Internal Regret in On-line Portfolio Selection *

Gilles Stoltz (gilles.stoltz@ens.fr) †
Département de Mathématiques et Applications,
Ecole Normale Supérieure,
75005 Paris, France

Gábor Lugosi (lugosi@upf.es) ‡
Department of Economics,
Pompeu Fabra University,
08005 Barcelona, Spain

Abstract. This paper extends the game-theoretic notion of internal regret to the case of on-line potfolio selection problems. New sequential investment strategies are designed to minimize the cumulative internal regret for all possible market behaviors. Some of the introduced strategies, apart from achieving a small internal regret, achieve an accumulated wealth almost as large as that of the best constantly rebalanced portfolio. It is argued that the low-internal-regret property is related to stability and experiments on real stock exchange data demonstrate that the new strategies achieve better returns compared to some known algorithms.

1. Introduction

The problem of sequential portfolio allocation is well-known to be closely related to the on-line prediction of individual sequences under expert advice, see, for example, Cover (1991), Cover and Ordentlich (1996), Helmbold, Schapire, Singer, and Warmuth (1998), Ordentlich and Cover (1998), Blum and Kalai (1999), Cesa-Bianchi and Lugosi (2000). In the on-line prediction problem the goal is to minimize the predictor's cumulative loss with respect to the best cumulative loss in a pool of "experts". In a certain equivalent game-theoretic formulation of the problem, this is the same as minimizing the predictor's *external regret*, see Foster and Vohra (1999). External regret measures the difference between the predictor's cumulative loss and that of the best expert. However, another notion of regret, called *internal regret* in Foster and Vohra (1999) has also been in the focus of attention mostly in the theory of playing repeated games, see Foster and Vohra (1998; 1999), Fudenberg and Levine (1999), Hart and Mas-

* An extended abstract appeared in the *Proceedings of the 16th Annual Conference on Learning Theory and 7th Kernel Workshop*, Springer, 2003.

† The work of the first author was supported by PAI Picasso grant 02543RM and by the French CNRS research network AS66 (SVM and kernel algorithms).

‡ The work of the second author was supported by DGI grant BMF2000-0807.

Colell (2000; 2001), Cesa-Bianchi and Lugosi (2003). Roughly speaking, a predictor has a small internal regret if for each pair of experts (i, j), the predictor does not regret of not having followed expert i each time it followed expert j. It is easy to see that requiring a small internal regret is a more difficult problem since a small internal regret in the prediction problem implies small external regret as well. A brief summary of the basic properties is given below.

The goal in the sequential investment problem is to distribute one's capital in each trading period among a certain number of stocks such that the total achieved wealth is almost as large as the wealth of the largest in a certain class of investment strategies. This problem, known as the minimization of the worst-case logarithmic wealth ratio, is easily seen to be the generalization of an external regret minimization problem in the "expert" setting under the logarithmic loss function. The main purpose of this paper is to extend the notion of internal regret to the sequential investment problem, understand its relationship to the worst-case logarithmic wealth ratio, and design investment strategies minimizing this new notion of regret. The definition of internal regret given here has a natural interpretation and the investment strategies designed to minimize it have several desirable properties both in theory and in the experimental study described in the Appendix.

The paper is organized as follows. In Sections 2 and 3 we briefly summarize the sequential prediction problem based on expert advice and describe the notions of internal and external regrets. In Section 4 the sequential portfolio selection problem is described, and basic properties of Cover's universal portfolio and the EG investment strategy are discussed. In Section 5 we introduce the notion of internal regret for sequential portfolio selection, and describe some basic properties. In Section 6 new investment strategies are presented aiming at the minimization of the internal regret. Finally, in Section 7 the notion of internal regret is generalized for an uncountable class of investment strategies and an algorithm inspired by Cover's universal portfolio is proposed which minimizes the new notion of internal regret.

2. Sequential prediction: external regret

In the (randomized) sequential prediction problem the predictor, at each time instance $t = 1, 2, \ldots$, chooses a probability distribution $\mathbf{P}_t = (P_{1,t}, \ldots, P_{N,t})$ over the set $\{1, 2, \ldots, N\}$ of experts. After the choice is made, expert i suffers loss $\ell_{i,t}$, and the predictor suffers loss

$$\ell_t(\mathbf{P}_t) = \sum_{i=1}^{N} P_{i,t}\ell_{i,t} \ .$$

This loss may be interpreted as the expected loss if the predictor chooses an expert randomly, according to the distribution $\mathbf{P}_t$, and predicts as the selected expert's advice. The *external regret* of the predictor, after n rounds of play, is

$$\sum_{t=1}^{n} \ell_t(\mathbf{P}_t) - \min_{i=1,\dots,N} \sum_{t=1}^{n} \ell_{i,t} = \max_{i=1,\dots,N} \sum_{j=1}^{N} \sum_{t=1}^{n} P_{j,t}(\ell_{j,t} - \ell_{i,t}) \ .$$

If this external regret is $o(n)$ uniformly over all possible values of the losses, then the corresponding predictor is said to suffer no external regret. This problem has been extensively studied since Hannan (1957) and Blackwell (1956).

For example, it is well known (see, e.g., Cesa-Bianchi and Lugosi (1999)) that if the losses $\ell_{i,t}$ are all bounded between zero and $B > 0$, then the exponentially weighted average predictor defined, for $t = 1, 2, \dots$, by

$$P_{i,t+1} = \frac{\exp(-\eta \sum_{s=1}^{t} \ell_{i,s})}{\sum_{j=1}^{N} \exp(-\eta \sum_{s=1}^{t} \ell_{j,s})}$$

has an external regret bounded by

$$\frac{\ln N}{\eta} + n\frac{\eta}{8}B^2 = B\sqrt{(n/2)\ln N} \ , \qquad (1)$$

with the (optimal) choice $\eta = B^{-1}\sqrt{8 \ln N/n}$. The tuning parameter η can be set optimally only when the time length n is known in advance. However, we recall a simple modification of the exponentially weighted average algorithm, proposed by Auer, Cesa-Bianchi, and Gentile (2002), which does not need to know n in advance.

A natural adaptive version of the optimal parameter η determined in the case of known time length is formed by defining the tuning parameter at round t by $\eta_t = B^{-1}\sqrt{8 \ln N/t}$. Now, the exponentially weighted average forecaster with time-varying tuning parameter predicts, at rounds $t = 1, 2, \dots$, with

$$P_{i,t+1} = \frac{\exp\left(-\eta_{t+1} L_{i,t}\right)}{\sum_{j=1}^{N} \exp\left(-\eta_{t+1} L_{j,t}\right)} \ ,$$

where $L_{i,t} = \sum_{s=1}^{t} \ell_{i,s}$. Denote the (expected) cumulative loss of the algorithm by $\widehat{L}_n = \sum_{t=1}^{n} \ell_t(\mathbf{P}_t)$. The following result is proved in Auer, Cesa-Bianchi, and Gentile (2002).

Theorem 1. The exponentially weighted average forecaster with time-varying tuning parameter achieves, uniformly over all possible values of the losses $\ell_{i,t} \in [0, B]$,

$$\widehat{L}_n - \min_{i=1,\dots,N} L_{i,n} \le B \left(2\sqrt{\frac{n}{2}\ln N} + \sqrt{\frac{\ln N}{8}} \right) .$$

This result will be used in Section 6.1 to define an investment algorithm which does not need to know in advance the trading time length. A whole family of predictors with performance guarantees similar to those of the exponentially weighted forecaster may be defined, see, for example, Cesa-Bianchi and Lugosi (2003). Some of them do not require the knowledge of the time length, as is the case of the polynomial forecaster described below. Nevertheless, it is important to design a time-adaptive version of the exponentially weighted forecaster, for the latter is a popular method, usually achieving good results in practical situations (see also our experimental results in the Appendix).

An important class of "polynomial" forecasters are those of the form

$$P_{i,t+1} = \frac{\left(\sum_{s=1}^{t} \ell_s(\mathbf{P}_s) - \ell_{i,s}\right)_+^{p-1}}{\sum_{j=1}^{N}\left(\sum_{s=1}^{t} \ell_s(\mathbf{P}_s) - \ell_{j,s}\right)_+^{p-1}} .$$

where $p \ge 1$ and $(x)_+ = \max\{x, 0\}$ denotes the nonnegative part of the real number x.

These forecasters satisfy the following bound, see Cesa-Bianchi and Lugosi (2003).

Theorem 2. The polynomial forecaster with $p \ge 1$ achieves, uniformly over all possible values of the losses $\ell_{i,t} \in [0, B]$,

$$\widehat{L}_n - \min_{i=1,\dots,N} L_{i,n} \le B\sqrt{(p-1)nN^{2/p}} .$$

3. Sequential prediction: internal regret

3.1. Definition of the internal regret

The definition of external regret is based on the comparison to an external pool of strategies, the ones given by each expert. In the definition of the internal regret one is interested in modifications of the predictor's strategy obtained by replacing the action of the forecaster by expert j each time it chooses expert i. This is equivalent to selecting an expert according to the distribution $\mathbf{P}_t^{i\to j}$ obtained from $\mathbf{P}_t$ by putting probability mass 0 on i and $P_{i,t}+P_{j,t}$ on j. This transformation is called the $i \to j$ *modified strategy.*

We require that none of these modified strategies is much better than the original strategy, that is, we seek strategies such that the difference between their (expected) cumulative loss and that of the best modified strategy is small. Thus,

$$\sum_{t=1}^{n} \ell_t(\mathbf{P}_t) - \min_{i,j\in\{1,\ldots,N\}} \sum_{t=1}^{n} \ell_t(\mathbf{P}_t^{i\to j})$$

should be as small as possible. This quantity is called the *internal regret* of the sequential predictor $\mathbf{P}_t$. The internal regret may be re-written as

$$\max_{i,j\in\{1,\ldots,N\}} \sum_{t=1}^{n} r_{(i,j),t}$$

where $r_{(i,j),t} = P_{i,t}(\ell_{i,t} - \ell_{j,t})$. Thus, $r_{(i,j),t}$ expresses the predictor's regret of having put the probability mass $P_{i,t}$ on the i-th expert instead of on the j-th one, and

$$R_{(i,j),n} = \sum_{t=1}^{n} r_{(i,j),t} = \sum_{t=1}^{n} P_{i,t}(\ell_{i,t} - \ell_{j,t})$$

is the corresponding cumulative regret. Similarly to the case of the external regret, if this quantity is uniformly $o(n)$ over all possible values of the losses, then the corresponding predictor is said to exhibit no internal regret.

Now clearly, the external regret of the predictor $\mathbf{P}_t$ equals

$$\max_{j=1,\ldots,N} \sum_{i=1}^{N} R_{(i,j),n} \le N \max_{i,j\in\{1,\ldots,N\}} R_{(i,j),n} \tag{2}$$

which shows that any algorithm with a small (i.e., sublinear in n) internal regret also has a small external regret. On the other hand, it is easy to see that a small external regret does not imply small internal regret. In fact, as it is shown in the next example, even the exponential weighted average algorithm defined above may have a linearly growing internal regret.

Example 1. (Weighted average predictor has a large internal regret.) Consider the following example with three experts, A, B, and C. Let n be a large multiple of 3 and assume that time is divided in three equally long regimes, characterized by a constant loss for each expert. These

Table I. The losses for Example 1.

Regimes	$\ell_{A,t}$	$\ell_{B,t}$	$\ell_{C,t}$
$1 \le t \le n/3$	0	1	5
$n/3+1 \le t \le 2n/3$	1	0	5
$2n/3+1 \le t \le n$	2	1	0

losses are summarized in Table I. We claim that the regret $R_{(B,C),n}$ of B versus C grows linearly with n, that is,

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} P_{B,t}\left(\ell_{B,t} - \ell_{C,t}\right) = \gamma > 0 \ ,$$

where

$$P_{B,t} = \frac{e^{-\eta L_{B,t}}}{e^{-\eta L_{A,t}} + e^{-\eta L_{B,t}} + e^{-\eta L_{C,t}}}$$

denotes the weight assigned by the exponential weighted average predictor to expert B, where $L_{i,t} = \sum_{s=1}^{t} \ell(i,s)$ denotes the cumulative loss of expert i and η is chosen to minimize the external regret, that is, $\eta = (1/5)\sqrt{(8\ln 3)/n} = 1/(K\sqrt{n})$ with $K = 5/\sqrt{8\ln 3}$. (Note that the same argument leads to a similar lower bound for $\eta = \gamma/\sqrt{n}$, where $\gamma > 0$ is any constant.) The intuition behind this example is that at the end of the second regime the predictor quickly switches from A to B, and the weight of expert C can never recover because of its disastrous behavior in the first two regimes. But since expert C behaves much better than B in the third regime, the weighted average predictor will regret of not having followed the advice of C each time it followed B.

More precisely, we show that during the first two regimes, the number of times when $P_{B,t}$ is more than ε is of the order of $\sqrt{n}$ and that, in the third regime, $P_{B,t}$ is always more than a fixed constant (1/3, say). This is illustrated in Figure 1. In the first regime, a sufficient condition for $P_{B,t} \le \varepsilon$ is that $e^{-\eta L_{B,t}} \le \varepsilon$. This occurs whenever $t \ge t_0 = K(-\ln\varepsilon)\sqrt{n}$. For the second regime, we lower bound the time instant t_1 when $P_{B,t}$ gets larger than ε. To this end, note that $P_{B,t} \ge \varepsilon$ implies

$$(1-\varepsilon)e^{-\eta L_{B,t}} \ge \varepsilon\left(e^{-\eta L_{A,t}} + e^{-\eta L_{C,t}}\right) \ge \varepsilon e^{-\eta L_{A,t}} \ ,$$

which leads to $t_1 \ge \frac{2n}{3} + K\left(\ln\frac{\varepsilon}{1-\varepsilon}\right)\sqrt{n}$. Finally, in the third regime, we have at each time instant $L_{B,t} \le L_{A,t}$ and $L_{B,t} \le L_{C,t}$, so that $P_{B,t} \ge 1/3$. Putting these three steps together, we obtain the following

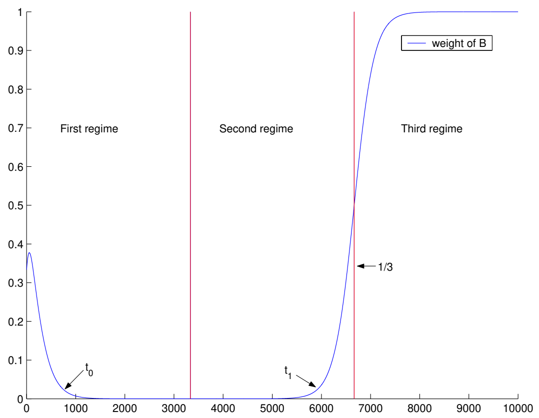

Figure 1. The evolution of the weight assigned to B in Example 1 for $n = 10000$.

lower bound for the internal regret of B versus C:

$$\sum_{t=1}^{n} P_{B,t}\left(\ell_{B,t} - \ell_{C,t}\right) \geq \frac{n}{9} - 5\left(\frac{2n}{3}\varepsilon + K\left(\ln\frac{1-\varepsilon}{\varepsilon^2}\right)\sqrt{n}\right) \ ,$$

which is of the order n, for a sufficiently small $\varepsilon > 0$.

3.2. A general way to design internal regret minimizing algorithms

The example above shows that special algorithms need to be designed to guarantee a small internal regret. Indeed, such predictors exist, as was shown by Foster and Vohra (1998), see also Fudenberg and Levine (1999), Hart and Mas-Colell (2000; 2001). Here we briefly give a new insight on predictors studied in Cesa-Bianchi and Lugosi (2003) (see the remark at the end of this section), and based on Hart and Mas-Colell (2001), as well as a new, simple analysis of their performance guarantees.

Consider the case of sequential prediction under expert advice, with N experts. At round t, the forecaster has already chosen the probability distributions $\mathbf{P}_1, \ldots, \mathbf{P}_{t-1}$. We define $N(N-1)$ fictitious experts, indexed by pairs of integers $i \neq j$, by their losses at time instants $1 \leq s \leq t-1$, which equal $\ell_s(\mathbf{P}_s^{i \to j})$.

Define now a probability distribution $\boldsymbol{\Delta}_t$ over the pairs $i \neq j$ by running one of the algorithms of Section 2 on this pool of fictitious experts, and choose $\mathbf{P}_t$ such that the fixed point equality

$$\mathbf{P}_t = \sum_{(i,j):i\neq j} \Delta_{(i,j),t} \mathbf{P}_t^{i\to j} \ , \tag{3}$$

holds. (We say that $\mathbf{\Delta}_t$ *induces* $\mathbf{P}_t$.) The existence and the practical computation of such a $\mathbf{P}_t$ is an application of Lemma 1 below.

For instance, $\mathbf{\Delta}_t = (\Delta_{(i,j),t})_{i \neq j}$ may be given by

$$\Delta_{(i,j),t} = \frac{\exp\left(-\eta \sum_{s=1}^{t-1} \ell_s(\mathbf{P}_s^{i \to j})\right)}{\sum_{(k,l):k \neq l} \exp\left(-\eta \sum_{s=1}^{t-1} \ell_s(\mathbf{P}_s^{k \to l})\right)} ,$$

tuned, as suggested by the theory, with $\eta = 4B^{-1}\sqrt{\ln N/n}$ in case of known time horizon n.

Indeed, this choice of η and the application of the bound (1) (with $N(N-1)$ upper bounded by N^2) lead to

$$\sum_{t=1}^{n} \sum_{i \neq j} \Delta_{(i,j),t} \ell_t(\mathbf{P}_t^{i \to j}) \leq \min_{i \neq j} \sum_{t=1}^{n} \ell_t(\mathbf{P}_t^{i \to j}) + B\sqrt{n \ln N} ,$$

that is, recalling the fixed point equality (3), the cumulative internal regret of the above strategy is bounded by

$$\max_{i \neq j} R_{(i,j),n} \leq B\sqrt{n \ln N} .$$

Note that this improves the bound given in Corollary 8 of Cesa-Bianchi and Lugosi (2003), by a factor of two.

The same analysis can be carried over for the polynomial forecasters or the time-adaptive version of the exponentially weighted forecaster, using Theorems 1 and 2, and is summarized in the following theorem.

Theorem 3. The above exponentially weighted predictor achieves, uniformly over all possible values of the losses $\ell_{i,t} \in [0, B]$,

$$\max_{i \neq j} R_{(i,j),n} \leq B\sqrt{n \ln N}.$$

With a time-adaptive tuning parameter the upper bound becomes

$$\max_{i \neq j} R_{(i,j),n} \leq B \left(2\sqrt{n \ln N} + \frac{\sqrt{\ln N}}{2} \right) .$$

Finally, with a polynomial predictor of order $p \geq 1$,

$$\max_{i \neq j} R_{(i,j),n} \leq B\sqrt{(p-1)nN^{4/p}} .$$

Remark. The conversion trick illustrated above is a general trick which extends to any weighted average predictor, that is, to any predictor which, at each round, maintains one weight per expert. More precisely, any weighted average predictor whose external regret is small may be converted into a strategy whose internal regret remains small. This will be illustrated for convex loss functions in Section 6.1 and for exp-concave ones in Section 7.1 and 7.2. Note also that in the case of randomized prediction under expert advice Blum and Mansour (2004) propose a different conversion trick, with about the same algorithmic complexity.

Such tricks are valuable to extend results in an effortless way from the case of no external to no internal regret, like the time-adaptive exponentially weighted average predictor suited for the minimization of internal regret proposed by Theorem 3.

It only remains to see the existence and the way to compute a fixed point of the equality (3). The following lemma proposes a more general result, needed for subsequent analysis in Section 7.1. The meaning of this result is that each probability distribution over the expert pairs induces naturally a probability distribution over the experts.

Lemma 1. Let $\mathbf{Q}$ be a probability distributions over the N experts. For all probability distributions $\boldsymbol{\Delta}$ over the pairs of different experts $i \neq j$ and $\alpha \in [0,1]$, there exists a probability distribution $\mathbf{P}$ over the experts such that

$$\mathbf{P} = (1-\alpha) \sum_{i \neq j} \Delta_{(i,j)} \mathbf{P}^{i \to j} + \alpha \mathbf{Q} \ .$$

Moreover, $\mathbf{P}$ may be easily computed by a Gaussian elimination over a simple $N \times N$ matrix.

Proof. The equality

$$\mathbf{P} = (1-\alpha) \sum_{i \neq j} \Delta_{(i,j)} \mathbf{P}^{i \to j} + \alpha \mathbf{Q}$$

means that for all $m \in \{1, \ldots, N\}$,

$$P_m = (1-\alpha) \sum_{i \neq j} \Delta_{(i,j)} P_m^{i \to j} + \alpha Q_m \left(\sum_{j=1}^{N} P_j \right) \ ,$$

or equivalently,

$$\left(\alpha \left(1 - Q_m\right) + (1-\alpha) \sum_{j \neq m} \Delta_{(m,j)} \right) P_m = \sum_{i \neq m} \big((1-\alpha) \Delta_{(i,m)} + \alpha Q_m \big) \, P_i \ ,$$

that is, $\mathbf{P}$ is an element of the kernel of the matrix A defined by

- if $i \neq m$, $A_{m,i} = w_{m,i}$,
- $A_{m,m} = -\sum_{j \neq m,\ 1 \leq j \leq N} w_{j,m}$,

where, for $i \neq m$,

$$w_{m,i} = (1-\alpha)\Delta_{(i,m)} + \alpha Q_m \ .$$

The elements of A have a modulus less than 1. An element of the kernel of A is a fixed point of the matrix $S = A + I_N$, where I_N is the $N \times N$ identity matrix. But S is a column stochastic matrix (its columns are probability distributions), and thus admits a probability distribution $\mathbf{P}$ as a fixed point.

Foster and Vohra (1999) suggest a Gaussian elimination method over A for the practical computation of $\mathbf{P}$.

Remark. Cesa-Bianchi and Lugosi (2003) show that, writing $\mathbf{r}_t$ for the $N(N-1)$-vector with components $r_{(i,j),t}$ and $\mathbf{R}_t = \sum_{s=1}^{t} \mathbf{r}_s$, any predictor satisfying the so-called "Blackwell condition"

$$\nabla\Phi(\mathbf{R}_{t-1}) \cdot \mathbf{r}_t \leq 0 \tag{4}$$

for all $t \geq 1$, with Φ being either an exponential potential

$$\Phi(\mathbf{u}) = \sum_{i=1}^{N} \exp\left(\eta u_i\right) \ ,$$

with η possibly depending on t (when time-adaptive versions are considered) or a polynomial potential

$$\Phi(\mathbf{u}) = \sum_{i=1}^{N} (u_i)_+^p \ ,$$

has the performance guarantees given by Theorem 3.

But the choice (3) ensures that the Blackwell condition is satisfied with an equality, as

$$\nabla\Phi(\mathbf{R}_{t-1}) \cdot \mathbf{r}_t$$
$$= \sum_{i=1}^{N} \ell_{i,t} \left(\sum_{j=1,\ldots,N,\ j \neq i} \nabla_{(i,j)}\Phi(\mathbf{R}_{t-1}) P_{i,t} - \sum_{j=1,\ldots,N,\ j \neq i} \nabla_{(j,i)}\Phi(\mathbf{R}_{t-1}) P_{j,t} \right)$$

(see, e.g., Cesa-Bianchi and Lugosi (2003) for the details), which equals 0 as soon as

$$\sum_{j=1,\ldots,N,\ j \neq i} \nabla_{(i,j)}\Phi(\mathbf{R}_{t-1}) P_{i,t} - \sum_{j=1,\ldots,N,\ j \neq i} \nabla_{(j,i)}\Phi(\mathbf{R}_{t-1}) P_{j,t} = 0$$

for all $i = 1, \ldots, N$. The latter set of equations may be seen to be equivalent to (3), with the choice

$$\Delta_{(i,j),t} = \frac{\nabla_{(i,j)} \Phi(\mathbf{R}_{t-1})}{\sum_{k \neq l} \nabla_{(k,l)} \Phi(\mathbf{R}_{t-1})} ,$$

which was indeed the probability distribution proposed by the conversion trick introduced at the beginning of this section.

4. Sequential portfolio selection

In this section we describe the problem of sequential portfolio selection, recall some previous results, and take a new look at the EG strategy of Helmbold, Schapire, Singer, and Warmuth (1998).

A *market vector* $\mathbf{x} = (x_1, \ldots, x_N)$ for N assets is a vector of nonnegative numbers representing price relatives for a given trading period. In other words, the quantity $x_i \geq 0$ denotes the ratio of closing to opening price of the i-th asset for that period. Hence, an initial wealth invested in the N assets according to fractions $Q_1, \ldots, Q_N$ multiplies by a factor of $\sum_{i=1}^{N} x_i Q_i$ at the end of period. The market behavior during n trading periods is represented by a sequence $\mathbf{x}_1^n = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ of market vectors. $x_{j,t}$, the j-th component of $\mathbf{x}_t$, denotes the factor by which the wealth invested in asset j increases in the t-th period. We denote the probability simplex in $\mathbb{R}^N$ by $\mathcal{X}$.

An *investment strategy* Q for n trading periods consists in a sequence $\mathbf{Q}_1, \ldots, \mathbf{Q}_n$ of vector-valued functions $\mathbf{Q}_t : (\mathbb{R}_+^N)^{t-1} \to \mathcal{X}$, where the i-th component $Q_{i,t}(\mathbf{x}_1^{t-1})$ of the vector $\mathbf{Q}_t(\mathbf{x}_1^{t-1})$ denotes the fraction of the current wealth invested in the i-th asset at the beginning of the t-th period based on the past market behavior $\mathbf{x}_1^{t-1}$. We use

$$S_n(Q, \mathbf{x}_1^n) = \prod_{t=1}^{n} \left(\sum_{i=1}^{N} x_{i,t} Q_{i,t}(\mathbf{x}_1^{t-1}) \right)$$

to denote the wealth factor of strategy Q after n trading periods.

The simplest examples of investment strategies are the so called *buy-and-hold* strategies. A buy-and-hold strategy simply distributes its initial wealth among the N assets according to some distribution $\mathbf{Q}_1 \in \mathcal{X}$ before the first trading period, and does not trade anymore, which amounts to investing, at day t and for $1 \leq i \leq N$, as

$$Q_{i,t}(\mathbf{x}_1^{t-1}) = \frac{Q_{i,1} \prod_{s=1}^{t-1} x_{i,s}}{\sum_{k=1}^{N} Q_{k,1} \prod_{s=1}^{t-1} x_{k,s}} .$$

The wealth factor of such a strategy, after n periods, is simply

$$S_n(Q, \mathbf{x}_1^n) = \sum_{j=1}^{N} Q_{j,1} S_n(j) \ ,$$

where

$$S_n(j) = \prod_{t=1}^{n} x_{j,t}$$

is the accumulated wealth of stock j. Clearly, the wealth factor of any buy-and-hold strategy is at most as large as the gain $\max_{j=1,\ldots,N} S_n(j)$ of the best stock over the investment period, and achieves this maximal wealth if $\mathbf{Q}_1$ concentrates on the best stock.

Another simple and important class of investment strategies is the class of *constantly rebalanced portfolios*. Such a strategy B is parametrized by a probability vector $\mathbf{B} = (B_1, \ldots, B_N) \in \mathcal{X}$, and simply $\mathbf{Q}_t(\mathbf{x}_1^{t-1}) = \mathbf{B}$ regardless of t and the past market behavior $\mathbf{x}_1^{t-1}$. Thus, an investor following such a strategy rebalances, at every trading period, his current wealth according to the distribution $\mathbf{B}$ by investing a proportion B_1 of his wealth in the first stock, a proportion B_2 in the second stock, etc. The wealth factor achieved after n trading periods is

$$S_n(\mathbf{B}, \mathbf{x}_1^n) = \prod_{t=1}^{n} \left(\sum_{i=1}^{N} x_{i,t} B_i \right) \ .$$

Now given a class $\mathcal{Q}$ of investment strategies, we define the *worst-case logarithmic wealth ratio* of strategy P by

$$W_n(P, \mathcal{Q}) = \sup_{\mathbf{x}_1^n} \sup_{Q \in \mathcal{Q}} \ln \frac{S_n(Q, \mathbf{x}_1^n)}{S_n(P, \mathbf{x}_1^n)} \ .$$

The worst-case logarithmic wealth ratio is the analog of the external regret in the sequential portfolio selection problem. $W_n(P, \mathcal{Q}) = o(n)$ means that the investment strategy P achieves the same exponent of growth as the best reference strategy in the class $\mathcal{Q}$ for all possible market behaviors.

For example, it is immediate to see that if $\mathcal{Q}$ is the class of all buy-and-hold strategies, then if P is chosen to be the buy-and-hold strategy based on the uniform distribution $\mathbf{Q}_1$, then $W_n(P, \mathcal{Q}) \leq \ln N$.

The class of constantly rebalanced portfolios is a significantly richer class and achieving a small worst-case logarithmic wealth ratio is a greater challenge. Cover's *universal portfolio* (1991) was the first example to achieve this goal. The universal portfolio strategy P is defined by

$$P_{j,t}(\mathbf{x}_1^{t-1}) = \frac{\int_{\mathcal{X}} B_j S_{t-1}(\mathbf{B}, \mathbf{x}_1^{t-1})\phi(\mathbf{B})d\mathbf{B}}{\int_{\mathcal{X}} S_{t-1}(\mathbf{B}, \mathbf{x}_1^{t-1})\phi(\mathbf{B})d\mathbf{B}} , \quad j = 1, \dots, N, t = 1, \dots, n$$

where ϕ is a density function on $\mathcal{X}$. In the simplest case ϕ is the uniform density over $\mathcal{X}$. In that case, the worst-case logarithmic wealth ratio of P with respect to the class $\mathcal{Q}$ of all universal portfolios satisfies

$$W_n(P, \mathcal{Q}) \leq (N-1)\ln(n+1) \ .$$

If the universal portfolio is defined using the Dirichlet$(1/2, \cdots, 1/2)$ density ϕ, then the bound improves to

$$W_n(P, \mathcal{Q}) \leq \frac{N-1}{2}\ln n + \ln \frac{\Gamma(1/2)^N}{\Gamma(N/2)} + \frac{N-1}{2}\ln 2 + o(1) \ ,$$

see Cover and Ordentlich (1996). The worst-case performance of the universal portfolio is basically unimprovable (see Ordentlich and Cover (1998)) but it has some practical disadvantages, including computational difficulties for not very small values of N. Helmbold, Schapire, Singer, and Warmuth (1998) suggest their EG strategy to overcome these difficulties.

The EG strategy is defined by

$$P_{i,t+1} = \frac{P_{i,t}\exp\left(\eta x_{i,t}/\mathbf{P}_t \cdot \mathbf{x}_t\right)}{\sum_{j=1}^N P_{j,t}\exp\left(\eta x_{j,t}/\mathbf{P}_t \cdot \mathbf{x}_t\right)} \ . \tag{5}$$

Helmbold, Schapire, Singer, and Warmuth (1998) prove that if the market values $x_{i,t}$ all fall between the positive constants m and M, then the worst-case logarithmic wealth ratio of the EG investment strategy is bounded by

$$\frac{\ln N}{\eta} + \frac{n\eta}{8}\frac{M^2}{m^2} = \frac{M}{m}\sqrt{\frac{n}{2}\ln N} \ ,$$

where the equality holds for the choice $\eta = (m/M)\sqrt{(8\ln N)/n}$. Here we give a simple new proof of this result, mostly because the main idea is at the basis of other arguments that follow. Recall that the worst-case logarithmic wealth ratio is

$$\max_{\mathbf{x}_1^n} \max_{\mathbf{B}\in\mathcal{X}} \ln \frac{\prod_{t=1}^n \mathbf{B} \cdot \mathbf{x}_t}{\prod_{t=1}^n \mathbf{P}_t \cdot \mathbf{x}_t}$$

where in this case the first maximum is taken over market sequences satisfying the boundedness assumption. By using the elementary inequality $\ln(1+u) \le u$, we obtain

$$\begin{aligned}
\ln \frac{\prod_{t=1}^n \mathbf{B}\cdot\mathbf{x}_t}{\prod_{t=1}^n \mathbf{P}_t\cdot\mathbf{x}_t} &= \sum_{t=1}^n \ln\left(1+\frac{(\mathbf{B}-\mathbf{P}_t)\cdot\mathbf{x}_t}{\mathbf{P}_t\cdot\mathbf{x}_t}\right) \\
&\le \sum_{t=1}^n \sum_{i=1}^N \frac{(B_i - P_{i,t})\, x_{i,t}}{\mathbf{P}_t\cdot\mathbf{x}_t} \\
&= \sum_{t=1}^n \left(\sum_{j=1}^N \sum_{i=1}^N P_{i,t} \frac{B_j x_{j,t}}{\mathbf{P}_t\cdot\mathbf{x}_t} - \sum_{i=1}^N \sum_{j=1}^N B_j \frac{P_{i,t} x_{i,t}}{\mathbf{P}_t\cdot\mathbf{x}_t} \right) \\
&= \sum_{j=1}^N B_j \left(\sum_{t=1}^n \sum_{i=1}^N P_{i,t} \left(\frac{x_{j,t}}{\mathbf{P}_t\cdot\mathbf{x}_t} - \frac{x_{i,t}}{\mathbf{P}_t\cdot\mathbf{x}_t} \right)\right) . \qquad (6)
\end{aligned}$$

Under the boundedness assumption $0 < m \le x_{i,t} \le M$, the quantities

$$\ell_{i,t} = M/m - x_{i,t}/(\mathbf{P}_t\cdot\mathbf{x}_t)$$

are within $[0, M/m]$ and can therefore be interpreted as bounded loss functions. Thus, the minimization of the above upper bound on the worst-case logarithmic wealth ratio may be cast as a sequential prediction problem as described in Section 2. Observing that the EG investment algorithm is just the exponentially weighted average predictor for this prediction problem, and using the performance bound (1) we obtain the cited inequality of Helmbold, Schapire, Singer, and Warmuth (1998).

Note that in (5), we could replace the fixed η by a time-adaptive $\eta_t = (m/M)\sqrt{(8\ln N)/t}$. Applying Theorem 3 to the linear upper bound (6), we may prove that this still leads to a worst-case logarithmic wealth ratio less than something of the order of $(M/m)\sqrt{n \ln N}$.

Remark. (Sub-optimality of the EG *investment strategy.)* Using the approach of bounding the worst-case logarithmic wealth ratio linearly as above is inevitably suboptimal. Indeed, the right-hand side of the linear upper bounding

$$\sum_{j=1}^N B_j \left(\sum_{t=1}^n \left(\sum_{i=1}^N P_{i,t}\ell_{i,t} \right) - \ell_{j,t} \right) = \sum_{j=1}^N B_j \sum_{i=1}^N \left(\sum_{t=1}^n P_{i,t} \left(\ell_{i,t} - \ell_{j,t}\right) \right)$$

is maximized for a constantly rebalanced portfolio $\mathbf{B}$ lying in a corner of the simplex $\mathcal{X}$, whereas the left-hand side is concave in $\mathbf{B}$ and therefore is possibly maximized in the interior of the simplex. Thus, no algorithm trying to minimize (in a worst-case sense) the linear upper bound on the external regret can be minimax optimal. However, as it is shown in Helmbold, Schapire, Singer, and Warmuth (1998), on real data good performance may be achieved.

Note also that the bound obtained for the worst-case logarithmic wealth ratio of the EG strategy grows as $\sqrt{n}$ whereas that of Cover's universal portfolio has only a logarithmic growth. In Helmbold, Schapire, Singer, and Warmuth (1998) it is asked whether the suboptimal bound for the EG strategy is an artifact of the analysis or it is inherent in the algorithm. The next simple example shows that no bound of a smaller order than $\sqrt{n}$ holds. Consider a market with two assets and market vectors $\mathbf{x}_t = (1, 1-\varepsilon)$, for all t. Then every wealth allocation $\mathbf{P}_t$ satisfies $1 - \varepsilon \le \mathbf{P}_t \cdot \mathbf{x}_t \le 1$. Now, the best constantly rebalanced portfolio is clearly $(1, 0)$, and the worst-case logarithmic wealth ratio is simply

$$\sum_{t=1}^{n} \ln \frac{1}{1 - P_{2,t}\varepsilon} \ge \sum_{t=1}^{n} P_{2,t}\varepsilon \ .$$

In the case of the EG strategy,

$$\begin{aligned}
P_{2,t} &= \frac{\exp\left(\eta \sum_{s=1}^{t-1} \frac{(1-\varepsilon)}{\mathbf{P}_s \cdot \mathbf{x}_s}\right)}{\exp\left(\eta \sum_{s=1}^{t-1} \frac{1}{\mathbf{P}_s \cdot \mathbf{x}_s}\right) + \exp\left(\eta \sum_{s=1}^{t-1} \frac{(1-\varepsilon)}{\mathbf{P}_s \cdot \mathbf{x}_s}\right)} \\
&= \frac{\exp\left(-\eta\varepsilon \sum_{s=1}^{t-1} \frac{1}{\mathbf{P}_s \cdot \mathbf{x}_s}\right)}{1 + \exp\left(-\eta\varepsilon \sum_{s=1}^{t-1} \frac{1}{\mathbf{P}_s \cdot \mathbf{x}_s}\right)} \\
&\ge \frac{\exp\left(-\eta\left(\varepsilon/(1-\varepsilon)\right)(t-1)\right)}{2} \ .
\end{aligned}$$

Thus, the logarithmic wealth ratio of the EG algorithm is lower bounded by

$$\begin{aligned}
\sum_{t=1}^{n} \varepsilon \frac{\exp\left(-\eta\left(\varepsilon/(1-\varepsilon)\right)(t-1)\right)}{2} &= \frac{\varepsilon}{2} \frac{1 - \exp\left(-\eta\left(\varepsilon/(1-\varepsilon)\right)n\right)}{1 - \exp\left(-\eta\left(\varepsilon/(1-\varepsilon)\right)\right)} \\
&= \frac{1}{2}\sqrt{\frac{n}{8\ln N}} + o(\sqrt{n}) \ .
\end{aligned}$$

5. Internal regret of investment strategies

The aim of this section is to introduce the notion of internal regret to the sequential investment problem. In the latter, the loss function we consider is defined by $\ell'(\mathbf{Q},\mathbf{x}) = -\ln \mathbf{Q}\cdot\mathbf{x}$ for a portfolio $\mathbf{Q}$ and a market vector $\mathbf{x}$. This is no longer a linear function of $\mathbf{Q}$ (as this was the case in Sections 2 and 3 for the expected loss of the predictor).

Recall that in the framework of sequential prediction described in Section 2, the cumulative internal regret $R_{(i,j),n}$ for the pair of experts (i,j) may be interpreted as how much the predictor would have gained, had he replaced all values $P_{i,t}$ $(t \le n)$ by zero and all values $P_{j,t}$ by $P_{i,t}+P_{j,t}$. Analogously, given an investment strategy $P = (\mathbf{P}_1, \mathbf{P}_2, \ldots)$, we may define the *internal regret of P with respect to the pair of assets (i,j) at day t* (where $1 \le i,j \le N$) by

$$\widetilde{r}_{(i,j),t} = \ell'(\mathbf{P}_t,\mathbf{x}_t) - \ell'(\mathbf{P}_t^{i\to j},\mathbf{x}_t) = \ln \frac{\mathbf{P}_t^{i\to j}\cdot\mathbf{x}_t}{\mathbf{P}_t\cdot\mathbf{x}_t}$$

where the probability vector $\mathbf{P}_t^{i\to j}$ is defined such that its i-th component equals zero, its j-th component equals $P_{j,t}+P_{i,t}$, and all other components are equal to those of $\mathbf{P}_t$. $\widetilde{r}_{(i,j),t}$ expresses the regret the investor using strategy P suffers after trading day t of not having invested all the capital he invested in stock i in stock j instead. The *cumulative internal regret of P with respect to the pair (i,j)* after n trading periods is simply

$$\widetilde{R}_{(i,j),n} = \sum_{t=1}^{n} \widetilde{r}_{(i,j),n} \ .$$

This notion of internal regret in on-line portfolio selection may be seen as a special case of the definition of internal regret for general loss functions proposed in Stoltz and Lugosi (2004), with the class of departure functions given by those functions that move all probability mass from a given component to another one. In Section 7.2, we study internal regret with respect to a much larger class, whose size is of the power of the continuum. It is a desirable property of an investment strategy that its cumulative internal regret grows sub-linearly for all possible pairs of assets, independently of the market outcomes. Indeed, otherwise the investor could exhibit a simple modification of his betting strategy which would have led to exponentially larger wealth. In this sense, the notion of internal regret is a measure of the efficiency of the strategy: the aim of the broker is not that the owner of the stocks gets rich, but that the owner cannot criticize easily the chosen strategy. Note that the worst-case logarithmic wealth ratio corresponds to the case when the owner compares his achieved wealths to those obtained by others who have different brokers. Based on this, we define the *internal regret* of the investment strategy P by

$$\widetilde{R}_n = \max_{1\le i,j\le N} \widetilde{R}_{(i,j),n}$$

and ask whether it is possible to guarantee that $\widetilde{R}_n = o(n)$ for all possible market sequences. Thus, an investor using a strategy with a small internal regret is guaranteed that for any pair of stocks the total regret of not investing in one stock instead of the other becomes negligible. (Note that in Section 7.2 we introduce a richer class of possible departures from the original investment strategies.)

The next two examples show that it is not trivial to achieve a small internal regret. Indeed, the buy-and-hold and EG investment strategies have linearly increasing internal regret for some bounded market sequences.

Example 2. (Buy-and-hold strategies may have large internal regret.) Consider a market with $N = 3$ assets which evolves according to the following repeated scheme:

$$(1-\varepsilon,\varepsilon,\varepsilon),\ (\varepsilon,1-\varepsilon,1-\varepsilon),\ (1-\varepsilon,\varepsilon,\varepsilon),\ (\varepsilon,1-\varepsilon,1-\varepsilon),\ \ldots$$

where $\varepsilon < 1$ is a fixed positive number. The buy-and-hold strategy, which distributes its initial wealth uniformly among the assets invests, at odd t's, with

$$\mathbf{P}_t = \left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right) \ , \quad \text{so that } \mathbf{P}_t^{2\to 1} = \left(\frac{2}{3},0,\frac{1}{3}\right) \ ,$$

and at even t's, with

$$\mathbf{P}_t = \left(\frac{1-\varepsilon}{1+\varepsilon},\frac{\varepsilon}{1+\varepsilon},\frac{\varepsilon}{1+\varepsilon}\right) \ , \quad \text{so that } \mathbf{P}_t^{2\to 1} = \left(\frac{1}{1+\varepsilon},0,\frac{\varepsilon}{1+\varepsilon}\right) \ .$$

Straightforward calculation now shows that for an even n, the cumulative internal regret $\widetilde{R}_{(2,1),n}$ of this strategy equals

$$\frac{n}{2}\left(\ln\frac{(2-\varepsilon)^2}{3(1-\varepsilon)(1+\varepsilon)}\right) \ ,$$

showing that even for bounded markets, the naive buy-and-hold strategy may incur a large internal regret. Later we will see a generalization of buy-and-hold with small internal regret.

Example 3. (The EG *strategy may have large internal regret.)* The next example, showing that for some market sequence the EG algorithm of Helmbold, Schapire, Singer, and Warmuth (1998) has a linearly growing internal regret, is inspired by Example 1 above. Consider a market of three stocks A, B, and C. Divide the n trading periods into three different regimes of lengths n_1, n_2, and n_3. The wealth ratios

Table II. The market vectors for Example 3.

Regimes	$x_{A,t}$	$x_{B,t}$	$x_{C,t}$
$1 \le t \le T_1 = n_1$	2	1	0.5
$T_1 + 1 \le t \le T_2 = n_1 + n_2$	1	2	0.5
$T_2 + 1 \le t \le T_3 = n$	1	2	2.05

(which are constant in each regime) are summarized in the Table II. We show that it is possible to set n_1, n_2, and n_3 in such a way that the cumulative internal regret $R_{(B,C),n}$ is lower bounded by a positive constant times n for n sufficiently large.

The internal regret of B versus C can be lower bounded by using the inequality $\ln(1+u) \le u$:

$$\sum_{t=1}^{n} \ln \frac{\mathbf{Q}_t^{B\to C} \cdot \mathbf{x}_t}{\mathbf{Q}_t \cdot \mathbf{x}_t} \ge \sum_{t=1}^{n} Q_{B,t} \left(\frac{x_{C,t}}{\mathbf{Q}_t^{B\to C} \cdot \mathbf{x}_t} - \frac{x_{B,t}}{\mathbf{Q}_t^{B\to C} \cdot \mathbf{x}_t} \right),$$

where the difference in the parenthesis is larger than -1 in the first regime, -3 in the second one and $0.05/2.05$ in the third one. It suffices now to estimate $Q_{B,t}$:

$$Q_{B,t} = \frac{e^{\eta G_{B,t}}}{e^{\eta G_{A,t}} + e^{\eta G_{B,t}} + e^{\eta G_{C,t}}}, \tag{7}$$

where

$$\eta = 4.1 \sqrt{\frac{8 \ln 3}{n}} \hat{=} \frac{1}{C_\eta \sqrt{n}} \quad \text{and} \quad G_{B,t} = \sum_{s=1}^{t} \frac{x_{B,s}}{\mathbf{Q}_s \cdot \mathbf{x}_s}$$

(and similarly for the two other stocks).

We take $n_1 = dn$, where $d > 0$ will be determined later. In the first regime, a sufficient condition for $Q_{B,t} \le \varepsilon$ is that $e^{\eta G_{B,t}}/e^{\eta G_{A,t}} \le \varepsilon$, which can be ensured by

$$G_{A,t} - G_{B,t} = \sum_{s=1}^{t} \frac{1}{\mathbf{Q}_s \cdot \mathbf{x}_s} \ge \frac{-\ln \varepsilon}{\eta},$$

which is implied, since $\mathbf{Q}_s \cdot \mathbf{x}_s \le 2$, by

$$t \ge t_0 = 2 C_\eta \left(-\ln \varepsilon\right) \sqrt{n}.$$

In the second regime, the $Q_{B,t}$'s increase. Let T_2 denote the first time instant t when $Q_{B,t} \ge 1/2$, and denote by $n_2 = T_2 - T_1$ the length of this second regime. Now, it is easy to see that $n_2 \geq n_1/4$ and $n_2 \leq 4n_1 + (2\ln 2)C_\eta\sqrt{n} \leq 5dn$, for n sufficiently large. Moreover, the number of times that $Q_{B,t}$ is larger than ε in this regime is less than

$$C_\eta \left(\ln \left(2\frac{1-\varepsilon}{\varepsilon} \right) \right) \sqrt{n}.$$

At the beginning of the third regime, we then have $Q_{B,t} \geq 1/2$, which means that $G_{A,t} \leq G_{B,t}$ and $G_{C,t} \leq G_{B,t}$. The first inequality remains true during the whole regime and we set n_3 such that the second one also remains true. This will imply that $Q_{B,t} \geq 1/3$ during the third regime. Now by the bounds on $\mathbf{Q}_s \cdot \mathbf{x}_s$ in the different regimes, a sufficient condition on n_3 is

$$0.05 n_3 \leq \frac{n_1}{4} + \frac{3n_2}{4},$$

which, recalling the lower bound $n_2 \geq n_1/4$, is implied by

$$n_3 \leq \frac{35}{4} dn.$$

It remains to set the value of d. We have to ensure that n_3 is not larger than $35dn/4$ and that it is larger than γn, where γ is a universal constant denoting the fraction of time spent in the third regime. That is, we have to find d and γ such that

$$\begin{cases} d + 5d + \gamma & \leq & 1 \\ d + \frac{1}{4}d + \frac{35}{4}d & \geq & 1 \end{cases},$$

where we used $n_1/n + n_2/n + n_3/n = 1$ and the various bounds and constraints described above. $\gamma = 1/7$ and $d = 1/7$ are adequate choices.

Summarizing, we have proved the following lower bound on the internal regret

$$\sum_{t=1}^{n} \ln \frac{\mathbf{Q}_t^{B \to C} \cdot \mathbf{x}_t}{\mathbf{Q}_t \cdot \mathbf{x}_t} \geq \frac{1}{3}\gamma \frac{0.05}{2.05} n - \varepsilon \left(3(1-\gamma) \right) n + \Omega \left((-\ln \varepsilon)\sqrt{n} \right),$$

and the proof that the EG strategy has a large internal regret is concluded by choosing $\varepsilon > 0$ small enough (for instance, $\varepsilon = 1/5000$).

6. Investment strategies with small internal regret

The investment algorithm introduced in the next section has the surprising property that, apart from a guaranteed sublinear internal regret, it also achieves a sublinear worst-case logarithmic wealth ratio not only with respect to the class of buy-and-hold strategies, but also with respect to the class of all constantly rebalanced portfolios.

6.1. A strategy with small internal and external regrets

The investment strategy introduced in this section–which we call B1EXP–is based on the same kind of linear upper bound on the internal regret as the one that was used in our proof of the performance of the EG strategy in Section 4. This strategy may be seen as the algorithm that results from an application of the conversion trick explained in Section 3 to the EG strategy. However, this only proves the no-internal-regret property. Since the worst-case logarithmic wealth ratio is also minimized, we provide a detailed analysis below.

The same argument as for the EG strategy may be used to upper bound the cumulative internal regret as

$$\begin{aligned}\widetilde{R}_{(i,j),n} &= \sum_{t=1}^{n} \ln\left(\mathbf{P}_t^{i\to j}\cdot\mathbf{x}_t\right) - \ln\left(\mathbf{P}_t\cdot\mathbf{x}_t\right)\\ &\le \sum_{t=1}^{n} P_{i,t}\left(\frac{x_{j,t}}{\mathbf{P}_t\cdot\mathbf{x}_t} - \frac{x_{i,t}}{\mathbf{P}_t\cdot\mathbf{x}_t}\right)\ .\end{aligned}$$

Introducing again

$$\ell_{i,t} = -\frac{x_{i,t}}{\mathbf{P}_t\cdot\mathbf{x}_t}\ ,$$

we may use the internal-regret minimizing prediction algorithm of Section 3. For simplicity, we use exponential weighting. This definition, of course, requires the boundedness of the values of $\ell_{i,t}$. This may be guaranteed by the same assumption as in the analysis of the EG investment strategy, that is, by assuming that the returns $x_{i,t}$ all fall in the interval $[m, M]$ where $m < M$ are positive constants. Then the internal regret of the algorithm B1EXP may be bounded by the result of Theorem 3. An important additional property of the algorithm is that its worst-case logarithmic wealth ratio, with respect to the class of all constantly rebalanced portfolios, may be bounded similarly as that of the EG algorithm. These main properties are summarized in the following theorem.

Theorem 4. Assume that $m \le x_{i,t} \le M$ for all $1 \le i \le N$ and $1 \le t \le n$. Then the cumulative internal regret of the B1EXP strategy

P is bounded by

$$\widetilde{R}_n \le \frac{\ln N(N-1)}{\eta} + \frac{n\eta}{8}\frac{M^2}{m^2} = \frac{M}{m}\sqrt{n \ln N}\ ,$$

where we set $\eta = 4(m/M)\sqrt{(\ln N)/n}$. In addition, if $\mathcal{Q}$ denotes the class of all constantly rebalanced portfolios, then the worst-case logarithmic wealth ratio of P is bounded by

$$W_n(P, \mathcal{Q}) \le N\frac{M}{m}\sqrt{n \ln N}\ .$$

Proof. The bound for the internal regret $\widetilde{R}_n$ follows from the linear upper bound described above and Theorem 3.

To bound the worst-case logarithmic wealth ratio $W_n(P, \mathcal{Q})$, recall that by inequality (6), for any constantly rebalanced portfolio $\mathbf{B}$,

$$\begin{aligned} W_n(P, \mathcal{Q}) \ &\le\ \sum_{j=1}^{N} B_j \sum_{i=1}^{N} \left(\sum_{t=1}^{n} P_{i,t} \left(\ell_{i,t} - \ell_{j,t}\right) \right) \\ &\le\ N \max_{1 \le i,j \le N} \sum_{t=1}^{n} P_{i,t} \left(\frac{x_{j,t}}{\mathbf{P}_t \cdot \mathbf{x}_t} - \frac{x_{i,t}}{\mathbf{P}_t \cdot \mathbf{x}_t} \right) \end{aligned}$$

which is not larger than N times the upper bound obtained on the cumulative internal regret $\widetilde{R}_n$ which completes the proof.

Remark. The computation of the investment strategy requires the inversion of an $N \times N$ matrix at each trading period (see Lemma 1). This is quite feasible even for large markets in which N may be as large as about 100.

Remark. Recalling Section 3 we observe that the B1EXP strategy may be considered as an instance of the exponentially weighted average predictor, which uses the fictitious strategies $\mathbf{P}_t^{i \to j}$ as experts. Thus, instead of considering the single stocks, as EG, B1EXP considers pairs of stocks and their relative behaviors. This may explain the greater stability observed on real data (see the Appendix).

Remark. Just like in the case of the sequential prediction problem, exponential weighting may be replaced by others such as polynomial weighting. In that case the cumulative internal regret is bounded by $\frac{M}{m}\sqrt{n(p-1)}N^{2/p}$ which is approximately optimized by the choice $p =$

$4 \ln N$. We call this investment strategy B1POL. Even though this strategy has comparable theoretical guarantees to those of B1EXP, our experiments show a clear superiority of the use of exponential weighting. This and other practical issues are discussed in the Appendix.

Remark. Similarly to EG, the strategy B1EXP requires the knowledge of the time horizon n and the ratio M/m of the bounds assumed on the market. This first disadvantage may be avoided by either using the well-known "doubling trick" or considering a time-varying value of η and applying the second bound of Theorem 3. Both methods lead to internal regret and worst-case logarithmic wealth ratios bounded by quantities of the order of $\sqrt{n}$.

6.2. Another strategy with small internal regret

In this section we introduce a new algorithm, called B2POL. We use polynomial weighting and assume bounded market evolutions. The Blackwell condition (4) is sufficient to ensure the property of small internal regret. It may be written as

$$\sum_{i \neq j} \Delta_{(i,j),t} \widetilde{r}_{(i,j),t} \leq 0 \tag{8}$$

where

$$\Delta_{(i,j),t} = \frac{\left(\widetilde{R}_{(i,j),t-1}\right)_+^{p-1}}{\sum_{a \neq b} \left(\widetilde{R}_{(a,b),t-1}\right)_+^{p-1}}.$$

Note that the $\Delta_{(i,j),t}$'s are nonnegative and sum up to one. The concavity of the logarithm and the definition of $\widetilde{r}_{(i,j),t}$ lead to

$$\begin{aligned}
\sum_{i \neq j} \Delta_{(i,j),t} \widetilde{r}_{(i,j),t} &= \left(\sum_{i \neq j} \Delta_{(i,j),t} \ln \left(\mathbf{P}_t^{i \to j} \cdot \mathbf{x}_t \right) \right) - \ln \left(\mathbf{P}_t \cdot \mathbf{x}_t \right) \\
&\leq \ln \left(\sum_{i \neq j} \Delta_{(i,j),t} \mathbf{P}_t^{i \to j} \cdot \mathbf{x}_t \right) - \ln \left(\mathbf{P}_t \cdot \mathbf{x}_t \right).
\end{aligned}$$

It is now obvious that the Blackwell condition (4) is satisfied whenever

$$\mathbf{P}_t = \sum_{i \neq j} \Delta_{(i,j),t} \mathbf{P}_t^{i \to j}.$$

Lemma 1 shows that such a portfolio $\mathbf{P}_t$ indeed exists for all t. This defines a strategy which we call B2POL. The following theorem is an immediate consequence of Corollary 1 of Cesa-Bianchi and Lugosi (2003).

Theorem 5. Assume that $m \le x_{i,t} \le M$ for all $1 \le i \le N$ and $1 \le t \le n$. Then the cumulative internal regret of the B2POL strategy P is bounded by

$$\widetilde{R}_n \le \left(\ln \frac{M}{m}\right) \sqrt{n(p-1)} N^{2/p} .$$

The above bound is approximately minimized for $p = 4 \ln N$. Note also that it only differs from the bound on the cumulative internal regret of the B1POL strategy by a constant factor which is smaller here ($\ln(M/m)$ instead of M/m).

7. Generalizations

7.1. Generalized buy-and-hold strategy

The GBH strategy performs buy-and-hold on the $N(N-1)$ fictitious modified strategies, using the conversion trick explained in Section 3 (and, in the particular case of $N = 2$ assets, it reduces to the simple buy-and-hold strategy–hence its name). The main property of this investment strategy is that its internal regret is bounded by a constant, as stated by the theorem below.

More precisely, the GBH strategy is defined such that at each round t, we have the fixed point equality

$$\mathbf{P}_t = \sum_{i \ne j} \frac{W_{t-1}^{i \to j}}{\sum_{k \ne l} W_{t-1}^{k \to l}} \mathbf{P}_t^{i \to j} , \tag{9}$$

where $W_t = \prod_{s=1}^{t} \mathbf{P}_s \cdot \mathbf{x}_s$ is the wealth achieved by the investment strategy we consider and $W_t^{i \to j} = \prod_{s=1}^{t} \mathbf{P}_s^{i \to j} \cdot \mathbf{x}_s$ is the fictitious wealth obtained by the $i \to j$ modified version of it. The existence and the practical computation of such a portfolio $\mathbf{P}_t$ are given by Lemma 1.

Theorem 6. The GBH investment strategy incurs a cumulative internal regret $\widetilde{R}_n \le \ln N(N-1)$ for all n.

Proof. The proof is done by a simple telescoping argument:

$$W_n = \prod_{t=1}^{n} \mathbf{P}_t \cdot \mathbf{x}_t \; = \; \prod_{t=1}^{n} \sum_{i \ne j} \frac{W_{t-1}^{i \to j} \mathbf{P}_t^{i \to j} \cdot \mathbf{x}_t}{\sum_{k \ne l} W_{t-1}^{k \to l}}$$

$$= \frac{\sum_{i\neq j} W_n^{i\to j}}{N(N-1)} \; .$$

The advantage of this algorithm is that its performance bounds do not depend on the market.

Remark. Unlike in the sequential prediction problem described in Section 2, a small internal regret in the problem of sequential portfolio selection does not necessarily imply a small worst-case logarithmic wealth ratio, not even with respect to the class of all buy-and-hold strategies. This may be seen by considering the following numerical counterexample. Let the market be formed by three stocks and let it be cyclic such that at odd-indexed rounds the wealth ratios are respectively 1/2, 1, 2 and at even ones they equal 2, 1.1, 1/2. The accumulated wealth of the best stock increases exponentially fast whereas the one of the GBH strategy is bounded.

The reason is that the loss function ℓ' associated to this problem is no longer linear, and therefore, the argument of Equation (2) does not extend to it.

However, there is a simple modification of the GBH strategy leading to internal regret less than $2\ln N$ and external regret with respect to buy-and-hold strategies less than $2\ln N$. We call this modification the GBH2 algorithm.

Instead of (9), the GBH2 strategy is such that

$$\mathbf{P}_t = \frac{\sum_{1\le k\le N} S_{t-1}(k)\mathbf{e}_k + \sum_{i\neq j} W_{t-1}^{i\to j}\mathbf{P}_t^{i\to j}}{\sum_{1\le k\le N} S_{t-1}(k) + \sum_{i\neq j} W_{t-1}^{i\to j}} \; , \tag{10}$$

for every t, where $\mathbf{e}_k$ denotes the portfolio that invests all its wealth in the k-th stock. Now a telescoping argument similar to that of the proof of Theorem 6 shows that the final wealth equals

$$W_n = \frac{1}{N^2}\left(\sum_{1\le k\le N} S_n(k) + \sum_{i\neq j} W_n^{i\to j} \right) \; ,$$

thus ensuring that both regrets are less than the claimed upper bound $2\ln N$. Lemma 1 shows that (10) can be satisfied and how the portfolios $\mathbf{P}_t$ are computed.

The next section is an extension of GBH and GBH2 strategies to a continuum of fictitious experts.

7.2. A GENERALIZED UNIVERSAL PORTFOLIO

Next we extend the notion of internal regret for investment strategies. Recall that the definition of internal regret $\widetilde{R}_n$ considers the regret suffered by not moving one's capital from one stock to another. Moving the capital from one stock to another may be considered as a simple linear function from the probability simplex $\mathcal{X}$ to $\mathcal{X}$. A more exigent definition is obtained by considering all linear functions $g : \mathcal{X} \to \mathcal{X}$. Clearly, any such function may be written as $g(\mathbf{P}_t) = \mathbf{A}\mathbf{P}_t$ where $\mathbf{A}$ is a column-stochastic matrix. Denote the set of all column-stochastic matrices of order N by $\mathcal{A}$ and let the linear modifications $\mathbf{A}\mathbf{P}_t$ of the master strategy be denoted by $\mathbf{P}_t^{\mathbf{A}}$. The generalized internal regret is defined as

$$\max_{\mathbf{A}\in\mathcal{A}} \ln \frac{W_n^{\mathbf{A}}}{W_n}$$

where $W_n^{\mathbf{A}} = \prod_{t=1}^n \sum_{i=1}^N P_{i,t}^{\mathbf{A}} x_{i,t}$.

Linear modifications were already considered (in finite number) by Greenwald and Jafari (2003) in the case of sequential prediction. In that case, due to the linearity of the loss function $\ell(\mathbf{P}_t)$, it is not more difficult to have a low generalized internal regret than the usual internal regret. On the contrary here, due to the concavity of the logarithm, minimizing the generalized internal regret turns out to be a greater challenge. Since the algorithms B1EXP and B1POL are based on a linear upper bounding of the internal regret, it is easy to see that their generalized internal regret is bounded by N times the bounds derived for the internal regret in Sections 6.1, leading to upper bounds both of the order of $N\sqrt{n \ln N}$.

The main result of this section is that there exist investment strategies that achieve a much smaller generalized internal regret. The proof below is inspired by Theorem 6 and uses some techniques introduced by Blum and Kalai (1999). The investment strategy presented above may be seen as a modification of Cover's universal portfolio (1991) through a conversion trick to deal with generalized internal regret of the same flavor as the one explained in Section 3.

Theorem 7. There exists an investment strategy P such that

$$\max_{\mathbf{A}\in\mathcal{A}} \ln \frac{W_n^{\mathbf{A}}}{W_n} \le N(N-1)\ln(n+1) + 1 .$$

Remark. The algorithm given in the proof has a computational complexity exponential in the number of stocks (at least in its straightforward implementation). However, it provides a theoretical bound which is likely to be of the best achievable order.

The algorithm could also be easily modified, using the techniques of Section 7.1, to be competitive with respect to the best constantly rebalanced portfolio as well as to suffer a low generalized internal regret, with associated performance bounds for both of the order $N^2 \ln n$.

Proof. Denote a column-stochastic matrix $\mathbf{A}$ by $[\mathbf{a}_1, \ldots, \mathbf{a}_N]$, where the $\mathbf{a}_j$'s are the columns of $\mathbf{A}$. Let μ be the uniform measure over the simplex and let ν be the measure over $\mathcal{A}$ given by the product of N independent instances of μ:

$$\nu(\mathbf{A}) = \prod_{j=1}^{N} \mu(\mathbf{a}_j) \ .$$

If the investment strategy, at each time instant t, satisfied the equality

$$\mathbf{P}_t = \frac{\int_{\mathbf{A}\in\mathcal{A}} W_{t-1}^{\mathbf{A}} \mathbf{P}_t^{\mathbf{A}} d\nu(\mathbf{A})}{\int_{\mathbf{A}\in\mathcal{A}} W_{t-1}^{\mathbf{A}} d\nu(\mathbf{A})}, \tag{11}$$

then the final wealth would be given by an average over all modified strategies, that is,

$$W_n = \int_{\mathbf{A}\in\mathcal{A}} W_n^{\mathbf{A}} d\nu(\mathbf{A}). \tag{12}$$

Fix a matrix $\mathbf{A}$ and consider the set $\chi_{\alpha,\mathbf{A}}$ of column-stochastic matrices of the form $(1-\alpha)\mathbf{A} + \alpha\mathbf{z}$, $\mathbf{z} \in \mathcal{A}$. Similarly, denote by $\chi_{\alpha,\mathbf{a}_j}$ the set of probability vectors of the form $(1-\alpha)\mathbf{a}_j + \alpha\mathbf{z}_j$, $\mathbf{z}_j \in \mathcal{X}$. It is easy to see that (with a slight abuse of notation)

$$\chi_{\alpha,\mathbf{A}} = \prod_{j=1}^{N} \chi_{\alpha,\mathbf{a}_j} \ . \tag{13}$$

Any element $\mathbf{A}'$ of $\chi_{\alpha,\mathbf{A}}$ may be seen to satisfy (component-wise)

$$P_t^{\mathbf{A}'} \geq (1-\alpha) P_t^{\mathbf{A}} \ ,$$

for all t and therefore

$$W_n^{\mathbf{A}'} \geq (1-\alpha)^n W_n^{\mathbf{A}} \ .$$

Finally, using equality (13), we have

$$\nu\left(\chi_{\alpha,\mathbf{A}}\right) = \prod_{j=1}^{N} \mu\left(\chi_{\alpha,\mathbf{a}_j}\right) = \left(\alpha^{N-1}\right)^N ,$$

implying

$$\int_{\mathbf{A}'\in\chi_{\alpha,\mathbf{A}}} W_n^{\mathbf{A}'} d\nu(\mathbf{A}') \geq (1-\alpha)^n \, \alpha^{N(N-1)} W_n^{\mathbf{A}} \ .$$

Taking $\alpha = 1/(n+1)$, recalling that

$$(1-\alpha)^n \, \alpha^{N(N-1)} \geq \frac{e^{-1}}{(n+1)^{N(N-1)}} \ ,$$

and combining this with 12, we obtain the theorem.

Thus, it suffices to see that one may satisfy the set of linear equations (11). We denote an element $\mathbf{A} \in \mathcal{A}$ by $\mathbf{A} = [A_{(i,j)}]$. Writing only the equality for the ith components of both sides,

$$\begin{aligned} &\left(\int_{\mathbf{A}\in\mathcal{A}} W_{t-1}^{\mathbf{A}} d\nu(\mathbf{A}) \right) P_{i,t} \\ &= \int_{\mathbf{A}\in\mathcal{A}} W_{t-1}^{\mathbf{A}} \left(\sum_{k=1}^{N} A_{(i,k)} P_{k,t} \right) d\nu(\mathbf{A}) \ , \end{aligned}$$

we see that $\mathbf{P}_t$ has to be an element of the kernel of the matrix T defined by

- if $i \neq k$, $T_{i,k} = w_{i,k}$,
- $T_{i,i} = -\sum_{j\neq i,\ 1\leq j\leq N} w_{j,i}$,

where

$$w_{i,k} = \int_{\mathbf{A}\in\mathcal{A}} W_{t-1}^{\mathbf{A}} A_{(i,k)} d\nu(\mathbf{A}).$$

The same argument as in the proof of Lemma 1 shows that such a vector exists (and the computability of the latter depends on how easy it is to compute the elements of the matrix T).

Acknowledgements

We thank Yoram Singer for sending us the NYSE data set used in the experiments. We also thank Dean Foster for his suggestions that lead us to the example showing that the exponential weighted average predictor has a large internal regret. We are grateful to the three anonymous reviewers for helpful comments.

References

P. Auer, N. Cesa-Bianchi, and C. Gentile. Adaptive and self-confident on-line learning algorithms. *Journal of Computer and System Sciences*, 64:48–75, 2002.

D. Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6:1–8, 1956.

A. Blum and A. Kalai. Universal portfolios with and without transaction costs. *Machine Learning*, 35:193–205, 1999.

A. Blum and Y. Mansour. From external to internal regret. Manuscript, 2004.

A. Borodin, R. El-Yaniv, and V. Gogan. On the competitive theory and practice of portfolio selection (extended abstract). In *Proc. of the 4th Latin American Symposium on Theoretical Informatics (LATIN'00)*, pages 173–196, Punta del Este, Uruguay, 2000.

N. Cesa-Bianchi and G. Lugosi. On prediction of individual sequences. *Annals of Statistics*, 27:1865–1895,1999.

N. Cesa-Bianchi and G. Lugosi. Minimax values and entropy bounds for portfolio selection problems. In *Proceedings of the First World Congress of the Game Theory Society*, 2000.

N. Cesa-Bianchi and G. Lugosi. Potential-based algorithms in on-line prediction and game theory. *Machine Learning*, 51, 2003.

T. Cover. An algorithm for maximizing expected log investment return. *IEEE Transactions on Information Theory*, 30:369–373, 1984.

T.M. Cover. Universal portfolios. *Mathematical Finance*, 1:1–29, 1991.

T.M. Cover and E. Ordentlich. Universal portfolios with side information. *IEEE Transactions on Information Theory*, 42:348–363, 1996.

D. Foster and R. Vohra. Asymptotic calibration. *Biometrica*, 85:379–390, 1998.

D. Foster and R. Vohra. Regret in the on-line decision problem. *Games and Economic Behavior*, 29:7–36, 1999.

D. Fudenberg and D. Levine. Universal conditional consistency. *Games and Economic Behavior*, 29:104–130, 1999.

A. Greenwald and A. Jafari. A general class of no-regret learning algorithms and game-theoretic equilibria. *Proceedings of the 16th Annual Conference on Learning Theory and 7th Kernel Workshop*, 2-12, 2003.

J. Hannan. Approximation to Bayes risk in repeated play. *Contributions to the theory of games*, 3:97–139, 1957.

S. Hart and A. Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68:1127–1150, 2000.

S. Hart and A. Mas-Colell. A general class of adaptive strategies. *Journal of Economic Theory*, 98:26–54, 2001.

D. P. Helmbold, R. E. Schapire, Y. Singer, and M. K. Warmuth. On-line portfolio selection using multiplicative updates. *Mathematical Finance*, 8:325–344, 1998.

E. Ordentlich and T.M. Cover. The cost of achieving the best portfolio in hindsight. *Mathematics of Operations Research*, 23:960–982, 1998.

Y. Singer. Switching portfolios. *International Journal of Neural Systems*, 8:445–455, 1997.

G. Stoltz and G. Lugosi. Learning correlated equilibria in games with compact sets of strategies. Manuscript, 2004.

Appendix

In this appendix we present an experimental comparison of the performance of the new algorithms with existing ones. In the experiments we used a data set of daily wealth ratios of 36 stocks of the New York Stock Exchange that has been used by various authors including Cover (1991), Cover and Ordentlich (1996), Helmbold, Schapire, Singer, and Warmuth (1998), Blum and Kalai (1999), Singer (1997), and Borodin, El-Yaniv, and Gogan (2000), We also considered monthly wealth ratios (taking 20 days for a month).

We first give an overview of the methodology we used to derive our investment algorithms. A strategy is given by the choice of a measure of the regret $\mathbf{r}_t$ and of a potential function Φ (see Sections 2 and 3). We consider three ways of measuring the regrets:

1. Linear approximation to the instantaneous external regret (see Section 4):
$$r_{i,t} = -\frac{x_{i,t}}{\mathbf{P}_t \cdot \mathbf{x}_t} \ ,$$
2. Instantaneous internal regret (see Sections 6.2 and 7.1):
$$\widetilde{r}_{(i,j),t} = \left(\mathbf{P}_t^{i \to j} \cdot \mathbf{x}_t\right) - \ln\left(\mathbf{P}_t \cdot \mathbf{x}_t\right) \ ,$$
3. Linear approximation to the instantaneous external regret (see Section 6.1):
$$r_{(i,j),t} = P_{i,t}\left(\frac{x_{j,t}}{\mathbf{P}_t \cdot \mathbf{x}_t} - \frac{x_{i,t}}{\mathbf{P}_t \cdot \mathbf{x}_t}\right) \ .$$

Also, both the exponential and the polynomial potentials are used. Each combination of $\mathbf{r}_t$ and Φ induces an investment strategy as summarized in Table III.

Table III. A summary of investment strategies.

Φ	$r_{i,t}$	$\widetilde{r}_{(i,j),t}$	$r_{(i,j),t}$
Exp	EG	GBH	B1EXP
Pol	–	B2POL	B1POL

The tuning of the EG and B1EXP strategies

The first experiment compares the behavior of the B1EXP and EG strategies whose results are summarized in Tables IV and V and Figure 2. We compared the strategies EG and B1EXP for various choices of the tuning parameter η. We used the parameters suggested by theory $\eta^* = \alpha\sqrt{8 \ln N/n}$ and $\eta^* = 4\alpha\sqrt{\ln N/n}$, respectively, in case of known time horizon n, and also the time varying versions $\eta_t^* = \alpha\sqrt{8 \ln N/t}$ and $\eta_t^* = 4\alpha\sqrt{\ln N/t}$ where the ratio $\alpha = m/M$ is taken to be 0.5 for daily rebalancing and 0.3 for monthly rebalancing. (These values are estimated on the data.)

Tables IV and V show the arithmetic averages of the wealths achieved on random samples of size 100. For example, the numbers in the columns "ten stocks" have been obtained by choosing ten of the 36 stocks randomly to form a market of $N = 10$ assets. This experiment was repeated 100 times and the averages of the achieved wealth factors appear in the table. The column "Freq." contains the number of times B1EXP outperformed EG of these 100 experiments. The average wealth ratios for both strategies were calculated for different fixed and time varying parameters. One of the interesting conclusions is that time varying updating never affects the performance of B1EXP while that of EG drops in case of monthly rebalancing or when the number of stocks is large.

In the rest of this experimental study both algorithms are used with their respective time varying theoretical optimal parameter η_t^*.

It is also seen in Tables IV and V that EG is less robust against a bad choice of η. Its performance degrades faster when η or η_t is increased.

Interestingly, the increase of the external regret when the tuning parameter is increased corresponds to an increase in the internal regret, as shown in Figure 2. (This behavior was checked to be typical indeed.) The increase of the internal regret is far larger for the EG strategy. This suggests that minimizing internal regret results in more stability.

Tuning of B1POL and B2POL

Table VI shows that for B1POL and B2POL the theoretically (almost) optimal parameter $p = 4 \ln N$ performs quite poorly in our experiments, for it leads to too fast wealth reallocations. The values of p with better numerical performance are usually far smaller than the ones prescribed by theory. Thus, for the rest of this experimental study and the subsequent simulations, we choose $p = 2$, as it was originally suggested by Blackwell (1956). (Note that in Table VI we show the geometric averages instead of the arithmetic ones, to take into account the huge

Table IV. Evolution of the achieved wealths according to the tuning parameter of EG and B1EXP both for fixed and time varying parameters. Computations are realized on random samples of size 100, arithmetic means are displayed. Monthly rebalancing.

η	Monthly rebalancing					
(para-	Three stocks			Ten stocks		
meter)	EG	B1EXP	Freq.	EG	B1EXP	Freq.
2	14.7	15.5	73	12.8	19.2	95
1.5	15.1	16.0	76	14.0	19.9	96
1	15.9	16.7	80	16.0	20.6	97
0.5	17.3	18.0	84	18.8	21.3	97
0.2	18.7	19.0	84	20.7	21.6	97
0.15	18.9	19.2	84	21.0	21.7	95
0.1	19.2	19.4	84	21.3	21.8	94
0.05	19.5	19.6	82	21.6	21.8	94
0.03	19.6	19.7	82	21.7	21.8	94
0.02	19.7	19.7	82	21.8	21.9	94
0.01	19.7	19.7	82	21.8	21.9	94
η^*	19.5	19.5	80	21.4	21.8	94
η_t^*	19.3	19.4	80	21.2	21.7	95
$0.1\,\eta_t^*$	19.7	19.7	81	21.8	21.9	95
$0.2\,\eta_t^*$	19.7	19.7	80	21.7	21.8	95
$0.5\,\eta_t^*$	19.6	19.6	79	21.5	21.8	95
$2\,\eta_t^*$	18.9	19.0	81	20.5	21.5	95
$5\,\eta_t^*$	17.8	17.9	77	18.7	20.8	97
$10\,\eta_t^*$	16.5	16.7	71	16.1	19.8	94
$25\,\eta_t^*$	14.7	15.4	61	12.5	17.8	92

dispersion of the wealths achieved by these two investment strategies – see also Table X and the related comments.)

Global comparison

In the next experiment various different investment strategies are compared, which we denominate by EG, B1EXP, B1POL, GBH, GBH2, B2POL, Cover's, UBH, B-CRP, and U-CRP. For the first six strategies we have already described how to tune (some of them do not require any tuning). The algorithm "Cover's" stands for Cover's universal portfolio

Table V. Evolution of the achieved wealths according to the tuning parameter of EG and B1EXP both for fixed and time varying parameters. Computations are realized on random samples of size 100, arithmetic means are displayed. Daily rebalancing.

η	Daily rebalancing					
(parameter)	Three stocks			Ten stocks		
	EG	B1EXP	Freq.	EG	B1EXP	Freq.
2	13.2	14.5	77	12.4	21.7	93
1.5	14.1	15.6	80	14.0	23.2	95
1	15.7	17.4	86	17.0	24.7	95
0.5	18.8	20.4	89	22.0	25.8	94
0.2	22.1	23.1	89	25.2	26.3	92
0.15	22.8	23.6	89	25.6	26.3	91
0.1	23.6	24.2	89	26.0	26.4	88
0.05	24.5	24.8	88	26.3	26.4	83
0.03	24.8	25.0	88	26.4	26.5	82
0.02	25.0	25.1	88	26.4	26.5	82
0.01	25.2	25.3	88	26.4	26.5	82
η^*	25.0	25.0	89	26.4	26.5	82
η_t^*	24.8	24.8	86	26.2	26.4	94
$0.1\,\eta_t^*$	25.3	25.3	88	26.5	26.5	91
$0.2\,\eta_t^*$	25.3	25.3	88	26.4	26.5	91
$0.5\,\eta_t^*$	25.1	25.1	87	26.3	26.4	92
$2\,\eta_t^*$	24.2	24.3	86	25.8	26.3	94
$5\,\eta_t^*$	22.6	22.7	85	24.5	26.0	98
$10\,\eta_t^*$	20.4	20.5	82	22.0	25.2	98
$25\,\eta_t^*$	16.2	16.4	72	15.2	22.3	99

based on the uniform density. To compute the universal portfolio, we drew at random 10^3 different constantly rebalanced portfolios and took the average on the wealth ratio sequences to compute each instance of Cover's algorithm. (The value 10^3 may seem to be too small in view of the 10^8 used in Helmbold, Schapire, Singer, and Warmuth (1998) but calculations using the Chebyshev bound of Blum and Kalai (1999) indicate that this value is sufficient to have a good idea of the order of the wealth achieved by the universal portfolio.) To compute the best constantly rebalanced portfolio (called B-CRP) we used a technique

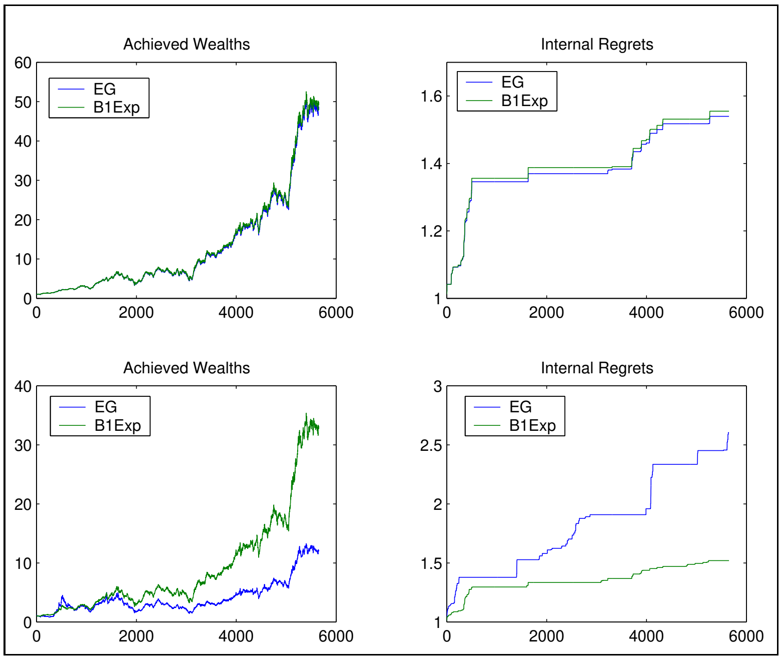

Figure 2. Evolution of both external and internal regrets for the optimal time varying tuning parameter (top) and a 25 times too large one (bottom). Stocks used: Dow Chemical, Coke, GTE, Mei Corp., Gulf, Iroquois, Kin Arc, Amer Brands, Fischbach, Lukens.

described in Cover (1984), with (according to the notations therein) $\varepsilon = 10^{-4}$ for daily rebalancing and $\varepsilon = 10^{-5}$ for monthly rebalancing. This guarantees an estimate within a multiplicative factor of 1.0028 of the wealth achieved by the best contantly rebalanced portfolio in case of a monthly rebalancing and 1.7596 in case of a daily rebalancing. Nevertheless, the values thus obtained are often even closer to the optimal, despite the weak guarantees in case of daily rebalancing. We also considered the uniform buy-and-hold strategy (denoted by UBH) and, following Borodin, El-Yaniv, and Gogan (2000), the uniform constantly rebalanced portfolio (U-CRP).

Transaction costs were also taken into account (whose amount is indicated in the column TC of the tables) according to the model defined in Blum and Kalai (1999). In particular, transaction fees are paid at purchase only. We implemented Blum and Kalai's optimal rebalancing algorithm, using different transaction costs. Here, we summarize the results for zero transaction cost and a heavy 2% at-purchase transaction cost in case of monthly rebalancing and a milder 1% transaction cost when the rebalancing occurs daily.

Table VI. Evolution of the achieved wealths according to the tuning parameter of B1POL and B2POL. Computations are realized on random samples of size 100, geometric means are displayed.

p	Monthly rebalancing				Daily rebalancing			
(para-	Three stocks		Ten stocks		Three stocks		Ten stocks	
meter)	B1POL	B2POL	B1POL	B2POL	B1POL	B2POL	B1POL	B2POL
p^*	11.5	9.5	15.7	12.4	9.1	7.3	11.1	9.7
1.1	13.3	10.9	16.2	13.5	12.7	9.5	16.5	13.5
1.2	13.1	10.9	16.0	13.9	12.3	9.5	16.4	13.5
1.3	13.0	10.9	16.0	13.8	12.1	9.3	16.4	13.8
1.5	12.9	11.0	16.5	14.1	11.5	8.9	16.5	13.5
2	12.3	10.4	16.9	13.5	10.7	8.5	15.9	13.5
2.5	12.0	10.1	16.1	14.4	10.3	8.1	15.6	13.2
3	11.8	9.9	16.9	15.4	9.9	7.8	15.4	12.9
3.5	11.7	9.8	17.0	15.2	9.5	7.5	15.0	12.6
4	11.5	9.7	17.8	14.3	9.3	7.4	14.8	12.5
4.5	11.5	9.5	17.1	14.5	9.1	7.3	14.7	12.0
5	11.5	9.4	17.1	14.6	9.1	7.3	14.5	11.7
6	11.5	9.4	16.2	14.0	8.6	7.1	13.8	11.8
8	11.2	9.4	14.5	11.9	8.1	7.0	12.3	9.8
10	10.4	9.0	14.3	11.9	7.8	6.8	10.7	8.4

All these algorithms were run on randomly chosen sets of stocks. The number of selected stocks is shown in the first column of Tables VII and VIII. These tables indicate the arithmetic averages of the wealths achieved. In each line, the results of the algorithm which outperformed its competitors the more often are set in bold face. Globally, B1EXP seems to have the best results in terms of accumulated wealth, but there are some fine variations which should be mentioned. First, EG is better than B1EXP when the portfolio is reduced to two stocks only. The reason that in this case the internal regret is nothing else than the external regret and the exponential weighted algorithm on which EG is based is known to be optimal for the minimization of the external regret. Second, in the presence of transaction costs and for a daily rebalancing, GBH performs well. This is due to its closeness to buy-and-hold. Interestingly enough, it performs considerably better than buy-and-hold, which is known to be valuable in the presence of such heavy transaction costs. Surprisingly enough, GBH2, which was

Table VII. Arithmetic means of the wealths achieved on randomly selected sets of stocks, repeated 100 times. Monthly rebalancing. A different sample was drawn for each line of this table. Top lines correspond to a no transaction cost setting, whereas the bottom lines consider the case of 2% transaction costs.

ST.	EG	B1EXP	B1POL	GBH	GBH2	B2POL	Cover's	UBH	B-CRP
2	**16.2**	16.2	12.4	13.6	13.6	12.0	15.5	13.6	21.0
3	19.3	**19.4**	15.6	16.1	15.5	13.8	18.4	14.9	30.2
5	20.0	**20.3**	16.6	18.0	16.3	13.2	19.6	14.9	39.6
8	21.3	**21.7**	20.9	20.2	17.6	17.4	21.2	15.4	53.9
10	21.2	**21.7**	19.3	20.6	17.7	15.3	21.3	15.2	61.2
12	20.9	**21.5**	18.1	20.5	17.3	16.0	21.1	14.6	62.4
15	21.9	**22.5**	20.4	21.8	18.3	17.6	22.2	15.3	72.3
18	21.0	**21.6**	17.8	21.1	17.9	16.0	21.4	15.0	76.3
20	21.3	**21.9**	19.7	21.5	18.1	17.5	21.8	15.2	80.3
25	21.4	**22.0**	20.5	21.6	18.2	17.1	21.9	15.2	85.9
2	**14.9**	14.9	10.6	13.7	13.7	10.6	14.5	13.7	20.2
3	16.8	**16.8**	11.1	14.9	14.5	9.9	16.2	14.2	26.9
5	18.5	**18.6**	11.5	17.2	16.1	9.6	18.1	15.0	36.3
8	17.8	**17.9**	9.6	17.2	15.9	9.3	17.6	14.7	46.1
10	18.9	**19.1**	10.3	18.3	16.5	8.6	18.8	14.9	51.2
12	19.0	**19.2**	10.4	18.7	17.0	9.4	19.0	15.4	57.4
15	19.9	**20.1**	10.2	19.7	17.6	9.0	19.9	15.7	65.1
18	19.1	**19.3**	8.9	19.0	17.0	7.7	19.2	15.1	67.3
20	18.5	**18.7**	9.2	18.5	16.6	7.7	18.6	14.9	68.1
25	19.1	**19.3**	10.0	19.2	17.2	7.7	19.3	15.3	75.8

designed to be a modification of GBH suffering a low external regret with respect to buy-and-hold, performs quite poorly compared to GBH. Actually, the wealths achieved by GBH2 seem to interpolate those of GBH and the uniform buy-and-hold strategy. Finally, the at first sight naive U-CRP strategy seems to have interesting results, as already noted in Borodin, El-Yaniv, and Gogan (2000), even though there are no theoretical guarantees for its universality (see for instance Table XII).

Finer comparisons

After this global comparison, we compare B1EXP more carefully with the best opponents in case of no transaction costs, which are EG and B1POL. The comparison to EG is done in table IX which shows the geo-

Table VIII. Arithmetic means of the wealths achieved on randomly selected sets of stocks, repeated 100 times. Daily rebalancing. A different sample was drawn for each line of this table. Top lines correspond to a no transaction cost setting, whereas the bottom lines consider the case of 1% transaction costs.

ST.	EG	B1EXP	B1POL	GBH	GBH2	B2POL	Cover's	UBH	B-CRP
2	**19.3**	19.2	11.4	13.6	13.6	10.3	17.2	13.6	20.4
3	24.8	**24.8**	13.0	16.2	15.1	10.8	21.6	13.9	28.8
5	31.6	**32.0**	16.5	23.6	19.4	11.9	29.1	15.6	47.9
8	28.2	**28.5**	16.7	25.2	19.6	13.9	27.4	15.1	59.5
10	26.2	**26.4**	17.5	24.7	19.1	15.2	25.8	14.5	67.3
12	29.0	**29.3**	18.5	27.8	20.4	15.5	28.7	14.6	87.1
15	27.6	**27.8**	18.0	27.2	20.2	15.3	27.7	14.7	98.6
18	29.3	**29.5**	19.1	29.0	21.2	16.2	29.3	15.1	121.8
20	28.1	**28.4**	18.3	28.0	20.8	16.4	28.3	15.0	120.3
25	28.9	**29.0**	19.1	28.9	21.2	17.3	29.0	15.1	153.9
2	**18.4**	18.3	9.7	15.9	15.9	8.3	17.5	15.9	19.0
3	17.4	**17.4**	8.0	15.3	14.9	6.8	16.6	14.4	21.1
5	18.6	**18.6**	5.7	17.0	15.8	4.4	18.0	14.5	28.2
8	18.9	**18.9**	5.0	18.0	15.9	3.9	18.5	13.7	36.7
10	20.3	**20.3**	5.2	19.9	17.5	3.7	20.1	15.1	43.5
12	20.9	20.9	5.3	**20.5**	17.4	4.0	20.7	14.5	51.3
15	19.7	19.6	4.6	**19.8**	17.0	3.7	19.6	14.5	55.3
18	20.7	20.6	4.8	**20.8**	17.8	3.9	20.6	14.9	66.3
20	20.3	20.2	4.2	**20.4**	17.4	3.4	20.2	14.7	71.6
25	20.5	20.3	4.5	**20.6**	17.7	3.6	20.4	15.0	83.7

metric and arithmetic averages obtained, as well as the number of times B1EXP won and also by how much each algorithm outperformed the other. The value of Δ^+ indicates the maximal gap between B1EXP and EG (in the favour of the former) on the 100 elements of the randomly selected sample and Δ^- is in favour of the latter. We conclude from this table that (in case of no transaction costs) B1EXP is quite often better than EG, and even when it is outperformed by EG, the wealth then achieved by EG is just a bit smaller. The difference between the two algorithms seems to be especially large when η is large, that is, for monthly rebalancing and/or many stocks.

Table X reveals that B1POL and B2POL are not serious contenders because of their huge standard deviation and the extreme values. This is also illustrated by the catastrophic results of these algorithms in the

Table IX. Extensive comparison between the performances of EG and B1EXP on the samples of Table VII.

		Geom. Avg.		Arith. Avg.			Max.	
ST.	TC	EG	B1EXP	EG	B1EXP	Freq.	Δ^-	Δ^+
2	0 %	14.0	13.9	16.2	16.2	12	0.47	0.19
3	0 %	17.0	17.0	19.3	19.4	80	0.02	0.17
5	0 %	18.5	18.6	20.0	20.3	82	0.12	2.23
8	0 %	20.4	20.8	21.3	21.7	92	0.17	2.30
10	0 %	20.6	21.1	21.2	21.7	95	0.21	1.53
12	0 %	20.5	21.0	20.9	21.5	99	0.05	1.66
15	0 %	21.5	22.1	21.9	22.5	98	0.08	1.45
18	0 %	20.7	21.3	21.0	21.6	100		1.65
20	0 %	21.2	21.7	21.3	21.9	100		1.74
25	0 %	21.3	21.9	21.4	22.0	100		1.18
2	2 %	13.0	12.9	14.9	14.9	27	0.30	0.22
3	2 %	15.0	15.0	16.8	16.8	65	0.05	0.09
5	2 %	17.4	17.5	18.5	18.6	72	0.20	1.42
8	2 %	17.2	17.3	17.8	17.9	72	0.42	1.36
10	2 %	18.2	18.4	18.9	19.1	82	0.19	1.46
12	2 %	18.6	18.8	19.0	19.2	73	0.27	1.30
15	2 %	19.6	19.8	19.9	20.1	84	0.18	0.85
18	2 %	18.8	19.0	19.1	19.3	81	0.19	1.20
20	2 %	18.3	18.5	18.5	18.7	84	0.30	0.70
25	2 %	19.1	19.3	19.1	19.3	88	0.23	0.50

presence of transaction costs and for a daily rebalancing, see Table VIII. The reason is that B1POL and B2POL reallocate just too quickly, which can be good or bad. This happens because of the property of the polynomial potential that only the nonnegative internal regrets count in the computation of the wealth allocation, and therefore when one stock dominates, almost all the weight is put on it, which is of course dangerous.

Tables XI and XII are given for sake of completeness as well as to allow comparison with Helmbold, Schapire, Singer, and Warmuth (1998). The algorithms are run on portfolios chosen according to the volatilities of the stocks. Three groups were formed by putting the 12 lowest volatility stocks in the first group (L12), then the 12 highest in the second (H12) and the 12 remaining in the third group (M12).

Table X. Statistical characterization of the wealths achieved on the random sample corresponding to 12 stocks without transaction costs and monthly rebalancing. The minimum, arithmetic and geometric averages, maximum, and standard deviation of the achieved wealths are shown.

Stat.	EG	B1EXP	B1POL	GBH	GBH2	B2POL	Cover's	UBH
Min.	13.2	13.6	6.6	13.0	11.5	4.7	13.4	8.8
Ar. av.	20.9	21.5	18.1	20.5	17.3	16.0	21.1	14.6
Geo. av.	20.5	21.0	16.1	20.1	17.0	13.8	20.7	14.4
Max.	32.9	34.6	56.3	31.7	24.9	60.9	33.7	20.9
St. dev.	4.6	4.9	9.3	4.3	3.2	9.5	4.7	2.8

Table XI. Volatilities (multiplied by 100) for portfolios chosen according to their volatilities, for monthly rebalancing (top lines) as well as for daily rebalancing (bottom lines).

Ptf.	EG	B1EXP	B1POL	GBH	GBH2	B2POL	Cover's	UBH	U-CRP
L12	4.20	4.20	4.61	4.21	4.25	4.64	4.20	4.31	4.20
M12	4.68	4.67	6.32	4.68	4.77	6.71	4.68	4.93	4.67
H12	6.79	6.74	8.12	6.78	6.89	8.32	6.77	7.13	6.73
L24	4.32	4.30	5.66	4.31	4.40	5.84	4.31	4.55	4.30
H24	5.40	5.35	7.40	5.37	5.44	7.94	5.35	5.61	5.35
A36	4.87	4.81	6.94	4.83	4.94	7.21	4.81	5.13	4.81
L12	0.83	0.83	0.88	0.83	0.84	0.89	0.83	0.85	0.83
M12	0.88	0.88	1.11	0.88	0.90	1.14	0.88	0.93	0.88
H12	1.17	1.16	1.82	1.20	1.20	1.96	1.17	1.28	1.15
L24	0.82	0.82	1.01	0.83	0.84	1.03	0.82	0.86	0.82
H24	0.92	0.91	1.45	0.93	0.96	1.54	0.92	1.03	0.91
A36	0.85	0.85	1.25	0.85	0.88	1.28	0.85	0.94	0.85

The group formed by L12 and M12 is called L24, the one of M12 and H12 is denoted by H24. Finally, the set of all 36 stocks is refered to as A36. Note that the B1EXP strategy has almost always the lowest volatilities. Thanks to its agressive rebalancing, the B1POL strategy has interesting achieved wealths for monthly rebalancing. Nevertheless, the B1EXP investment scheme has globally the higher returns.

Table XII. Wealths achieved by the portfolios of Table XI. In each line, the wealth obtained by the best adaptive algorithm is set in bold face.

Ptf.	EG	B1EXP	B1POL	GBH	GBH2	B2POL	Cover's	UBH	U-CRP
L12	10.9	**11.1**	7.6	10.8	10.1	7.7	11.0	9.4	11.2
M12	17.2	17.1	**22.9**	17.1	16.9	21.9	17.0	16.7	17.1
H12	36.3	**39.0**	12.8	34.6	25.3	10.2	37.8	17.6	39.8
L24	13.9	14.0	**19.8**	14.0	13.5	15.7	14.1	13.1	14.1
H24	26.7	27.8	**41.3**	27.1	21.8	21.7	27.6	17.2	28.0
A36	20.5	21.1	**30.9**	20.8	17.5	22.5	20.7	14.5	21.1
L12	12.3	**12.4**	6.7	12.0	11.1	6.5	12.2	10.1	12.4
M12	16.1	**16.2**	9.9	15.8	14.8	9.4	16.0	13.9	16.2
H12	78.1	**81.0**	40.8	67.9	40.2	21.9	76.0	19.5	81.9
L24	14.3	**14.4**	9.3	14.2	13.1	9.0	14.4	12.0	14.4
H24	38.2	**38.7**	25.6	38.1	26.1	21.9	38.6	16.7	38.8
A36	26.9	**27.1**	20.2	27.1	20.2	17.4	27.0	14.5	27.1