The trade-off between growth and risk in Kelly's gambling and beyond

S. Cavallero[a], A. Rousselot[a], R. Pugatch[b], L. Dinis[c], D. Lacoste[a]

[a] *Gulliver Laboratory, UMR CNRS 7083, PSL Research University, ESPCI, 10 rue Vauquelin, Paris, 75231, France*
[b] *Department of Industrial Engineering and Management, Ben-Gurion University of the Negev, Beer-Sheva, Israel*
[c] *Grupo Interdisciplinar de Sistemas Complejos (GISC), Department of Structure of Matter, Thermal Physics and Electronics, Universidad Complutense de Madrid, Madrid, 28040, Spain*

Abstract

We study a generalization of Kelly's horse model to situations where gambling on horses other than the winning horse does not lead to a complete loss of the investment. In such cases, the odds matrix is non-diagonal, which is particularly interesting for biological applications. We examine the trade-off between the mean growth rate and its asymptotic variance, an approximation for risk. Because the consequences of fluctuations around the average growth rate are asymmetric, we further explore a better alternative definition of risk: the extinction probability and its implications for Kelly gambling and the risk-return trade-off. We discuss some applications of these concepts in biology and ecology.

Keywords:
bet-hedging, information theory, game theory, optimization

Introduction

In his seminal work from 1948, Shannon founded information theory [1]. A pivotal contribution of Shannon's theory was an existence proof he provided for a code that can allow signals to pass through a noisy channel with a negligible loss of information as long as the rate is smaller than the channel capacity. This was a big surprise back in the day since the belief up to that point was that noise monotonically reduces the information rate. Hence, from its outset, information theory involved coding. Enter J. R. Kelly, Shannon's friend and colleague. In 1956, Kelly found a surprising example that hinted at a deep connection between information theory and gambling in horse races. With Shannon's support, he published a paper, which is by now well-known and highly influential, titled "A new interpretation of information rate" [2]. In his article, Kelly showed, surprisingly, that the channel capacity also appears in the context of repeated bet-hedging on stationary horse races as the growth rate of the bettor's capital.

A horse race is said to be stationary if both the odds and winning probabilities for each horse are constant in time. It can be shown that if one is trying to optimize the expected capital, one has to bet all the capital on the horse with the maximal distribution of winning. However, this will lead to the complete ruin of all the bettor's capital in the long run. Kelly offered an alternative optimization criterion, which is by no means unique, namely, to optimize the growth rate of the capital. It then proceeded to show that in scenarios where partial information regarding the winning probability is provided by an insider ("side information"), the optimal strategy using his criterion can be calculated; it depends on the conditional probability of winning each horse and leads to a growth rate that is equal to the channel capacity of the insider-bettor information channel (where the information is the uncertain identity of the winning horse). As mentioned, there is no apparent coding involved in the scheme. Later, Cover et al. devised a coding scheme similar to arithmetic coding that uses two identical bettors to code and decode messages (see [3], chapter 6).

Over the years, Kelly's idea grew into a whole branch of information theory [3] with theoretical and practical implications for portfolio management [4] and more generally for investment strategies in finance [5]. Kelly's model can also be formulated as a non-linear control problem with fruitful implications in finance and applied mathematics [6].

In biology, Kelly's work is important because it connects information and fitness [7, 8], a central question in evolutionary biology. This connection is made through bet-hedging—a strategy that spreads risks among various phenotypes within a population, increasing the overall chance of survival under uncertain conditions [9, 10, 11]. Such a strategy is employed, for instance, by cells to cope with antibiotics [12], by phages to optimize their infection strategy against bacteria [13], or by plants to adapt to a fluctuating climate [14]. The latter three examples, in particular, involve a dormant state that protects individuals from harsh environmental conditions while preserving biodiversity [15].

In the context of gambling, Kelly's strategy is known to be risky, and for this reason, most gamblers use fractional Kelly's strategies, with reduced risk and growth rate [4]. This observation hints at a trade-off between the risk the gambler is ready to take and the average long-term growth rate of his capital, which is known in finance under the name of the risk-return trade-off. In previous work, we have studied this trade-off in a version of the Kelly model with a risk constraint [16]. In subsequent work, we found a similar trade-off in the context of a biological population with phenotypic switching in a fluctuating environment [17] by building on a piece-wise Markov model introduced earlier [18]. We have also proposed an adaptive version of Kelly's gambling based on Bayesian inference [19].

In section 1, we introduce Kelly's gambling model, then in section 2 we present the game theoretical formulation of the model building on Ref. [20]. A generalization of Kelly's model is presented with non-diagonal odds, in which the gambling on horses other than the winning horse does not lead to a complete investment loss [7, 21, 22]. Such an extension was not considered in the classic paper on phenotypic adaptation in varying environments by Kussel and Leibler [8] which instead relied on diagonal odds. In fact, this extension of Kelly's model to non-diagonal odds is particularly important for applications to biology because a given phenotype is never only adapted to one specific state of the environment; instead, there is a distribution of environment states that correspond to a given phenotype—the equivalent of the bets.

In section 3, we derive inequalities that characterize the risk-return trade-off for Kelly's model and for its generalization with non-diagonal odds. Finally, in section 4, we explore an alternate measure of risk, based on the capital drawdown [23] rather than the volatility. We explore the consequences of this alternate definition of risk for a formulation of risk-constrained Kelly gambling [24] and for the risk-return trade-off.

1. Definition of Kelly's model

Let us recall Kelly's horse race model [2]. A race involves M horses and is described by a normalized vector of winning probabilities $\mathbf{p}$, an inverse-odds vector $\mathbf{r}$ (or equivalently an odds vector $\mathbf{o}$) and a vector of bets which defines the gambler strategy $\mathbf{b}$. The latter corresponds to a specific allocation of the gambler's capital on the M horses: if we denote by C_t the gambler's capital at time t, the amount of capital invested on horse x reads $\mathrm{b}_x C_t$. We further assume that, after each race, the gambler invests his whole capital, i.e., $\sum_{x=1}^{M} \mathrm{b}_x = 1$, always betting a non-zero amount on all horses, i.e., $\forall x \in [1, M]: \ \mathrm{b}_x \neq 0$. The inverse-odds vector $\mathbf{r}$ is set by the bookmaker. When $\sum_x \mathrm{r}_x = 1$, the odds are *fair*, when $\sum_x \mathrm{r}_x > 1$ the odds are *unfair* and when $\sum_x \mathrm{r}_x < 1$ the odds are *superfair*. To define a probability distribution from the vector $\mathbf{r}$ which is normalized in all cases, an obvious choice is to introduce the distribution $\tilde{\mathrm{r}}_x = \mathrm{r}_x / \sum_x \mathrm{r}_x$.

The evolution of the gambler's capital after one race reads:

$$C_{t+1} = \frac{\mathrm{b}_x}{\mathrm{r}_x} C_t, \quad \text{with a probability } \mathrm{p}_x, \tag{1}$$

which implies that the log of the capital, $\text{log-cap}(t) \equiv \log C_t$, evolves additively:

$$\text{log-cap}(t) = \sum_{\tau=1}^{t} \log\left(\frac{\mathrm{b}_{x_\tau}}{\mathrm{r}_{x_\tau}}\right), \tag{2}$$

where x_τ denotes the index of the winner of the τ-th race and we assumed log-cap$(0) = 0$ (i.e. $C_0 = 1$). Since the races are assumed to be independent, the terms $\log(\mathrm{b}_{x_\tau}/\mathrm{r}_{x_\tau})$ in (2) are independent and identically distributed, and we can use the weak law of the large numbers:

$$\frac{\text{log-cap}(t)}{t} \xrightarrow[t\to\infty]{} \mathbb{E}\left[\log\left(\frac{\mathrm{b}_x}{\mathrm{r}_x}\right)\right] \tag{3}$$

in probability. It follows from this relation that under multiplicative dynamics, the rate of change of the logarithm of the capital is an ergodic observable [25, 26], a key point related to the differences between arithmetic and geometric averages [27]. Then, we define the growth rate as the long-term increase of the log-capital:

$$\begin{aligned}\langle W(\mathbf{p},\mathbf{b})\rangle &= \mathbb{E}\left[\log\left(\frac{\mathrm{b}_x}{\mathrm{r}_x}\right)\right] \equiv \sum_x \mathrm{p}_x \log\left(\frac{\mathrm{b}_x}{\mathrm{r}_x}\right) &(4)\\ &= D_{KL}\left(\mathbf{p}||\tilde{\mathrm{r}}\right) - D_{KL}\left(\mathbf{p}||\mathbf{b}\right) - \ln(\sum_x \mathrm{r}_x), &(5)\end{aligned}$$

where D_{KL} stands for the Kullback-Leibler divergence between two probability distributions [3, sec. 2.3]. From an information-theoretic point of view, (3) and (4) imply that the capital of the gambler increases in the long term only if the gambler has a better knowledge of $\mathbf{p}$ than the bookmaker; otherwise, it decreases.

It also follows from this analysis that when the odds are fair, the optimal strategy $\mathbf{b}^{\text{KELLY}} = \mathbf{p}$, called *Kelly's strategy* [2], overtakes any other strategies in the long-term. The corresponding optimum growth rate is the positive quantity $D_{\mathrm{KL}}\left(\mathbf{p}\|\mathbf{r}\right)$, and the strategy $\mathbf{b}^{\text{NULL}} = \mathbf{r}$ also plays an specific role. We have called this strategy the *null strategy* [16] because it yields asymptotically a constant capital as can be seen from Eqs. (3)-(4).

Risk can be estimated using volatility, which is the asymptotic variance of fluctuations in the capital growth rate. This measure is known to be imperfect and less appropriate than methods that account for asymmetry in fluctuation directions because positive fluctuations of gain relative to the mean are not considered risk, while negative fluctuations are. In section 4, we will explore an alternative risk measure for this reason. Nevertheless, we first use volatility because it allows for tractable calculations.

Since we have considered independent races, by the central limit theorem, the rescaled log-capital converges in law towards a centered Gaussian distribution of unit variance:

$$\frac{1}{\sigma_W\sqrt{t}}\left(\log\frac{C_t}{C_0} - t\langle W\rangle\right) \xrightarrow[t\to\infty]{} \mathcal{N}(0,1), \tag{6}$$

where

$$\sigma_W^2 = \mathrm{Var}\left[\log\left(\frac{\mathrm{b}_x}{\mathrm{r}_x}\right)\right], \tag{7}$$

is the volatility. From this definition, one can see that the *null strategy* has a zero volatility, *i.e.* it is risk free. Note also that for an arbitrary strategy, risk is relevant at intermediate times, $t \ll (\sigma_W/\langle W \rangle)^2$, long-enough for the central limit theorem to apply but not too long for deviations from exponential growth to become negligible.

2. Game-theoretic formulation of the asymptotic growth rate

We start by asking a simple question: What is the maximum average growth rate that can be secured by a bettor who has no prior knowledge of the winning probabilities of the horses? To secure this value of the growth rate V, the bettor will play a specific strategy $\mathbf{b} = \mathbf{r}$, which we will look for. The value V and the betting strategy $\mathbf{r}$ will then serve us as a benchmark to judge other betting strategies. Clearly, we expect strategies that correctly employ information about the winning probabilities to yield higher growth rates.

First, let us define the value V, which is the largest possible growth rate that can be secured by the bettor if there is no information regarding the horses' winning probabilities. By definition this will be the maximal growth rate that can be obtained against the worst possible winning probability vector $\mathbf{p}^*$, given the odds, because the probability of winning will increase for any probability vector that differs from $\mathbf{p}^*$:

$$V = \max_{\mathbf{b}} \min_{\mathbf{p}} \langle W(\mathbf{p}, \mathbf{b}) \rangle, \tag{8}$$

$$\begin{aligned} \mathbf{p}^* &= \mathrm{argmax}_{\mathbf{p}}\ \mathrm{argmin}_{\mathbf{b}} \langle W(\mathbf{p}, \mathbf{b}) \rangle. && (9) \\ \mathbf{b}^* &= \mathrm{argmax}_{\mathbf{b}}\ \mathrm{argmin}_{\mathbf{p}} \langle W(\mathbf{p}, \mathbf{b}) \rangle. && (10) \end{aligned}$$

Clearly, for any $\mathbf{p}$, we have that $\langle W(\mathbf{p}, \mathbf{b}^*) \rangle \geq V$, which shows this is indeed a guarantee for the expected growth. There are three possibilities. The first is known as a sub-fair game which occurs when $V < 0$. In this case, the bettor cannot secure a gain, just secure a loss with a minimal rate of loss. This might or might not turn to gain for certain $\mathbf{p}$'s. The second case is $V = 0$, i.e., the bettor can guarantee a minimal amount of growth of the investment by investing in a way that will keep the average growth rate to zero. And finally, super-fair case—-also known as a Dutch book—where $V > 0$ in which case the bettor has a secured minimum strategy.

In game theory, the strategy that guarantees maximal growth, assuming the worst possible probability of winning for the horses, is known as the min-max solution to the game, or more generally, a Nash equilibrium [28]. The game under consideration is rather untypical; it is a zero-sum game where the expected growth rate is the payoff. The first player is the bettor who attempts to place the bets $\mathbf{b}$ to maximize the growth rate, while the second player, quite unusually, controls the horses' probability of winning $\mathbf{p}$, and his payoff is minus the payoff of the bettor. In biological applications, the second player could represent a fluctuating environment. In the following, we apply these concepts to the cases of diagonal and non-diagonal odds in Kelly's model.

2.1. Kelly's case (diagonal odds)

For Kelly's optimal strategy $\mathbf{b}^{\mathrm{KELLY}} = \mathbf{p}$, the growth rate is

$$\langle W(\mathbf{p}, \mathbf{b}^{\mathrm{KELLY}}) \rangle = \sum_x \mathrm{p}_x \ln(\mathrm{o}_x \mathrm{p}_x), \tag{11}$$

with $\mathrm{o}_x = 1/\mathrm{r}_x$. Let us now evaluate the worst possible scenario with the given odds o_x, i.e. the value of $\mathbf{p}$ such that his/her growth rate is minimal. This can be done by minimizing the function $\Psi(\mathbf{p}) = \langle W(\mathbf{p}, \mathbf{b}^{\mathrm{KELLY}}) \rangle - \lambda \sum_x \mathrm{p}_x$ with respect to $\mathbf{p}$, where the Lagrange multiplier enforces the normalization of $\mathbf{p}$. One obtains that the worst scenario occurs when

$$\mathrm{p}_x = \mathrm{p}^*_x = \frac{\mathrm{r}_x}{\sum_x \mathrm{r}_x}. \tag{12}$$

One can then write

$$\begin{aligned} \langle W(\mathbf{p}, \mathbf{b}) \rangle &= \sum_x \mathrm{p}_x \ln(\mathrm{o}_x \mathrm{b}_x), \qquad (13) \\ &= \sum_x \mathrm{p}_x \ln \left(\frac{\mathrm{b}_x}{\mathrm{r}_x / \sum_x \mathrm{r}_x} \cdot \left(\sum_x \mathrm{r}_x \right)^{-1} \right), \\ &= \sum_x \mathrm{p}_x \ln \left(\frac{\mathrm{p}_x}{\mathrm{r}_x / \sum_x \mathrm{r}_x} \cdot \left(\sum_x \mathrm{r}_x \right)^{-1} \cdot \frac{\mathrm{b}_x}{\mathrm{p}_x} \right), \\ &= D_{KL}(\mathbf{p}||\mathbf{p}^*) - D_{KL}(\mathbf{p}||\mathbf{b}) + V. \qquad (14) \end{aligned}$$

These three terms have the following interpretations:

- $D_{KL}(\mathbf{p}||\mathbf{p}^*)$ is the 'pessimist's surprise,' which measures by how much the growth rate is larger than the worst expected growth rate the bettor can guarantee under the worst conditions.

- $-D_{KL}(\mathbf{p}||\mathbf{b})$ is the 'bettor's regret', the loss in expected growth caused by playing sub-optimally.

- $V = -\ln \sum_x \mathrm{r}_x$ is the value of the game, the minimal growth rate that the gambler can expect irrespective of how $\mathbf{p}$ is chosen. In practice, this minimum is attained for the *null strategy* where $\mathbf{b} = \mathbf{p}^*$. Further, when the odds are *unfair*, $V < 0$, whereas when the odds are *super-fair*, $V > 0$. This definition of the *null strategy* generalizes the previous one when the odds are not fair.

We also use the following two concepts from game theory namely, dominance and mixing. Zero sum games in strategic form are represented by a payoff matrix, whose rows and columns represent players' pure strategies and the value of the matrix represent the payoff for the first players, and the payoff for the second players is minus the payoff to the first. A zero-sum game either has a min-max solution which is a pure strategy, or a mixed strategy solution, where the action is chosen at random with different probabilities. A fully mixing game is a game for which the min-max solutions take the form of probability vectors such that every possible pure strategy has a finite probability.

In Kelly's model, this corresponds to all elements in $\mathbf{b} \neq 0$ and $\mathbf{p} \neq 0$. Later we will state the conditions for the odd matrix such that the optimal solution is fully mixing and we refer to that game as a *fully mixing* game. Finally, a related concept is strategy dominance. If the game is not fully mixing, then the optimal solution has some zero bet or zero probability element, for instance, "never bet on horse number 2" in words. The pure strategies that bet some quantity on horse 2 become irrelevant and are usually referred to as being dominated by other more performing strategies. The dominated strategies can be removed and the game reduced to its essential part. We provide an example of game reduction in section 2.3.

2.2. Non-diagonal odds

In the general case, the matrix of the odds o_{xy} that gives the reward to a bet y when the winning horse is x, is non-diagonal, and the corresponding growth rate may be written as :

$$\langle W(\mathbf{p}, \mathbf{b}) \rangle = \sum_x \mathrm{p}_x \ln \left(\sum_y \mathrm{o}_{xy} \mathrm{b}_y \right). \tag{15}$$

A particular case of non-diagonal odds corresponds to the situation described in the original Kelly's paper has a 'track take', where the gambler has the option to not bet a fraction of his/her capital. In that case the growth rate is

$$\langle W(\mathbf{p}, \mathbf{b}) \rangle_{TT} = \sum_x \mathrm{p}_x \ln \left(b_0 + \mathrm{o}_x \mathrm{b}_x \right), \tag{16}$$

where b_0 is the fraction of the capital which is not bet. This case corresponds to a non-diagonal matrix of odds which contains a diagonal part and an isolated full column filled with ones. The optimal solution for the bets has been considered in the original Kelly's paper. This solution can be recovered with the Karush-Kuhn-Tücker (KKT) method as shown explicitly in Ref. [29].

An explicit optimization solution with respect to the bets can be obtained provided two conditions (i) and (ii) are met for the odds matrix.

The first condition (i) requires that this matrix is invertible and simplex preserving. Simplex preserving implies multiplying any probability vector by the inverse matrix will keep the vector inside the simplex. This means that the odds matrix, viewed as a game, is *fully mixing* [28]. An equivalent mathematical condition is that the inverse-odds matrix $\mathrm{r} = \mathrm{o}^{-1}$ has no zero in it. When this condition is satisfied, we can build the following game-theoretic solution:

$$\Omega_{xy} = \frac{\mathrm{r}_{xy}}{\sum_l \mathrm{r}_{ly}}, \tag{17}$$

which is such that

$$\sum_x \Omega_{xy} = \frac{\sum_x \mathrm{r}_{xy}}{\sum_l \mathrm{r}_{ly}} = 1. \tag{18}$$

Then, one can show that the optimal bets are given by

$$\mathrm{b}^*_x = \sum_y \Omega_{xy} \mathrm{p}_y. \tag{19}$$

To proceed with the game-theoretic analysis, we need to look for the worst-case scenario, i.e., for the value $\mathbf{p}^*$ of $\mathbf{p}$ yielding the minimal growth rate for the optimal strategy of the gambler. Using again the method of Lagrange multiplier, one finds

$$\mathrm{p}^*_x = \frac{\sum_l \mathrm{r}_{lx}}{\sum_{xy} \mathrm{r}_{xy}}, \tag{20}$$

which is acceptable, provided all the components of p_x are non-negative. This requires that for all x, $\sum_l \mathrm{r}_{lx} > 0$, which is our second condition (ii). When both conditions hold, the matrix Ω is stochastic (or more precisely, pseudo-stochastic because it can contain negative elements), and there is unique pair $(\mathbf{p}^*, \mathbf{b}^*)$, which represents a Nash equilibrium for the matrix game defined by the odds matrix o.

To obtain the equivalent of the decomposition of Eq. (13) for the case of non-diagonal odds, we start by evaluating the optimal growth rate when the bets are optimal and given by Eq. (19):

$$\begin{aligned} \langle W(\mathbf{p}, \mathbf{b}^*) \rangle &= \sum_x \mathrm{p}_x \ln \left(\sum_y \mathrm{o}_{xy} \sum_z \Omega_{yz} \mathrm{p}_z \right), \\ &= D_{KL}(\mathbf{p} || \mathbf{p}^*) + V, \end{aligned} \tag{21}$$

where $\mathbf{p}^*$ is the one of Eq. (20) and now the value of the game is $V = -\ln(\sum_{xy} \mathrm{r}_{xy})$. In a second step, one can then check that the growth rate for non-optimal bet $\mathbf{b}$ is less or equal than $\langle W(\mathbf{p}, \mathbf{b}^*) \rangle$ and that the difference is the term associated to the better's regret. Thus, the general decomposition of the asymptotic growth rate for non-diagonal odds takes a form similar to Eq. (13):

$$\begin{aligned} \langle W(\mathbf{p},\mathbf{b})\rangle &= W(\mathbf{p},\mathbf{b}^*) - D_{KL}(\mathbf{p}||\mathbf{b}), \\ &= D_{KL}(\mathbf{p}||\mathbf{p}^*) - D_{KL}(\mathbf{p}||\mathbf{b}) + V. \end{aligned} \tag{22}$$

Note that the value of the game has the same interpretation as before and the minimum of the growth rate is still attained for the *null strategy*.

2.3. Illustration with a three horses example

Let us illustrate this game-theoretic framework for a simple case of three horses in which only the gambler can play optimally. Let's consider specifically a diagonal and a non-diagonal odds matrix given by:

$$\mathbf{O}_d = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \qquad \mathbf{O}_{nd} = \begin{pmatrix} 2 & \frac{2}{3} & 1 \\ \frac{5}{6} & \frac{5}{3} & \frac{5}{6} \\ 1 & \frac{2}{3} & 2 \end{pmatrix}, \tag{23}$$

and an environment characterized by the vector $p = (p_0, p_1, p_2)$ where p_0 is varied in $(0, 1)$, $p_1 = \frac{p[1]}{p[1]+p[2]}(1 - p_0)$ and $p_2 = \frac{p[2]}{p[1]+p[2]}(1 - p_0)$, and $(p[1], p[2]) = (0.5, 0.3)$.

When the game is fully mixing, there is an analytical solution for the optimal solution namely (17)-(20). When it is not, we need to resort to numerical optimization. Simulated annealing and KKT optimization are two possible methods to do this. We have found empirically that the later gives better results than the former for low dimensions problems, which is the case here, since we only consider three horses. Below, we only use the KKT method. The optimization problem we want to solve is :

$$\max_{\mathbf{b}} \langle W \rangle = \max_{\mathbf{b}} \mathbf{E}\left(\log\left(\sum_y \frac{b_x}{r_{xy}}\right)\right), \tag{24}$$

subject to

$$\sum_x b_x = 1, \quad \text{and} \quad \forall x, b_x \geq 0 \tag{25}$$

To solve it with KKT method, one introduces the functional :

$$\mathcal{L}(\mathbf{b}, \lambda, \mu) = \mathbf{E}\left(\log\left(\sum_y \frac{b_x}{r_{xy}}\right)\right) + \sum_x \lambda_x b_x + \mu\left(\sum_x b_x - 1\right) \tag{26}$$

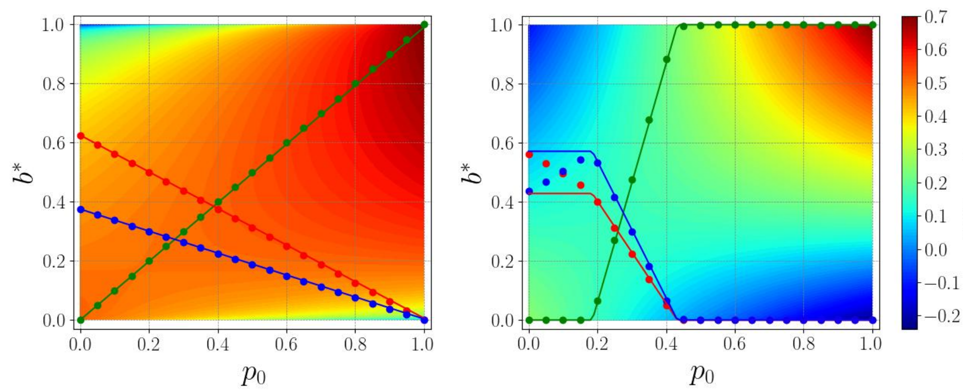
where λ_x and μ are Lagrange multipliers. Since the problem is concave for $\mathbf{b}$, we set the first derivative to zero to obtain the point of maximum, i.e. the optimal strategy $\mathbf{b}^*$.

$$\frac{\partial \mathcal{L}}{\partial b_x} = \sum_k p_k \frac{o_{kx}}{\sum_y o_{ky} b_y} + \lambda_x + \mu = 0.$$

The KKT solution is built by combining the solution of that equation for three mutually exclusive cases : (i) one bet b_i is zero while the other two b_j for $j \neq i$ are strictly positive. In that case, we have $\lambda_i > 0$. (ii) two bets are zero $b_i = b_j = 0$ while the last one is strictly positive. In that case, similarly $\lambda_i > 0$ and $\lambda_j > 0$, and (iii) none of the bets is zero, which is the same solution as that of (17)-(20).

A comparison between the result of KKT optimization (symbols) with the game-theoretical solution of equation (17)-(20) is shown in Fig. (1) which shows the optimal strategies b_0^*, b_1^*, b_2^* as a function of $p_0 \in (0,1)$, for the case of diagonal odds (a) and non-diagonal odds (b). It is interesting to notice that, as the environment probability vector varies, only in a certain interval of values of p_0 the game is fully mixing and (17)-(20) apply. Outside of this interval, one or more $b_i = 0$, the game is non-fully mixing and the game-theoretical solution is no longer correct. This explains the deviation with the correct result obtained from the KKT maximization for small value of p_0.

Fig 1. KKT solution vs. game theoretic solution. Optimal bets b_i^* on the three horses from KKT optimization (symbols) and game-theoretical solution deduced from (17)-(20) (solid lines) as function of the probability on the first horse p_0. Figure (a) corresponds to the diagonal odd matrix $\mathbf{O}_d$, figure (b) to the non-diagonal odd matrix $\mathbf{O}_{nd}$, the green color corresponds to first horse, red color for the second one, blue for the third one. The colored intensity represents the average growth rate $\langle W \rangle$.



The figures also show the average growth rate as a color plot, projected along the plane (p_0, b_0) for the diagonal and non-diagonal cases. Note that in this color plot, a particular choice is made for varying the parameters p_1, p_2, b_1, b_2 as p_0 and b_0 vary while satisfying normalization constraints. In the diagonal case, one can see that the growth rate is the highest for a fixed p_0, *i.e.* along the green line where $b_0 = p_0$ as expected from Kelly's gambling. Moreover, in the non-diagonal case, the game-theoretical solution takes instead the form of piece-wise linear functions. The colored intensity representing the average growth rate takes the form of a saddle point, which is visible here only in projection.

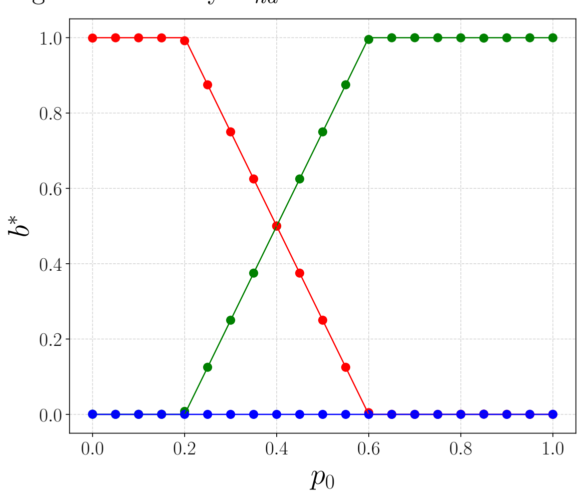
2.4. Illustration of the reduction of a game

The derivation of the game-theoretical solution of equation (17)-(20) assumes (i) a fully mixing game, (ii) an odd matrix which is invertible and simplex preserving. In the previous section, we have seen what happens when the game is not fully mixed. In that case, the strategies corresponding to zero bets or zero probability p_i become in a sense irrelevant because they are dominated by other strategies corresponding to non-zero b_i or p_i. These irrelevant strategies can be removed, and when doing so, one transforms the game into what is called the essential part of the game [28]. Let us illustrate this reduction procedure by starting with the 3x3 game defined by $\mathbf{O}_{nd}$, which we will reduce to a 2x2 game. The reason that we do not consider a reduction to a 3x2 game for instance, is because we need the reduced game to be a square matrix so that it can be invertible and (17)-(20) can apply.

Let us consider an input strategy vector for the environment given by $(p[1], p[2]) = (0.8, 0)$, so that while p_0 varies, $p_2 = 0$ always, and one can check that in that case one also has $b_2^* = 0 \; \forall p_0$. As a result, the last horse should never be played and the odd matrix can be reduced by removing the last row and last column in $\mathbf{O}_{nd}$. The odd matrix of the essential part of the game is then:

$$\mathbf{O}_{red} = \begin{pmatrix} 2 & \frac{2}{3} \\ \frac{5}{6} & \frac{5}{3} \end{pmatrix}. \tag{27}$$

Fig 2. Illustration of the reduction of the game to its essential part. Comparison between the result of KKT optimization (symbols) and the game-theoretical solution (solid lines) for the reduced odds matrix $\mathbf{O}_{red}$, which is the essential part of the game defined by $\mathbf{O}_{nd}$.



Now, since the essential part of the game is fully mixing, and fulfills the assumptions under which relations (17)-(20) have been derived, one can use these equations to obtain the optimal solution and check that they agree with the KKT solution as shown in Fig. 2.

3. Risk-return inequalities and their associated trade-off

3.1. Non-fair but diagonal odds

In recent work, we have studied the trade-off between the mean growth rate and the risk, measured by the volatility in the case of Kelly's original model with *fair* (and diagonal) odds [16]. This trade-off is embodied in the following mean-variance inequality :

$$\sigma_W \geq \frac{\langle W \rangle}{\sigma_{\mathrm{q}}}, \tag{28}$$

where σ_W is the volatility, $\langle W \rangle$ the average growth rate and σ_q is the standard deviation of the ratio $\mathrm{q}_x = \mathrm{r}_x/\mathrm{p}_x$, which compares the probability of races outcomes with the bets of the risk-free strategy (namely $\mathrm{b}_x = \mathrm{r}_x$). Inequality (28) holds for any $\langle W \rangle > 0$. In the negative growth region, it only applies near the null strategy where $D(\mathrm{r}||\mathrm{b}) \approx 0$. This inequality means in practice that a capital growing exponentially with a rate $\langle W \rangle > 0$ necessarily has a non-zero risk measured by the volatility. Recently, a similar bound has been derived for a wide class of models including the Black-Scholes and the Heston models [30]. In fact, bounds of this type are related to the Chapman-Robbins bound and to the thermodynamic uncertainty relations studied in Stochastic Thermodynamics.

Let us first generalize this result to the case where the odds are not *fair* but still diagonal. Note that the strategy $\mathrm{b}_x = \tilde{\mathrm{r}}_x$ corresponds to the *null strategy* introduced previously. For such a strategy, the asymptotic growth rate $\langle W \rangle$ is equal to V, independently of the choice of the bets and of the horse winning probabilities.

The definition of the quantity q is unchanged with respect to the case of fair odds, the only difference is that now $\langle \mathrm{q} \rangle \neq 1$. Thus, q itself is no longer a distribution since it is not normalized. Let us now go through the same steps which lead previously to Eq.(28). We start with

$$\langle \mathrm{q}W \rangle = \sum_x \mathrm{r}_x \ln \frac{\mathrm{b}_x}{\mathrm{r}_x} = \langle \mathrm{q} \rangle \sum_x \tilde{\mathrm{r}}_x \ln \frac{\mathrm{b}_x}{\mathrm{r}_x} = \langle \mathrm{q} \rangle \left(-D_{KL}(\tilde{\mathbf{r}}||\mathbf{b}) + V \right). \tag{29}$$

Now, we write the covariance between q and W as :

$$\mathrm{Cov}(\mathrm{q}, W) = \langle \mathrm{q}W \rangle - \langle \mathrm{q} \rangle \langle W \rangle = \langle \mathrm{q} \rangle \left(-D_{KL}(\tilde{\mathbf{r}}||\mathbf{b}) + V - \langle W \rangle \right). \tag{30}$$

Using Cauchy-Schwartz inequality, namely $\mathrm{Cov}(\mathrm{q}, W)^2 \leq \sigma_{\mathrm{q}}^2 \sigma_W^2$ and the positivity of the Kullback-Leibler divergence, we obtain the generalization of Eq. (28) for non-fair odds as :

$$\sigma_W \geq \frac{|\langle W \rangle - V|}{\sigma_{\mathrm{q}}} \langle \mathrm{q} \rangle. \tag{31}$$

valid for any $\langle W \rangle - V > 0$ and only close to the null strategy for $\langle W \rangle - V < 0$.

From the inequality of Eq. (31), we find that any strategy different from the *null strategy* which is risk free will have a non-zero volatility σ_W. Note also that the inequality is saturated for the risk free strategy for which $\sigma_W = 0$ and $\langle W \rangle = V$. A numerical illustration of that inequality is provided in Fig. 3 for a case where $V = 0$ and a case where $V > 0$.

3.2. Non-diagonal odds

The mean-variance trade-off for non-diagonal odds can be derived similarly to the diagonal case, provided the same conditions denoted (i) and (ii) hold. Condition (i) is the regularity and simplex preserving character of the odds matrix.

Now, the relevant probability ratio q has the form $\mathrm{q}_x = \sum_y \mathrm{r}_{yx}/\mathrm{p}_x$ so that $\langle \mathrm{q} \rangle = \sum_{xy} \mathrm{r}_{xy} = \exp(-V)$ in terms of the value of the game $V = -\ln(\sum_{xy} \mathrm{r}_{xy})$.

We start again with

$$\langle \mathrm{q}W \rangle = \sum_x \mathrm{p}_x \frac{1}{p_x} \sum_y \mathrm{r}_{yx} \ln\left(\sum_z \mathrm{o}_{xy}\mathrm{b}_z\right). \tag{32}$$

In order to write this term as a KL divergence, we introduce two new normalized distributions :

$$\mathrm{r}_x = \frac{\sum_y \mathrm{r}_{yx}}{\sum_{xy} \mathrm{r}_{yx}}, \tag{33}$$

which is acceptable as a distribution provided condition (ii) holds. Similarly, we introduce

$$t_x = \sum_y \mathrm{o}_{xy}\mathrm{b}_y \sum_l \mathrm{r}_{lx}. \tag{34}$$

It is easy to see then that

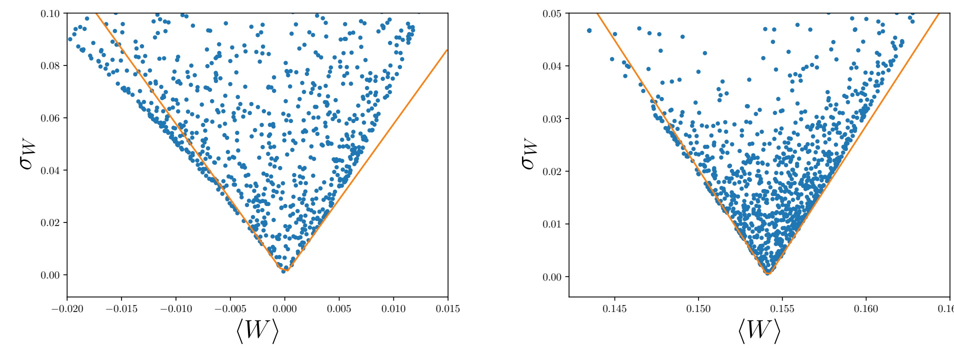
$$\langle \mathrm{q}W \rangle = -\langle \mathrm{q} \rangle D_{KL}(\mathbf{r}||\mathbf{t}) + \langle \mathrm{q} \rangle V. \tag{35}$$

In the end, we obtain the same relation as in Eq. (31), provided one takes into account the new definitions of the distribution q, $\langle W \rangle$ and V, and conditions (i)+(ii) hold.

3.3. Further consequences of the risk-return trade-off

When the odds are fixed, the clouds of points of Figs. 3 change when the probability vector of the horses to win (the $\mathbf{p}$ vector) changes as shown in Fig. 4. Each choice of this vector generates a separate cloud of points,

Fig 3. Pareto plots for fair and super-fair odds in the variables volatility σ_W versus growth rate $\langle W \rangle$. The cloud of points displays an ensemble of feasible random strategies for non-diagonal (a) fair and (b) super-fair odds, where $V > 0$. The solid lines test the inequality of Eq. (31), which is globally valid to the right of the risk-free strategy (where both lines meet) but just locally valid to the left of the risk-free strategy.



and all these clouds of points have the same lowest point in common, namely the risk free strategy, where $\langle W \rangle = V$ and $\sigma_W = 0$ independently of the $\mathbf{p}$ vector. Each cloud of points admits a tangent vector near this risk free strategy with a slope determined by an inequality of the form (31). There is a different slope for each tangent since the slope depends on the $\mathbf{p}$ vector. Now, if some information is known about the family of distributions of (the $\mathbf{p}$ vector), one can combine all these bounds to obtain a general bound on all the possible values of the slopes. Such a global bound would then inform on the minimum level of risk irrespective of the distribution $\mathbf{p}$.

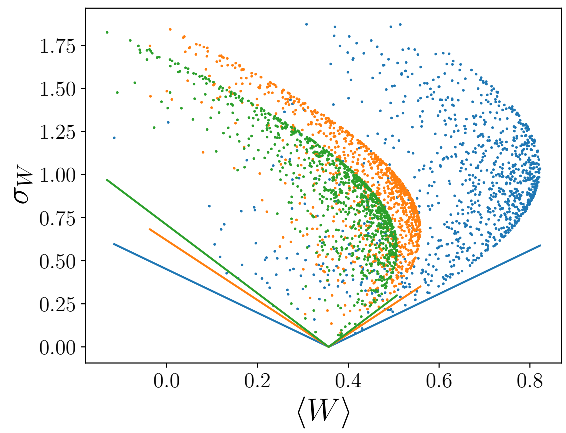
4. Risk quantification beyond volatility

4.1. Extinction probability for geometric brownian motion

Alternative measures of risk beyond volatility are needed because the volatility is symmetric, i.e. it describes positive or negative fluctuations. Therefore, it does not conform to the intuitive notion of risk, which is asymmetric since it is only related to negative fluctuations [23, 31]. To build a more appropriate measure of risk, we turn to a continuous approximation of the trajectory of the log-capital as a geometric Brownian motion.

This corresponds to the asymptotic regime for the central limit theorem of Eq. (6), in which the log-capital is distributed according to a normal law with $\langle W \rangle$ as mean and σ_W as standard deviation. Assuming the log-capital $y(t) = \log C_t$ starts from an initial value y_0, the probability that

Fig 4. Clouds of points from Kelly's horse race model in the plane (σ_W,$\langle W\rangle$) for various probability p vectors. The value of the game is unchanged as it is independent of the **p** vector, here it is $V = 0.36$. The solid lines have the same meaning as in Fig.3.



it reaches the value y at time t, namely $\phi_{y_0}(y,t)$ is:

$$\phi_{y_0}(y,t) = \frac{e^{-\frac{(y-\langle W\rangle t - y_0)^2}{2\sigma_W^2 t}}}{\sqrt{2\pi\sigma_W^2 t}}. \tag{36}$$

Then, the extinction probability is defined as the probability that the log-capital reaches a certain low threshold at any time $t' < t$ for the first time. Further, $\mathcal{P}(t) = 1 - \mathcal{S}(t)$, where $\mathcal{S}(t)$ denotes a survival probability, defined as the probability that the log-capital $y(t)$ did not ever reach the low threshold l *at any time* $t' < t$ assuming that it started with the value y_0 at time 0 with $y_0 > l$.

The survival probability $\mathcal{S}(t)$ can be evaluated from the classic image method. According to this method, one writes the probability $P(y,t)$ for the random walker to reach y at time t as a linear combination of $\phi_{y_0}(y,t)$ and $\phi_m(y,t)$ where m is the position of the image. By enforcing the condition $P(y = l) = 0$ at all times, one finds m and an explicit form for $P(y,t)$. Then, the survival probability is $\mathcal{S}(t) = \int_l^\infty P(y,t)dy$. One obtains

$$\mathcal{S}(t) = -e^{\frac{2\langle W\rangle(l-y_0)}{\sigma_W^2}} + \frac{1}{2}\text{erfc}\left(\frac{\langle W\rangle t + l - y_0}{\sqrt{2t}\sigma_W}\right)e^{\frac{2\langle W\rangle(l-y_0)}{\sigma_W^2}} + \frac{1}{2}\text{erfc}\left(\frac{-\langle W\rangle t + l - y_0}{\sqrt{2t}\sigma_W}\right), \tag{37}$$

where erfc(x) denotes the complementary error function (i.e. erfc(x) $= 1 -$ erf(x), where erf(x) is the error function). It is straightforward to check that $S(0) = 1$. One also finds that $\mathcal{S}(t \to \infty) = 1 - e^{2\langle W\rangle(l-y_0)/\sigma_W^2}$. Therefore, it is a meaningful survival probability provided $\langle W\rangle > 0$ if $l < y_0$ or $\langle W\rangle < 0$ if $l > y_0$. Let us focus on the case $\langle W \rangle > 0$, for which the capital is growing exponentially on long times. The larger $\langle W \rangle > 0$ or the higher the distance between the starting point y_0 and the threshold l, the less likely the log-capital reaches the low threshold, as one would expect. Since a negative fluctuation of the capital is needed to reach this low threshold, such an event can only occur at rather short times because at long times the capital is growing exponentially as illustrated in Fig 5a. Further, it can be shown that the event is guaranteed to occur when $\langle W \rangle \sim \sigma_W^2/(2(l - y_0))$.

From these considerations, an inequality similar to that of Eq. (28) can be derived to describe the mean growth rate-risk trade-off using the extinction probability $\mathcal{P}_{ext} = \mathcal{P}(t \to \infty)$ as a proxy of risk instead of the volatility. From the expression of $\mathcal{S}(t \to \infty)$ above, it is straightforward to obtain in the case of fair odds and when $\langle W \rangle > 0$:

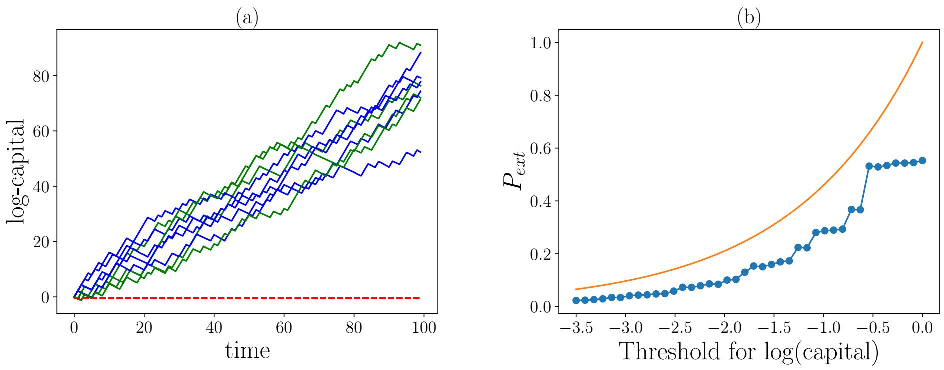
$$\mathcal{P}_{ext} \geq e^{\frac{2\sigma_q^2(l-y_0)}{\langle W \rangle}}, \tag{38}$$

which shows that in order to reduce risk (as measured by extinction probability), one needs to bring the threshold further away from the initial capital as one would expect or *reduce* the growth rate, rather counter-intuitively.

Further characterizations of risk could be considered. For instance, the distribution of first passage times for the log-capital to reach the threshold can be obtained from the opposite of the time derivative of $\mathcal{S}$; and using more advanced arguments, one can also compute analytically the distribution of the time where the log-capital reaches its maximum for an arbitrary value of the drift. This question has been studied in finance because it is related to the optimization of the time to sell/buy a stock [31].

In Fig. 5b, we compare the extinction probability $\mathcal{P}(t)$ for a fixed final time t as function of the threshold value l, for Kelly's horse race and for its approximation using geometric Brownian motion. In the case of Kelly's model, many stochastic trajectories are simulated from the model in the same conditions and from the statistics of these trajectories an empirical estimation of the extinction probability $\mathcal{P}(t)$ is obtained. The simulation results of Kelly's model displays steps, which follow the trend given by the continuous model. The presence of these steps can be traced back to the fact that in Kelly's model the log-capital changes by discrete increments at discrete time intervals. In Fig. 5b, one sees a comparison between the extinction probability evaluated from Eq. (37) using geometric brownian motion with a simulation of that quantity evaluated using Kelly's model. As shown in the figure, the prediction of geometric brownian motion is very close to that of Kelly's model in the left part of the figure where the threshold takes its minimum value. This is expected because in this regime the trajectory contains a large number of steps to reach the threshold and therefore the continuum approximation is well verified. In contrast, this

Fig 5. (a) Trajectories of Kelly's horse race model and (b) comparison between the extinction probability in Kelly's model and in the geometric Brownian motion that matches the parameters of Kelly's model. For figure (a), trajectories either never reach the target (blue solid curves) or do reach it (green solid curves), typically at short times. The threshold is set at $l = -0.5$ (red dashed line). For figure (b), the extinction probability is computed for Kelly's model vs. position of the threshold after 100 races, and in the long time limit for geometric Brownian motion. In both figures, horse probabilities and returns are $p = (0.36, 0.15, 0.49)$ and $r = (0.63, 0.31, 0.06)$.



does not happen on the right part of the figure, where the discreteness of Kelly's model is quite apparent.

4.2. Risk-constrained Kelly gambling

In our first study of risk-constrained Kelly gambling [16], we have introduced a penalization proportional to the volatility in the optimization of Kelly's growth rate with respect to the bet vector. This was done with the following objective function, which interpolates between the maximization of the growth rate and the minimization of the variance of the growth rate :

$$\tilde{J} = \rho \langle W \rangle - (1 - \rho)\sigma_W, \tag{39}$$

with $0 \leq \rho \leq 1$. In this approach, the parameter ρ plays the role of a risk aversion parameter, and the optimal bets are parameterized by it. From these optimal bets, one can build Pareto diagrams that represent the minimum amount of fluctuations for a given growth rate. An example of these Pareto diagrams is shown in Fig. 6a.

Instead of using the volatility to constrain the growth rate in Kelly's gambling, another approach is to introduce a constraint into the optimization of the growth rate to enforce that the extinction probability does not go beyond a certain threshold [24]. As usual, the constraint is taken into account with a Lagrange multiplier. To properly define that approach, it is convenient to introduce:

$$C_{min} = \min_{t=1,2..} C_t, \tag{40}$$

which represents the lowest value reached by the capital C_t during the observed time before it goes on increasing. The drawdown risk is quantified by the probability that this minimum goes below a target value, $P(C_{min} < \alpha)$, where α is the target value for the capital. Then the constraint on the probability of drawdown has the form $P(C_{min} < \alpha) < \beta$ with $\beta \in (0, 1)$. For example, we might take $\alpha = 0.7$ and $\beta = 0.1$, meaning that we require the probability of a drawdown of more than 30% to be less than 10%. This drawdown risk does not have in general a simple form as function of the bet vector, it can only be obtained numerically by solving a non-linear optimization problem with non-linear constraints. While this optimization problem is difficult, Boyd et al. introduced a bound on the drawdown risk that results in a tractable convex constraint [24]. This bound reads as follows:

$$\mathbb{E}\left(\frac{\mathrm{b}_x}{\mathrm{r}_x}\right)^{-\lambda} \leq 1 \quad \Longrightarrow \quad P(C_{min} < \alpha) < \beta, \tag{41}$$

where λ is defined as $\lambda = \ln\beta/\ln\alpha$. This means that, by varying the maximum extinction probability allowed, hence varying λ, our optimization is more or less sensitive to risk. For instance, when $\beta \to 1$ or $\alpha \to 0$, then $\lambda \to 0$ and we have an unconstrained optimization problem.

In the following, we fix the value of α. We consider the case of three horses, with an initial capital $C_0 = 1$, and we use $p = (0.1, 0.2, 0.7)$ and $r = (0.7, 0.1, 0.2)$. With these values, we obtain the optimal strategy $\mathbf{b}^*$ for different $\beta \in (0, 1)$. Once the optimal strategy is obtained for a fixed β, we can compute the growth rate $\langle W \rangle$ and the variance σ_W for that particular strategy. Hence, we obtain the diagram in the coordinates $(\langle W \rangle - \sigma_W)$ shown in Fig. 6a.

In this figure, we observe that the two measures of risk lead to comparable plots. Further, the blue line is always below the red line, which is expected since the blue plot represents the set of points where variance is minimized for a given growth rate. At Kelly's point, both curves meet since this corresponds to the case $\beta = 1$ for which Boyd's approach reduces to the simple optimization of the growth rate as done in Kelly's approach. We have observed that these features are robust with respect to the choice of α. Note that the red curve from Boyd's approach does not reach arbitrary low values of the growth rate because of the choice of the lowest value of β. In Boyd's approach, the null strategy is only reached asymptotically as β approaches zero.

In Fig. 6b we analyze the bound β on the actual extinction probability. Using simulations, we computed the probability of extinction for the optimal bets b^* obtained by Boyd's maximization. These simulations were run for the same parameters considered in Fig. 6a and for two other sets of horse probabilities and returns chosen at random. Finally, we show both the probability of extinction and the corresponding bound β (in matching color) as a function of average growth rate. As apparent from the figure, the probability of extinction is always below its bound, but depending on the parameters chosen, it may be tighter or looser. Although these curves do not represent a Pareto front for the probability of extinction and growth rate, they show that, in general, the probability of extinction increases with growth rate, making Kelly's the riskiest strategy.

We observe sudden increases in the extinction probability as a function of the growth rate at some points. This may be related again to the discreteness of the log-capital as in Fig. 5, where we see steps in probability of extinction when the threshold is modified. In fact, through the optimization procedure, the average gain and the threshold are connected so that one can parametrize the optimal solution with the threshold or the gain as in Fig. 6b.

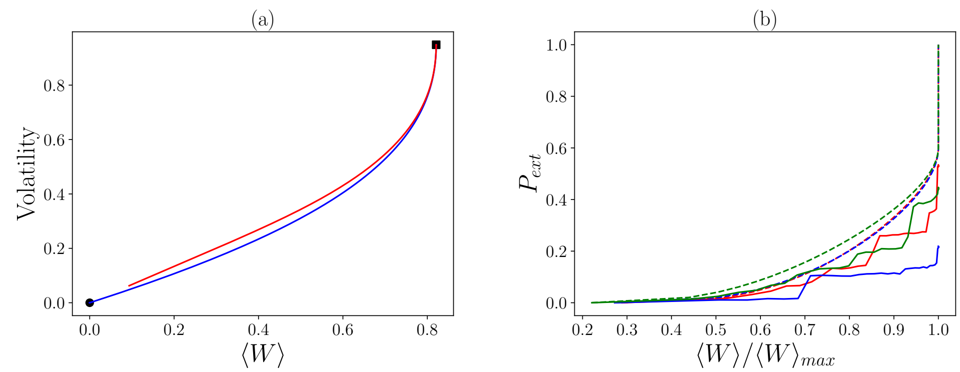
It is interesting to notice that above a certain value of $\beta = \beta^*$, the curve for the bound becomes vertical in Fig. 6b: Kelly' s strategy is always the optimal strategy when the bound imposed on the probability of extinction becomes high enough. This behavior is akin to a phase transition separating an optimal solution which is Kelly's like from a non-Kelly strategy. Indeed, when the constraint $\mathbb{E}\left(\frac{\mathrm{b}_x}{\mathrm{r}_x}\right)^{-\lambda} \leq 1$ is inactive for specific values of p and r, the solution of the optimization is the one without the constraint, i.e. Kelly's solution.

It is easy to check that in the region $\beta \in (\beta^*, 1)$, $\mathbf{E}\left(\left(\frac{b^*_x}{r_x}\right)^{-\lambda}\right) < 1$, for the values of p and r chosen above. There, Kelly's strategy is always optimal in this interval of β values where the probability of extinction is high. Note that the lower end of the vertical line corresponds to the case where the constraint becomes active and Kelly's strategy no longer fulfills the condition, so the optimal strategy then becomes different to Kelly's betting in order to lower the risk.

In some specific cases, the vertical line does not exists ($r_i > p_i \, \forall i$) which corresponds to unfair odds, or the plot shows only the vertical line corresponding to Kelly's regime for all β when $r_i < p_i \, \forall i$.

In the case where the vertical line does not exist, odds are unfair for the gambler, who should avoid Kelly's strategy because it leads to a high extinction probability (which means a high probability of bankruptcy for the gambler). In other words, when conditions are not favorable (in terms of the odds or of the distribution of the probabilities of the environment, gamblers (respectively, biological systems) cannot maintain themselves at Kelly's point except at the cost of a large extinction probability of the population (respectively bankruptcy probability). Instead, in good conditions for growth, Kelly's strategy is optimal.

Fig 6. (a) Pareto front of the volatility versus the average growth rate, and (b) extinction probability as a function of (normalized) growth rate for Boyd's optimization (solid lines). In figure (a), curves are calculated according to the mean-volatility trade-off approach (solid blue line) and volatility vs growth rate line for Boyd's optimization of growth rate with extinction probability constraint (solid red line). Kelly's strategy is shown with a square symbol and the null strategy with the circle in this diagram. In figure (b), the three different sets of parameters are: blue line is computed with the same horse probabilities and returns as in (a), red and green show two other different combinations taken at random. Dashed lines correspond to Boyd's method bound β for extinction probability ($P(C_{min} < \alpha) < \beta$). Each extinction probability curve is bound by the dashed line of matching color. Extinction probability is computed from 40000 simulations of 100 races for each value of the growth rate shown. For both plots $\alpha = \exp(l) = 0.6$



Conclusion

In this article, we have explored an extension of Kelly's gambling model to the case non-diagonal odds, an extension that is particularly relevant for finance or biological applications. For example, in the stock market, the odds matrix that codes for the daily returns from a list of stocks is non-diagonal, and the challenge is to deal with the day-to-day randomness in the daily returns themselves.

We also found that when the game is not fully mixing for certain environmental probabilities, it can be reduced to a smaller game, known as the 'essential part of the game' [28], which is fully mixing. This method is broadly interesting because it allows us to break the complexity of the initial problem into the study of an issue of reduced complexity without affecting the optimal strategies.

For a generalized Kelly gambling model, we have studied the trade-off between the average growth rate and volatility, which is known in the financial literature as a risk-return trade-off. We have also explored an alternate measure of risk beyond volatility, namely the extinction probability, which can easily be calculated if the races are uncorrelated as a realization of geometric Brownian motion.

Our main result is that this measure of risk leads to comparable results as obtained with the volatility as far as the risk-return trade-off is concerned. In particular, the inequality that embodies the trade-off between average growth rate and volatility can be expressed similarly as an inequality in which the extinction probability replaces the volatility. Given the derivation of these inequalities, we expect that they should hold in a broader context whenever a multiplicative process can characterize growth. In fact, it is indeed the case, Cohen and Gillespie [27, 32] both used a multiplicative process to describe population growth in random environments and observed that adding a random element to the number of offsprings of a particular genotype leads to a lower fitness as measured by a geometrical average. Their conclusions thus fully agree with the predictions from Kelly's model, although they fail to capture the beneficial side to fluctuations, which requires the alternate measures of risk mentioned above.

One area of application of our paper is evolutionary games in ecology, in which various forms of bet-hedging strategies have been considered together with their associated trade-offs: one example is the trade-off between egg size and number for birds [33] and another one is the emergence of cooperative breeding [34], which arise as consequences of the need to cope with environmental variations. Another central question in ecology is what determines the diversity of species and the coexistence between species. Biodiversity is regarded as a form of biological insurance against disruptive effects of the environment because biodiversity reduces the variability in ecosystem properties that arise due to differential responses of species to environmental variations [35]. This work again supports the idea of an ecological trade-off similar to the trade-off between growth and risk in economy and finance.

This trade-off shows that species persistence cannot be decided solely based on growth rate, because fluctuations matter in coexistence theory models. These observations have been confirmed by comprehensive studies in ecology, which underlined the role of the fluctuations of species abundances [36] and of fluctuations of the population growth rate [37]. In the end, we note that in all these works including ours, the ratio between the growth rate and the standard deviation of the fluctuations of the growth rate emerges as a central quantity both in the ecological context and in the finance field, where it is known under the name of Sharpe ratio.

Acknowledgments

DL acknowledges support from (ANR-11-LABX-0038, ANR-10-IDEX-0001-02). LD acknowledges support from Spanish Ministerio de Ciencia e Innovación through Grants PID2020-113455GB-I00 and PID2023-147067NB-I00 and Universidad Complutense de Madrid through "Convocatoria plurianual para la recualificación del Sistema

Universitario Español para 2021-2023 (MV24/21)" funded by NextGenerationEU. RP is supported by the Israeli Science Foundation (Grant No. 776/19).

We acknowledge fruitful discussions with Luca Peliti and Haim Permuter.

References

1. S. C. E., A mathematical theory of communication, Bell System Technical Journal (1948) 379–423.

2. J. L. Kelly, A new interpretation of information rate, IRE Transactions on Information Theory 2 (3) (1956) 185–189.

3. T. M. Cover, J. A. Thomas, Elements of Information Theory, Wiley Interscience, 2005.

4. L. C. MacLean, E. O. Thorp, W. T. Ziemba, The Kelly Capital Growth Investment Criterion: Theory and Practice, Vol. 3, Word Scientific, 2011.

5. D. G. Luenberger, Investment Science, Oxford Universily Press. Inc, 1998.

6. A. V. Proskurnikov, B. R. Barmish, On the benefit of nonlinear control for robust logarithmic growth: Coin flipping games as a demonstration case, IEEE Control Systems Letters (2023).

7. M. C. Donaldson-Matasci, C. T. Bergstrom, M. Lachmann, The fitness value of information, Oikos 119 (2) (2010) 219–230.

8. E. Kussell, S. Leibler, Ecology: Phenotypic diversity, population growth, and information in fluctuating environments, Science 309 (5743) (2005) 2075–2078.

9. M. Thattai, A. van Oudenaarden, Stochastic gene expression in fluctuating environments (January 2004).

10. E. Levien, J. Min, J. Kondev, A. Amir, Non-genetic variability: survival strategy or nuisance?, Rep. Prog. Phys. 84 (11) (2020) 116601.

11. D. Lacoste, O. Rivoire, D. S. Tourigny, Cell behavior in the face of uncertainty (2023).

12. N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, S. Leibler, Bacterial persistence as a phenotypic switch, Science 305 (5690) (2004) 1622–1625.

13. S. Maslov, K. Sneppen, Well-temperate phage: optimal bet-hedging against local environmental collapses, Sci. Rep. 5 (2015) 10523.

14. D. L. Venable, Bet hedging in a guild of desert annuals, Ecology 88 (5) (2007) 1086–1090.

15. A. Măgălie, D. A. Schwartz, J. T. Lennon, J. S. Weitz, Optimal dormancy strategies in fluctuating environments given delays in phenotypic switching, Journal of Theoretical Biology 561 (2023) 111413.

16. L. Dinis, J. Unterberger, D. Lacoste, Phase transitions in optimal betting strategies, EPL (Europhysics Letters) 131 (6) (2020) 1–23.

17. L. Dinis, J. Unterberger, D. Lacoste, Pareto-optimal trade-off for phenotypic switching of populations in a stochastic environment, Journal of Statistical Mechanics: Theory and Experiment 2022 (5) (2022) 053503.

18. P. G. Hufton, Y. T. Lin, T. Galla, A. J. McKane, Intrinsic noise in systems with switching environments, Phys. Rev. E 93 (5) (2016) 052119, publisher: American Physical Society.

19. L. P. A. Despons, D. Lacoste, Adaptive strategies in Kelly's horse races model, Journal of Statistical Mechanics: Theory and Experiment 2022 (2022) 093405.

20. R. Pugatch, N. Barkai, T. Tlusty, Asymptotic Cellular Growth Rate as the Effective Information Utilization Rate (2013).

21. O. Rivoire, S. Leibler, The Value of Information for Populations in Varying Environments, Journal of Statistical Physics 142 (6) (2011) 1124–1166.

22. O. Tal, T. D. Tran, Adaptive Bet-Hedging Revisited: Considerations of Risk and Time Horizon, Bull Math Biol 82 (4) (2020) 50.

23. S. Maslov, Y.-C. Zhang, Probability distribution of drawdowns in risky investments, Physica A: Statistical Mechanics and its Applications 262 (1) (1999) 232–241.

24. R. E. K. Busseti, Enzo, S. Boyd, Risk-constrained Kelly gambling, The Journal of Investing 25 (3) (2016) 118–134.

25. O. Peters, M. Gell-Mann, Evaluating gambles using dynamics, Chaos: An Interdisciplinary Journal of Nonlinear Science 26 (2) (2016) 023103.

26. O. Peters, The ergodicity problem in economics, Nature Physics 15 (12) (2019) 1216–1221.

27. D. Cohen, Optimizing reproduction in a randomly varying environment, Journal of Theoretical Biology 12 (1) (1966) 119–129.

28. E. N. Barron, Game Theory, an introduction, 2nd Edition, Wiley. Loyola University, Chicago, January 2010.

29. P. Smoczynski, D. Tomkins, An explicit solution to the problem of optimizing the allocations of a bettor's wealth when wagering on horse races, The Mathematical Scientist 35 (01 2010).

30. L. Ziyin, M. Ueda, Universal thermodynamic uncertainty relation in nonequilibrium dynamics, Phys. Rev. Res. 5 (2023) 013039.

31. S. N. Majumdar, J.-P. Bouchaud, Optimal time to sell a stock in the black–scholes model: comment on 'thou shalt buy and hold', by A. Shiryaev, Z. Xu and X.Y. Zhou, Quantitative Finance 8 (8) (2008) 753–760.

32. J. H. Gillespie, Natural selection for variances in offspring numbers: A new evolutionary principle, The American Naturalist 111 (981) (1977) 1010–1014.

33. M. Rees, C. Jessica, E. Metcalf, D. Z. Childs, Bet-hedging as an evolutionary game: the trade-off between egg size and number, Proceedings of the Royal Society B: Biological Sciences 277 (1685) (2010) 1149–1151.

34. D. R. Rubenstein, Spatiotemporal environmental variation, risk aversion, and the evolution of cooperative breeding as a bet-hedging strategy, Proceedings of the National Academy of Sciences 108 (supplement_2) (2011) 10816–10822.

35. M. Loreau, M. Barbier, E. Filotas, D. Gravel, F. Isbell, S. J. Miller, J. M. Montoya, S. Wang, R. Aussenac, R. Germain, P. L. Thompson, A. Gonzalez, L. E. Dee, Biodiversity as insurance: from concept to measurement and application, Biological Reviews 96 (5) (2021) 2333–2354.

36. J. Pande, T. Fung, R. Chisholm, N. M. Shnerb, Mean growth rate when rare is not a reliable metric for persistence of species, Ecology Letters 23 (2) (2020) 274–282.

37. J. Pande, Y. Tsubery, N. M. Shnerb, Quantifying invasibility, Ecology Letters 25 (8) (2022) 1783–1794.