual information between the cue and the environment (Fig. 6 and Appendix). Since we are dealing with perfect cues, this mutual information is simply the entropy of the environment $H(E) = -\sum_{i=1}^{3} p_{e_i} \log p_{e_i}$ (equal to $H(E_m)$) (yellow surface) in the asymmetric case).

2.5 Limits of bet-hedging in increasingly complex environments

So far we focused on cases in which the number of potential phenotypes m is lower than that of environmental states n, i.e. payoffs matrices with $m < n$. We now describe why these cases are relevant and likely universal in nature, by showing that even when more potential phenotypes are available and $m > n$, evolution leads to organisms using only a reduced set of phenotypes. Note that non-trivial instances of $m > n$ matrices correspond to cases in which at least a phenotype provides a non-zero payoff for multiple environmental states, thus offering an opportunity for conservative bet-hedging. An extreme case is that of a phenotype that provides the same payoff for all environmental states. An alternative case is that of an intermediate phenotype that can be adopted in spite of two or more optimal phenotypes for specific environmental states. Hence, our question corresponds to investigating when conservative bet-hedging strategies evolve over diversifying bet-hedging ones, i.e. when a matrix $m \times n$ reduces effectively to a matrix with a smaller number of rows/phenotypes. To do this

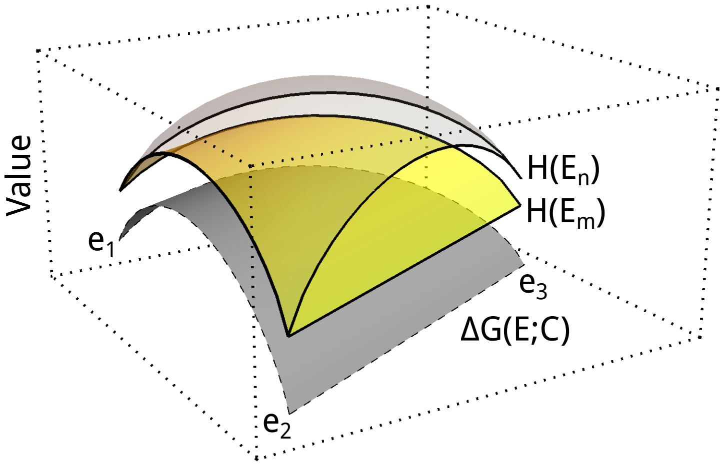

Figure 6: **Boundedness of the value of information by the entropy of the environment.** Considering all possible distributions, the parameter space of the possible environmental states is then defined on a simplex given by p_1, p_2 and p_3 where the vertices are the pure environmental states e_1, e_2 and e_3 respectively. The value of the cue $\Delta G(E;C)$ over such a space (gray triangle) is bounded by the mutual information between the environment and the cue (transparent triangle), that in the asymmetric case decreases to $H(E_m)$ (yellow surface). We demonstrate this for the most complicated case described by the interaction matrix (12).

we examine a matrix:

$$\begin{array}{c|ccccc} & e_1 & \dots & e_i & \dots & e_n \\ \hline \phi_1 & a_{1,1} & \dots & a_{1,i} & \dots & a_{1,n} \\ \hline \vdots & \dots & \dots & \dots & \dots & \dots \\ \hline \phi_n & a_{n,1} & \dots & a_{n,i} & \dots & a_{n,n} \\ \hline \phi_{n+1} & b_{n+1,1} & \dots & b_{n+1,i} & \dots & b_{m+1,n} \\ \hline \vdots & \dots & \dots & \dots & \dots & \dots \\ \hline \phi_m & b_{m,1} & \dots & b_{m,i} & \dots & b_{m,n} \\ \hline \end{array}, \tag{21}$$

where for each environmental state i we arranged the phenotype providing the highest payoff $a_{i,i}$ as phenotype i. Further potential phenotypes are denoted as ϕ_j with $j > n$, and by construction, payoffs $b_{j,i} \leq a_{i,i}$. Our previous analysis presents a simple method for investigating under which environmental circumstances such asymmetries might evolve. Consider a strategy $G(E)_m$ adopting a generalist phenotype ϕ_j with payoffs $b_{j,k} > a_{i,k}$ rather than the specialized phenotypes ϕ_k with payoffs $a_{k,k}$ for $k > i$, as strategy $G(E)_n$. By recalling Eq. (19)-(20), we can see that $G(E)_m$ is advantageous if

$$G(E)_m - G(E)_m = H(E|C_{i \geq m}) - \sum_{i=m}^{n} p_i \log a_i/b_i > 0 \tag{22}$$

indicating that the maximum decrease in payoff of the generalist strategy is bound by the entropy of the system that is *removed* by adopting it. This simple relationship helps us to illustrate the instances in which generalist phenotypes are employed. To do this, we devise an example in which specialized, intermediate and generalist phenotypes evolve, following the general payoff matrix described in Eq. (21). We consider n equally frequent environmental states, subdivided in n_k groups of size k, biologically representing environmental states with similar features. Intermediate phenotypes allow to respond to any of the similar environmental states, but with maximum payoff lower than the specialized phenotypes, i.e. for each of the n_k groups, an intermediate phenotype $\phi_{i \geq n}$ exists, providing payoff $b_{n_k,i}$ as in Eq.21. Also, a perfect generalist phenotype exists, providing the same payoff c in all environmental states.

By manipulating the number of environmental states, while keeping constant the size of the groups, the total entropy ($H(E_n)$) of the system increases, while the entropy within the groups ($H(E(_k))$) is not affected (Fig.7a). Furthermore, the conditional entropy of the environmental states with the same intermediate phenotype is unaffected; intermediate phenotypes are not favoured over specialized ones as the total entropy increases. Whether intermediate phenotypes evolve or not depends exclusively on the ratio of the payoffs. On the other hand, generalist phenotypes (i.e. conservative bet-hedging strategies) eventually become advantageous as entropy increases.

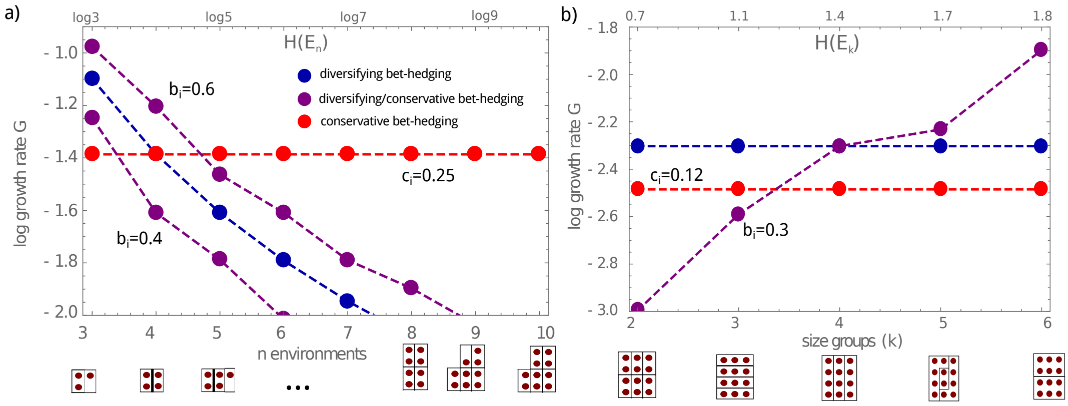

Figure 7: **Limits of bet-hedging in increasingly complex environments.** Long term log-growth rates for different bet-hedging strategies for environments with an increasing number of uniformly distributed environmental states (a) or an increasing size of the partitions of environmental states for which an intermediate generalist phenotype exist. These partitions are represented under the plots with environmental states (brown circles) grouped on the basis of a potential suitable generalist phenotype (represented as black rectangles). We assume a square environment payoff matrix in which the optimal payoff for any environmental state give payoff 1 ($a_{i,i} = 1$) and 0 in all non matching cases (and $a_{i \neq j} = 0$), leading to a fully diversifying bet-hedging strategy (blue dots). We also show the log-growth rate for a fully generalist strategy with fixed payoff in all environmental states c (red dots) and partially bet-hedging strategies that instead of hedging their bets on the optimal phenotypes, they hedge more generalist phenotypes matching the different partitions, providing payoff b in those and 0 in the rest. a) Partition size is fixed to two.

By manipulating the size of the groups, only the conditional entropy, $H(E|C_{i\geq m})$ (here $H(E_k)$), of the groups increases (Fig.7b). In this case, the growth rates of pure diversifying and conservative bet-hedging remain constant. However, strategies bet-hedging between intermediate phenotypes become more advantageous as $H(E_k)$ increases.

To summarize, in both cases, strategies employing generalist phenotypes evolve as the uncertainty of the system increases. In turn, the growth rate and the fitness value of information can be described using asymmetric matrices.

3 Discussion

Organisms facing complex and variable environments often evolve complex adaptive mechanisms to acquire information about the environment in the form of informative cues. What is the fitness values of these cues? What is the fitness value of these complex adaptations? Answers to these vital questions remain elusive in natural populations. A significant advantage of information theory is that it provides a theoretical boundary to these quantities, and could aid quantify the fitness benefits based on the uncertainty of the environment, thus informing empirical as well as theoretical research. However, its application has been so far limited [21], possibly because the models usually focused on simple systems for tractability.

Natural systems, on the other hand, can be exceedingly complex. Organisms have sophisticated bet-hedging strategies to deal with environmental uncertainty without relying on informative cues. In these cases, organisms do not need to respond flexibly to the environment. Spreading the risk of facing adverse environmental conditions by betting on diverse or intermediate phenotypes can work. This mechanism can explain several biological phenomena, from intergenerational phenotypic stochastic switching to animal personalities [38], although hard evidence for bet-hedging in the animal world has been so far elusive [18]. While it is possible to study bet-hedging experimentally in microbes [6, 7], the evolutionary mechanisms underpinning the use of informative cues is still elusive. Different levels of environmental unpredictability present along environmental or geographical clines exacerbate the complexity of bet-hedging and informative cues. Natural populations could then potentially display non-trivial bet-hedging patterns. We illustrate this through the example of organisms adapted only to intermediate states of environmental gradients such as temperature or salinity. Intermediate levels of such gradients are taken as discrete environmental states. For example, seasonal plants adapted to intermediate water levels: the total absence of flash floods might lead to a harsh dry year, as well as two consecutive floods could damage the growth of seedlings. In this example (Fig. 3) we show two cases in which bet-hedging occurs only within confined regions of the parameter space, i.e. regions with intermediate probabilities of rain, surrounded by more stable regions where only one phenotype is adopted. These hypothetical clines, described in Figures 5, could be geographic transects going from environmental extremes such as from desert to rain-forests. For example, dormancy is advantageous in response to multiple extreme events (e.g. drought, fire [39], extreme cold) and occurring from alpine [40] to wetland [41] and desert plants. We suggest that such studies might highlight interesting patterns of bet-hedging along environmental gradients, primarily if a species exist across such environmental transect [42].

We show that universally, the fitness benefit of an informative cue is bound by the mutual information between the cue and the environment. Besides, when the variety of phenotypes that can be adopted by phenotypic switching strategies and plasticity is limited, the maximum potential fitness benefit of an informative cue is further reduced. Precisely, the reduction is due to the conditional entropy between the environment and a cue informative about the inaccessible phenotypes, i.e. $H(E|C_{i\geq m})$, where a single phenotype has the highest environmental states, $i \geq m$. This result implies that plasticity is less advantageous when physiological constraints limit the phenotypes that can be adopted. This result is intriguing because counter-intuitively the number of phenotypes that can be adopted can decrease - and in-turn, the fitness benefit of information can decrease, rather than increase - as environmental uncertainty increases. In practice, under high environmental variation, an organism might decrease the amount of explored phenotypes explored through bet-hedging, further reducing the benefit of relying on informative cues.

Note that information theory provides insights into many possible scenarios in which the relationship between bet-hedging and cues is complicated. For example, when cues are unreliable and idiosyncratic (i.e. each individual - a different cue) relying on cues can be used as another means towards bet-hedging [19]. Hence, we advocate that the collection of more empirical data on the environmental uncertainty of species that live on environmental gradients might provide exciting insights into bet-hedging, and that would allow a quantification of the fitness benefits of informative cues using the framework described above. To this aim, we showed in the examples and the accompanying R notebook (Github) how such data could be analyzed. We emphasize that this approach is general, and we used parabolic environmental clines only as an example. Conversely, the approach applies to cases in which multiple environmental states and environmental factors could affect the fitness of an organism. For instance, density dependence and competition are essential predictors of bet-hedging [43], as in advancing pioneer plants, in which dormancy can evolve in response to the competition with plants adapted to prevailing environmental conditions. Most importantly, this approach provides theoretical upper limits for the fitness advantage of specific cues or adaptive mechanism even in the absence of precise knowledge of fitness, a task that is otherwise hard to achieve. To illustrate this, we look at a hypothetical case inspired by the marine midge *Clunio marinus* (Chironomidae, Diptera). Clunio is adapted to one of the most complex environments on earth, the intertidal zone of seacoasts. Synchronising the adult emergence time to multiple environmental cycles of tidal, diurnal and lunar cycles, the marine midge, has successfully adapted in processing complex informative cues [44]. Using cues, adults emerge when conditions are favourable. Assuming a narrow window of opportunity for successful reproduction, we provide a theoretical upper boundary for the fitness advantage of this trait equivalent to about 0.31 bits (see Github for additional explanation and examples). Providing a metric in bits, valid across species, a thorough quantification of environmental variability in different species along environmental gradients could allow for a clearer understanding of the ecological conditions promoting plasticity and the usage of informative cues.

In conclusion, we have extended the study of bet-hedging to asymmetric cases where phenotypes are scarcer than the number of different environmental states. We show that in realistic conditions, the patterns of bet-hedging might be more complicated than previously expected. The results could potentially help clarify some of the mixed results of empirical observations [18], and better describe the evolution of bet-hedging along complex environmental gradients. However, even the most complicated cases, still obey the limit on the fitness value of a cue imposed by the mutual information.

Acknowledgments

CSG acknowledges Arne Traulsen for helpful discussions and funding from the New Zealand Institute for Advanced Study, Massey University where part of the work was done. All authors acknowledge funding from the Max Planck Society.

References

[1] Hasson O. Cheating signals. Journal of Theoretical Biology. 1994;167(3):223–238.

[2] Maynard Smith J, Harper D. Animal Signals. Oxford University Press; 2003.

[3] Fusco G, Minelli A. Phenotypic plasticity in development and evolution: facts and concepts. Philosophical Transactions of the Royal Society B: Biological Sciences. 2010;365(1540):547–556.

[4] Kokko H, Chaturvedi A, Croll D, Fischer MC, Guillaume F, Karrenberg S, et al. Can Evolution Supply What Ecology Demands? Trends in Ecology & Evolution. 2017;32(3):187–197.

[5] Levins R. Theory of Fitness in a Heterogeneous Environment. I. The Fitness Set and Adaptive Function. The American Naturalist. 1962;96(891):361–373.

[6] Beaumont HJE, Gallie J, Kost C, Ferguson GC, Rainey PB. Experimental evolution of bet hedging. Nature. 2009;462:90–93.

[7] Ratcliff WC, Denison RF. Bacterial persistence and bet hedging in Sinorhizobium meliloti. Communicative & integrative biology. 2011;4(1):98–100.

[8] Starrfelt J, Kokko H. Bet-hedging—a triple trade-off between means, variances and correlations. Biological Reviews. 2012;87(3):742–755.

[9] Soll DR, Kraft B. A comparison of high frequency switching in the yeast Candida albicans and the slime mold Dictyostelium discoideum. Developmental Genetics. 1988;9(4-5):615–628.

[10] Moxon ER, Rainey PB, Nowak MA, Lenski RE. Adaptive evolution of highly mutable loci in pathogenic bacteria. Current Biology. 1994;4(1):24–33.

[11] Balaban NQ, Merrin J, Chait R, Kowalik L, Leibler S. Bacterial Persistence as a Phenotypic Switch. Science, New Series. 2004;305(5690):1622–1625.

[12] Cohen D. Optimizing reproduction in a randomly varying environment. Bioorganic Chemistry. 1966;12(1):119–129.

[13] Slatkin M. Hedging one's evolutionary bets. Nature. 1974;250:704–705.

[14] Bull JJ. Evolution of Phenotypic Variance. Evolution. 1987;41(2):303–315.

[15] Seger J, Brockmann HJ. What is bet-hedging? Oxford University Press; 1987. .

[16] King OD, Masel J. The evolution of bet-hedging adaptations to rare scenarios. Theoretical Population Biology. 2007 Dec;72(4):560–575.

[17] Childs DZ, Metcalf CJE, Rees M. Evolutionary bet-hedging in the real world: empirical evidence and challenges revealed by plants. Proceedings of the Royal Society B: Biological Sciences. 2010;277(1697):3055–3064.

[18] Simons AM. Modes of response to environmental change and the elusive empirical evidence for bet hedging. Proceedings of the Royal Society B: Biological Sciences. 2011;278(1712):1601–1609.

[19] Donaldson-Matasci MC, Bergstrom CT, Lachmann M. When Unreliable Cues Are Good Enough. The American Naturalist. 2013;182(3):313–327.

[20] Bergstrom CT, Lachmann M. Shannon information and biological fitness. In: 2004 IEEE Information Theory Workshop. IEEE; 2004. p. 50–54.

[21] van Baalen M. Biological information: why we need a good measure and the challenges ahead. Interface focus. 2013;3(6):20130030.

[22] Donaldson-Matasci MC, Bergstrom CT, Lachmann M. The fitness value of information. OIKOS. 2010;119(2):219–230.

[23] Gould JP. Risk, stochastic preference, and the value of information. Journal of Economic Theory. 1974;8(1):64–84.

[24] Stephens DW. Variance and the value of information. American Naturalist. 1989;134(1):128–140.

[25] Shannon CE. A Mathematical Theory of Comminication. The Bell System Technical Journal. 1948;27:379–423.

[26] Cover TM, Thomas JA. Elements of Information Theory. John Wiley & Sons, Inc.; 1991.

[27] Grafen A. Formal Darwinism, the individual-as-maximizing-agent analogy and bet-hedging. Proceedings of the Royal Society B: Biological Sciences. 1999;266(1421):799.

[28] Grafen A. Developments of the Price equation and natural selection under uncertainty. Proceedings of the Royal Society B: Biological Sciences. 2000;267(1449):1223–1227.

[29] Gillespie JH. The Causes of Molecular Evolution. Oxford University Press; 1991.

[30] Orr HA. Absolute fitness, relative fitness and utility. Evolution. 2007;61(12):2997–3000.

[31] Orr HA. Fitness and its role in evolutionary genetics. Nature Reviews Genetics. 2009;10:531–539.

[32] Dempster ER. Maintenance of genetic heterogeneity. Cold Spring Harb Symp Quant Biol. 1955;20:25–32.

[33] Haldane J, Jayakar SD. Polymorphism due to selection of varying direction. Journal of Genetics. 1963;58(2):237–242.

[34] Gillespie JH. Natural selection with varying selection coefficients—a haploid model. Genetical Research. 1973;.

[35] Lewontin RC, Cohen D. On population growth in a randomly varying environment. Proceedings of the National Academy of Sciences USA. 1969;62(4):1056–1060.

[36] Donaldson-Matasci MC, Bergstrom CT, Lachmann M. The fitness value of information. OIKOS. 2010;119(2):219–230.

[37] Donaldson-Matasci MC, Lachmann M, Bergstrom CT. Phenotypic diversity as an adaptation to environmental uncertainty. Evolutionary Ecology Research. 2008;10(4):493–515.

[38] Wolf M, Weissing FJ. An explanatory framework for adaptive personality differences. Philosophical Transactions of the Royal Society B: Biological Sciences. 2010;365(1560):3959–3968.

[39] Moreira B, Pausas JG. Tanned or Burned: The Role of Fire in Shaping Physical Seed Dormancy. PLoS ONE. 2012;7(12):e51523.

[40] Kaye TN. Seed dormancy in high elevation plants: implications for ecology and restoration. In: Kaye TN, Liston A, Love RN, Luoma DL, Meinke RJ, Wilson MV, editors. Conservation and Management of Native Plants and Fungi. Plant Society of Oregon, Corvallis, Oregon; 1997. .

[41] Xiao C, Xing W, Liu G. Seed germination of 14 wetland species in response to duration of cold-wet stratification and outdoor burial depth. Aquatic Biology. 2010;11(2):169–177.

[42] Wagmann K, Hautekèete NC, Piquot Y, Meunier C, Schmitt SE, Van Dijk H. Seed dormancy distribution: explanatory ecological factors. Annals of Botany. 2012 Nov;110(6):1205–1219.

[43] Gremer JR, Venable DL. Bet hedging in desert winter annual plants: optimal germination strategies in a variable environment. Ecology Letters. 2014;17(3):380–387.

[44] Kaiser TS, Neumann D, Heckel DG. Timing the tides: Genetic control of diurnal and lunar emergence times is correlated in the marine midge Clunio marinus. BMC Genetics. 2011;12(1):49.

Figure A.1: **Average fitnesses of the strategies Germinate and Dormant as a function of the probability of a single flood event p.** Panel (a) describes the situation where the intermediate environmental state is the only one favourable for germination. However knowing the exact state of the environment can lead to an increase in fitness given by $F(E|C)$. Hence the fitness benefit due to the cue, $\Delta F(E|C)$, is given by the shaded area. Panel (b) describes the case where the optimal gains for germinating are obtained in extreme environments. For this particular example the we get two internal fixed points given by $\{1/2, 3/4\}$. Below and above these fixed points it is best to germinate but in between it pays to stay dormant. Again for a perfect cue the fitness is given by increases by $\Delta F(E;C)$ to $F(E|C)$.

[45] Hauert C, Michor F, Nowak MA, Doebeli M. Synergy and discounting of cooperation in social dilemmas. Journal of Theoretical Biology. 2006;239:195–202.

[46] Gokhale CS, Traulsen A. Evolutionary games in the multiverse. Proceedings of the National Academy of Sciences USA. 2010;107:5500–5504.

A Examples with asymmetries in strategies and environments

Adaptation to intermediate environments. If the difference in benefits in the three environments switch their signs then this satisfies a necessary condition for getting two internal fixed points [45, 46]. Herein we set up an example where we do get two internal fixed points for the three environment case. Here

$$\begin{array}{cc} & \begin{array}{ccc} e_1 & e_3 & e_2 \end{array} \\ \begin{array}{l} \text{Germinate} \\ \text{Dormant} \end{array} & \begin{pmatrix} 1 & 4 & 1 \\ 2 & 2 & 2 \end{pmatrix}, \end{array} \tag{A.1}$$

The fitnesses of the two phenotypes are,

$$f_G = p^2 + 8p(1-p) + 1(1-p)^2 \tag{A.2}$$

$$f_D = 2p^2 + 4p(1-p) + 2(1-p)^2 \tag{A.3}$$

In the absence of any external information about the environment the single generational expected fitness is $F(E) = max[f_G, f_D]$. However with a perfectly informative cue C, the single generational expected fitness, $F(E|C) = 2p^2 + 8p(1-p) + 2(1-p)^2$. We look at the difference between the fitness with perfect information and the strategy of probabilistic allocation according to the game theoretic outcome. The value of information is measured as the difference between the fitness with the cue and without the cue $\Delta F(E;C)$ (Figure A.1).

The frequency with which seeds germinate is assumed to be x while they remain dormant with a complementary probability. The average fitnesses of the population in the different environments is then given by,

$$\begin{aligned} env(1,x) &= x + 2(1-x) \\ env(2,x) &= 4x + 2(1-x) \\ env(3,x) &= x + 2(1-x) \end{aligned} \tag{A.4}$$

Maximising the long term population growth rate means maximising,

$$g(x) = p^2 \log[env(1,x)] + 2p(1-p)\log[env(2,x)] + (1-p)^2 \log[env(3,x)] \tag{A.5}$$

The maxima of this function appears when $x_{max} = -1 + 6p - 6p^2$. Restricting to values of p which give us results for x between 0 and 1 we get,

$$x^*(p) = \begin{cases} 0 & \text{for } p \le \frac{1}{6}\left(3-\sqrt{3}\right) \text{ or } p \ge \frac{1}{6}\left(3+\sqrt{3}\right)\text{, Dormant} \\ x_{max} & \text{otherwise, Bet Hedging} \\ 1 & \text{never, Germinate.} \end{cases} \tag{A.6}$$

The growth rate without the cue therefore is given by,

$$G(E) = \begin{cases} \log(2) & \text{for } p \le \frac{1}{6}\left(3-\sqrt{3}\right) \text{ or } p \ge \frac{1}{6}\left(3+\sqrt{3}\right) \\ -H(E) + p^2 \log\left(\frac{3(1-2p(1-p))}{p^2}\right) & \\ \quad +2p(1-p)\log(6) & \\ \quad +(1-p)^2 \log\left(\frac{3(1-2p(1-p))}{(1-p)^2}\right) & \text{otherwise} \\ 2p(1-p)\log(4) & \text{never.} \end{cases} \tag{A.7}$$

If the cue provides perfect information then the long term fitness value is, $G(E|C) = p^2\log(2) + 2p(1-p)\log(4) + (1-p)^2\log(2)$. The gain in fitness due to knowing the environment, i.e. the fitness value of information is therefore given by, $\Delta G(E;C) = G(E|C) - G(E)$,

$$\Delta G(E;C) = \begin{cases} 2p(1-p)\log 2 & \text{for } p \le \frac{1}{6}\left(3-\sqrt{3}\right) \text{ or } p \ge \frac{1}{6}\left(3+\sqrt{3}\right) \\ H(E) + p^2 \log\left(\frac{2p^2}{3(p^2+(1-p)^2)}\right) & \\ \quad -2p(1-p)\log(3/2) & \\ \quad +(1-p)^2 \log\left(\frac{2(1-p)^2}{3(p^2+(1-p)^2)}\right) & \text{otherwise} \\ (p^2+(1-p)^2)\log(2) & \text{never.} \end{cases} \tag{A.8}$$

The bet-hedging region in this particular case explores the mixed phenotype space between the pure phenotypes. Therefore in the bet-hedging region such "exploratory bet-hedging" occurs that would never be affordable as a pure strategy.

Adaptation to extremes: Multiplicity of bet-hedging. Contrary to the previous example now we consider seeds which perform better germinating in extreme environments.

$$\begin{array}{l} \\ \text{Germinate} \\ \text{Dormant} \end{array} \begin{array}{c} \begin{array}{ccc} e_1 & e_2 & e_3 \end{array} \\ \begin{pmatrix} 4 & 1 & 6 \\ 3 & 3 & 3 \end{pmatrix} \end{array}, \tag{A.9}$$

Equating the average fitness of the two phenotypes which are given by,

$$f_G = 4p^2 + 2p(1-p) + 6(1-p)^2 \tag{A.10}$$

$$f_D = 3p^2 + 6p(1-p) + 3(1-p)^2, \tag{A.11}$$

we get two fixed points denoted by $p^* = \{1/2, 3/4\}$. Thus depending on p the best phenotype, according to the fitnesses, switches from germination to dormancy to germination again.

However if the cue provides perfect information about the environment then the fitness conditioned upon the cue is $F(E|C) = 4p^2 + 6p(1-p) + 6(1-p)^2$. The value of the information received due to the cue is the difference between the fitness with the cue and without the cue $\Delta F(E;C)$ (Figure A.1). Assuming that the seeds germinate with probability x and stay dormant with probability $1-x$, the average fitnesses of the population in the different environments are then given by,

$$\begin{aligned} env(1,x) &= 4x + 3(1-x) \\ env(2,x) &= x + 3(1-x) \\ env(3,x) &= 6x + 3(1-x) \end{aligned} \tag{A.12}$$

Maximising the long term population growth rate, we get

$$g(x) = p^2 \log[env(1,x)] + 2p(1-p)\log[env(2,x)] + (1-p)^2 \log[env(3,x)] \tag{A.13}$$

whose maxima is given by, $x_{max}^{\pm} = \frac{1}{4}\left(-3 - 10p + 14p^2 \pm \sqrt{196p^4 - 280p^3 + 208p^2 - 180p + 81}\right)$. While one of the solutions is outside the range of x ($0 \le x \le 1$) the other results into the following piecewise solution of x^*,

$$x^*(p) = \begin{cases} 0 & \text{for } 0.5 \le p \le 0.75 \text{ (Dormant)} \\ x_{max}^{+} & \text{for } \frac{1}{19}\left(10-\sqrt{62}\right) < p < 0.5 \text{ and } 0.75 < p < \frac{1}{19}\left(10+\sqrt{62}\right) \text{ (Bet Hedging)} \\ 1 & \text{for } p \le \frac{1}{19}\left(10-\sqrt{62}\right) \text{ or } p \ge \frac{1}{19}\left(10+\sqrt{62}\right) \text{ (Germinate).} \end{cases} \tag{A.14}$$

This shows an interesting situation where there are two regions of bet-hedging. The first one changes the phenotype from pure germination to pure dormancy and vice versa in the second region (Fig. A.1). The growth rate without the cue therefore is given by,

$$G(E) = \begin{cases} \log(3) & \text{for } 0.5 \leq p \leq 0.75 \\ -H(E) + p^2 \log\left(\frac{9-\phi}{4p^2}\right) & \\ \quad +2p(1-p)\log\left(\frac{\phi+9}{4(1-p)p}\right) & \\ \quad +(1-p)^2 \log\left(\frac{3(1-\phi)}{4(1-p)^2}\right) & \text{for } \frac{1}{19}\left(10-\sqrt{62}\right) < p < 0.5 \text{ and } 0.75 < p < \frac{1}{19}\left(10+\sqrt{62}\right) \\ p^2\log(4) + (1-p)^2\log(6) & \text{for } p \leq \frac{1}{19}\left(10-\sqrt{62}\right) \text{ or } p \geq \frac{1}{19}\left(10+\sqrt{62}\right) \end{cases} \tag{A.15}$$

where we use the abbreviation $\phi = 10p - 14p^2 - \sqrt{196p^4 - 280p^3 + 208p^2 - 180p + 81}$. For a perfectly informative cue the long term fitness value is, $G(E|C) = p^2 \log 4 + 2p(1-p)\log 3 + (1-p)^2 \log 6$. The gain in fitness due to knowing the environment, i.e. the fitness value of information is therefore given by, $\Delta G(E;C) = G(E|C) - G(E)$,

$$\Delta G(E;C) = \begin{cases} p^2 \log 4/3 + (1-p)^2 \log 2 & \text{for } 0.5 \leq p \leq 0.75 \\ H(E) + p^2 \log\left(\frac{4p^2}{4-\phi}\right) & \\ \quad +2p(1-p)\log\left(\frac{6(1-p)p}{2\phi+1}\right) & \\ \quad +(1-p)^2 \log\left(\frac{2(1-p)^2}{2-\phi}\right) & \text{for } \frac{1}{19}\left(10-\sqrt{62}\right) < p < 0.5 \text{ and } 0.75 < p < \frac{1}{19}\left(10+\sqrt{62}\right) \\ 2p(1-p)\log 3 & \text{for } p \leq \frac{1}{19}\left(10-\sqrt{62}\right) \text{ or } p \geq \frac{1}{19}\left(10+\sqrt{62}\right). \end{cases}$$

(A

Thus even for this complicated case with two regions of bet hedging we can see that the value of the cue is a composite of the mutual information between the environment and the cue and a non-linear function in p, which can be interpreted as the probability of a single flood.