Tugkan Batu and Pongphat Taptagaporn

# Competitive portfolio selection using stochastic predictions

## Book section

**Original citation:**
Originally published in Batu, Tugkan and Taptagaporn, Pongphat (2016) *Competitive portfolio selection using stochastic predictions.* In: Lecture Notes in Artificial Intelligence. Lecture Notes in Computer Science. Springer, Cham, Switzerland

# Competitive Portfolio Selection Using Stochastic Predictions

Tuğkan Batu[1] and Pongphat Taptagaporn[1]

Department of Mathematics, London School of Economics, London, U.K.
{t.batu,p.taptagaporn}@lse.ac.uk

**Abstract.** We study a portfolio selection problem where a player attempts to maximise a utility function that represents the growth rate of wealth. We show that, given some stochastic predictions of the asset prices in the next time step, a sublinear expected regret is attainable against an optimal greedy algorithm, subject to tradeoff against the "accuracy" of such predictions that learn (or improve) over time. We also study the effects of introducing transaction costs into the model.

## 1 Introduction

In the field of portfolio management, the problem of how to distribute wealth among a number of assets to maximise wealth gain (or some notion of utility, e.g., mean-variance tradeoff) has been the focus of much academic and industrial research. Most of the studies in this field were previously from the perspective of financial mathematics and economics, and would usually assume some underlying distribution for the price process, e.g., Brownian Motion.

In the 1990's, a new field emerged that uses online learning to design growth-optimal portfolio selection models, following Cover's original work [8]. This model was shown to be competitive to the best CRP: an investment strategy that maintains a fixed proportion of wealth in each of the $m$ assets for each time step, performing any required rebalancing as to maintain these proportions as the asset prices change. In particular, Cover showed sublinear regret on all possible outcomes of price sequence

$$\max_{x^T} \left( \log S_T^* - \log \hat{S}_T \right) = O(\log T),$$

where $S_T^*$ and $\hat{S}_T$ are the wealth obtained by the best CRP and Cover's universal portfolio over $T$ time steps (for some price sequence $x^T$), respectively. Most interestingly, the sublinear regret implies that the (per time step) log-wealth growth achieved by Cover's model converges to that of the best CRP as $T \to \infty$, without making any assumption on the price process (that is, in a model-free sense).

### 1.1 Our Contributions

Our result goes beyond the restriction imposed by the CRP, and instead, we devise a model that is competitive with the best greedy portfolio in a stochastic setting: one that makes the optimal decision as if it knows the next time step's price. To do this, we suppose that our model has access to a price prediction $\tilde{x}_t$ (of the next time step, $t+1$) that follows some probability distribution $\tilde{x}_t \sim \mathcal{D}_t(x_t)$, where $x_t$ is the later observed price change. In this model, we quantify the precise relationship between the expected regret and the accuracy of such predictions. Note that we allow the prediction accuracy to vary over time, as reflected by the dependence of $\mathcal{D}_t$ on the current time step $t$. We demonstrate that for certain probability distributions $\mathcal{D}_t$,

$$\mathbb{E}_{\tilde{x}_t \sim \mathcal{D}_t(x_t)} \left[ \max_{x^T} \left( \log S_T^* - \log \hat{S}_T \right) \right] = o(T)$$

is attainable, subject to some restrictions on the accuracy of $\tilde{x}_t$'s: namely, that the integral of the tail probabilities (of misestimation) must converge to zero as $t$ grows. Intuitively, this is equivalent to improving our predictions through learning from past outcomes, and the requirement is that the model must be learning at a rate fast enough as to satisfy a certain sufficient condition that we will later prove. We also show a bound on the variance of regret in these cases.

Note that we also consider transaction costs for transferring wealth between assets (similar to Blum and Kalai [6]), as there is usually costs associated with buying and selling financial assets in practice (spreads, brokerage charge, etc.). However, we will prove that sublinear expected regret (over all possible price paths) is not attainable in the case of non-zero transaction costs (unless we assume that the price increases in each time step are independently distributed), unlike in the case of zero transaction costs.

Lastly, we show that our portfolio selection model can be computed efficiently using linear programming.

### 1.2 Related Work

The first published work combining the studies of portfolio theory with regret minimisation was by Cover [8]. Since then, there has been much follow up work and extensions to Cover's original portfolio model. Of particular interest to us, Blum and Kalai [6] extended the original model to account for transaction costs. However, the transaction costs plays a minor role in the Blum and Kalai model as it does not affect the decision process beyond that the penalty reduces the wealth that was retained. In particular, there was no cost-versus-wealth tradeoff, to assess whether shifting the portfolio would be beneficial over the cost this would incur, due to the limitation of the CRP model. We introduced a counterpart to the above that balances the reward from rebalancing the portfolio (based on information received from a price prediction) against the transaction cost incurred, and find an optimal point in between as to maximise cost adjusted wealth again.

Transaction costs aside, we compare our model to a less restrictive benchmark than in [8] because the best greedy portfolio is at least as good as the best CRP (in terms of the wealth obtained). However, we instead proved a bound on expected regret (as a function of the distributions $\mathcal{D}_t$) rather than worst-case regret, as we assume that we have additional knowledge in the form of price predictions, bringing us from an adversarial setting to a stochastic one. Note that when considering non-zero transaction costs, neither the greedy portfolio nor the best CRP is strictly better than the other.

Some other works that introduced notions similar to predictions [9, 3] used a concept called "side information". This is where the adversary reveals a side information (say, an integer between 1 and $y$) and the CRP restriction is applied on each state separately. In particular, there is now $y$ different CRPs that may be used, depending on the side information in that particular time step. The benchmark in this case is the best set of $y$ CRP's that achieves the best wealth, given the observed sequence of side information. However, the regret bound of this model assumes that $y$ is finite and does not grow with $T$, meaning that sublinear regret does not hold if the benchmark model uses a different portfolio in every time step (i.e., the side information never repeats). We do not have such restriction in our model.

More recent efforts to incorporate predictions into online learning problems can be found in [7, 16]; these works look at the more general case of convex loss functions, but their regret is still benchmarked against the best CRP (which is substantially weaker than the best greedy portfolio). Some other variants of universal portfolio models can be found in [1, 2, 4, 18, 15, 11, 10, 13]. Most of these models are based on the idea of taking a weighted combination of CRPs over the set of all possible portfolio vectors.

Portfolio optimisation is a fundamental problem studied in mathematical finance literature [17, 14], wherein models with stochastic price changes is the norm. For example, price changes distributed log-normally is analogous to Geometric Brownian Motion [5, 12, 14], a well-understood model used in that field. However, our study and model, motivated by a machine learning perspective to maximise growth-rate of wealth (as opposed to, say, mean-variance optimisation in modern portfolio theory) yields incomparable results.

## 2 Preliminaries

Consider the scenario where we have $m$ assets available for trading over $T$ time steps. Define $x_t = (x_t(1), \ldots, x_t(m)) \in \mathbb{R}_+^m$ as a real-valued vector of price relatives at time step $t$; the $i$-th element of this vector is the ratio of the respective true market prices of Asset $i$ at time $t$ and time $t-1$. For convention, $x_t$ is defined for $1 \leq t \leq T$, and we denote by $x^T$ the vector $(x_1, \ldots, x_T)$. The space $\mathcal{B}$ of portfolio vectors is defined as

$$\mathcal{B} := \{b \in \mathbb{R}_+^m : \sum_{i=1}^{m} b(i) = 1\},$$

where $b(i)$ is the proportion of the portfolio $b$'s total wealth allocated to Asset $i$. Typically, we may need to redistribute wealth between assets as to obtain the portfolio vector chosen for the next time step. We will call this process of redistributing wealth *rebalancing*. We denote by $\theta(b, b', x)$ the multiplicative factor of decrease in wealth due to rebalancing from portfolio $b$ (after observing the price change $x$) to portfolio $b'$, which we will define in more details in the next section. Then, we can define the wealth of a portfolio model $(b_1, \ldots, b_T)$ as

$$S_T = \prod_{t=1}^{T} b_t x_t \theta(b_{t-1}, b_t, x_{t-1}).^1$$

As a convention, we assume that there are no transaction costs associated with the initial positioning before the first time step: that is, $b_0 := b_1$, $x_0 = (1, \ldots, 1)$, and, thus, $\theta(b_0, b_1, x_0) = 1$. Broadly speaking, $S_T$ is the product of the wealth change across all time steps $t = 1, \ldots, T$, where, at each step, we first pay a factor of $\theta(b_{t-1}, b_t, x_{t-1})$ transaction cost for rebalancing $b_{t-1}$ to $b_t$, and then experience a change $b_t x_t$ in wealth, once the price change is observed. Similarly, for the portfolio models denoted as $(\hat{b}_1, \ldots, \hat{b}_T)$ and $(b_1^*, \ldots, b_T^*)$, respectively, we will use $\hat{S}_T$ and $S_T^*$, respectively, to denote the wealth generated by the corresponding portfolio model.

Note that a CRP (from [8]) imposes the additional constraint that the portfolio vector is the same throughout every time step, that is, $b_1 = ... = b_T$.

Although the portfolio model investigated here has the restriction that all the wealth must be invested in one of the $m$ assets, this can be extended to a portfolio of $m + 1$ assets where the first $m$ asset is as before, and the last one represents cash. Therefore, the returns $x_t$ now has $m + 1$ dimension where the last element could represent risk-free interest rate, analogous to much of the work in financial mathematics.

### 2.1 Transaction Costs

The concept of transaction costs was first introduced into the study of online portfolios selection by Blum and Kalai [6], wherein their model charge a fixed percentage of commission on the purchase, but not on the sale, of assets. This is equivalent to charging commission on the purchase and sale of assets equally, as the wealth from any asset we sold will have to be used to purchase another asset (by the constraints of the problem setting). We will use the same model here, though the choice of model doesn't significantly affect our results.

Given portfolio vectors $b_{t-1}, b_t \in \mathcal{B}$ and price-relatives vector $x_{t-1}$, we want to rebalance from the vector $b'_{t-1} := b_{t-1} \cdot x_{t-1} \in \mathbb{R}^m$ to $b_t \in \mathcal{B} \subset \mathbb{R}^m$. Given a transaction cost factor $c \in [0, 1]$ indicating the proportion of cost to be paid from the value of assets purchased, the proportion of wealth retained after rebalancing can be expressed recursively as

$$\theta := \theta(b_{t-1}, b_t, x_{t-1}) = 1 - c \sum_{i:\beta_i > 0} \beta_i,$$

[1] The notations $b_t x_t$ is used as a short-hand for vector dot product.

where $\beta_i = \theta b_t(i) - b_{t-1}(i) \cdot x_{t-1}(i) = \theta b_t(i) - b'_{t-1}(i)$ indicates the quantity of Asset $i$ that needs to be sold or bought, depending on its sign. Intuitively, $\theta$ represents the proportion of the total wealth left after rebalancing. In the worst case, the market value of $b'$ is at least $1 - c$ of the market value of $b$ after rebalancing. In particular, rebalancing a portfolio will always retain at least $1-c$ proportion of its wealth.

### 2.2 Problem Setting

At time $t \in [T]$, suppose our model has access to a prediction such that it follows some probability distribution with respect to the later observed price change: that is, $\tilde{x}_t \sim \mathcal{D}_t(x_t)$. Note that the distribution $\mathcal{D}_t$ may depend on the current time step $t$ (hence, the subscript) and $x_t$, possibly hiding further dependencies on additional parameters such as variance. Based on this prediction, we can compute a portfolio vector as to optimise the wealth.

**Definition 1 (Portfolio Model).** *For each $t \in [T]$, given a predicted price-change $\tilde{x}_t$ of the observed price change $x_t$ such that $\tilde{x}_t \sim \mathcal{D}_t(x_t)$ for some probability distribution $\mathcal{D}_t$, the portfolio vector at time $t$ is specified by*

$$\hat{b}_t := \arg\max_{b \in \mathcal{B}} b\tilde{x}_t \theta(\hat{b}_{t-1}, b, x_{t-1}).$$

*Our benchmark model, which we call the optimal greedy portfolio, is defined similarly as, for each time $t$,*

$$b^*_t = \arg\max_{b \in \mathcal{B}} b x_t \theta(b^*_{t-1}, b, x_{t-1}).$$

Note that the above models considers the tradeoff between the transaction cost of shifting to a "better" portfolio against the expected benefit of doing such a rebalancing given the prediction or actual outcome, respectively. In the case where the optimisation yields multiple solutions, we canonically choose the one with the least transaction costs. This will be made more precise in Section 5.

## 3 Main Results

In this section, we present our technical contributions. In particular, we investigate how close the wealth of our portfolio model is to the benchmark model, in expectation over the random choices of $\tilde{x}_t \sim \mathcal{D}_t(x_t)$ and adversarially chosen $x_t$, for $t \in [T]$.

Firstly, we show the expected-regret bound of the portfolio model $\hat{b}$ against $b^*$, in terms of the distribution of the predicted price change $\tilde{x}_t$ relative to the later observed price change $x_t$. This will lead us to a sufficient condition to obtain a sublinear expected regret (and, additionally, sublinear variance of regret) in the case of zero transaction costs. Then, we show that sublinear expected regret is unattainable in general in the case of non-zero transaction costs, no matter how small $c > 0$ is.

### 3.1 Expected-Regret Bound

As a measure of performance, we consider the expected-regret $\mathbb{E}[R]$ of our portfolio model against the optimal greedy portfolio model: namely,

$$\mathbb{E}_{\tilde{x}_t \sim \mathcal{D}_t(x_t)} \Big[ \max_{x^T} \Big( \log S_T^* - \log \hat{S}_T \Big) \Big] .$$

This can be interpreted as enumerating through all possible price predictions $\tilde{x}^T$ and choosing the outcome of price sequence $x^T$ that maximises regret for each choice of $\tilde{x}^T$. Each of these choices of $\tilde{x}^T$ occurs with some probability depending on $x^T$ and $\mathcal{D}_t$ for $t \in [T]$, and we take the expectation over these probabilities.

We analyse the expected regret $\mathbb{E}[R]$, where the choices of portfolio vectors depend directly on the random choices of $\tilde{x}_t \sim \mathcal{D}_t(x_t)$ and $x_t$ is chosen adversarially, for each $t \in [T]$. The theorem below gives an upper bound on the expected regret as a function of the distributions $\mathcal{D}_t$ of predictions in each time step.

**Theorem 2.** *The expected regret of our portfolio model from Definition 1 can be bounded from above as*

$$\mathbb{E}[R] \leq \gamma + 2 \sum_{t=1}^{T} \int_0^\infty \Pr_{\tilde{x}_t \sim \mathcal{D}_t(x_t)} [\tilde{x}_t \notin (e^{-z} x_t, e^z x_t)] \, dz \, ,$$

*where $\gamma$ accounts for the regret arising from the positioning error of our portfolio and is defined as*

$$\gamma = - \sum_{t=1}^{T} \mathbb{E} \Big[ \log \frac{\theta(\hat{b}_{t-1}, b_t^*, x_{t-1})}{\theta(b_{t-1}^*, b_t^*, x_{t-1})} \Big] .$$

*Proof.* We fix some time $t$ and consider the ratio of the single-time-step wealth change of our portfolio to that of the benchmark at time $t$ in order to bound the regret arising from that time step. The regret associated with the time step $t$ has two sources: positioning error of the current portfolio that results in transaction costs and inaccurate price predictions. We define

$$\rho_t = \frac{\theta(\hat{b}_{t-1}, b_t^*, x_{t-1})}{\theta(b_{t-1}^*, b_t^*, x_{t-1})}$$

to capture the regret arising from the positioning error of the portfolio at time step $t$: for example, when $b_{t-1}^*$ was in a better position than $\hat{b}_{t-1}$ to minimise transaction costs when rebalancing at time $t$.

Now, suppose that $(1-\delta) x_t \preceq \tilde{x}_t \preceq (1-\delta)^{-1} x_t$,[2] at time step $t$, for some $\delta$ such that $0 \leq \delta < 1$. Then, for any $\hat{b}_t, b_t^*, \hat{b}_{t-1}, b_{t-1}^* \in \mathcal{B}$, we have the following

[2] The notations $\preceq$, $\succeq$, $\prec$, and $\succ$ denote component-wise vector inequalities.

bound on the ratio of the single-time-step wealths:

$$\frac{\hat{b}_t x_t \theta(\hat{b}_{t-1}, \hat{b}_t, x_{t-1})}{b_t^* x_t \theta(b_{t-1}^*, b_t^*, x_{t-1})} \geq (1-\delta) \frac{\hat{b}_t \tilde{x}_t \theta(\hat{b}_{t-1}, \hat{b}_t, x_{t-1})}{b_t^* x_t \theta(b_{t-1}^*, b_t^*, x_{t-1})} \tag{1}$$

$$\geq (1-\delta)^2 \frac{\hat{b}_t \tilde{x}_t \theta(\hat{b}_{t-1}, \hat{b}_t, x_{t-1})}{b_t^* \tilde{x}_t \theta(b_{t-1}^*, b_t^*, x_{t-1})} \tag{2}$$

$$\geq (1-\delta)^2 \rho_t. \tag{3}$$

In the above, (1) is due to $x_t \succeq (1-\delta)\tilde{x}_t$, (2) is due to $\tilde{x}_t \succeq (1-\delta)x_t$, and (3) is due to the fact that

$$\hat{b}_t \tilde{x}_t \theta(\hat{b}_{t-1}, \hat{b}_t, x_{t-1}) \geq b_t^* \tilde{x}_t \theta(\hat{b}_{t-1}, b_t^*, x_{t-1}) = \rho_t b_t^* \tilde{x}_t \theta(b_{t-1}^*, b_t^*, x_{t-1}),$$

as $\hat{b}_t$ was chosen to maximise its single-time-step wealth by Definition 1. For each time step $t \in [T]$, we define deviation $\delta_t$ of $x_t$ and $\tilde{x}_t$ as

$$\delta_t := \min\{\delta \geq 0 \mid (1-\delta)x_t \succeq \tilde{x}_t \succeq (1-\delta)^{-1} x_t\}.$$

Intuitively, this is the deviation of the predicted price change from the observed price change. We can now calculate the expected regret as follows.

$$\mathbb{E}[R] = \mathbb{E}\Big[\max_{x^T} \log\Big(\frac{S_T^*}{\hat{S}_T}\Big)\Big]$$

$$= \mathbb{E}\Big[\max_{x^T} \log\Big(\prod_{t=1}^{T} \frac{b_t^* x_t \theta(b_{t-1}^*, b_t^*, x_{t-1})}{\hat{b}_t x_t \theta(\hat{b}_{t-1}, \hat{b}_t, x_{t-1})}\Big)\Big]$$

$$\leq \mathbb{E}\Big[\log\Big(\prod_{t=1}^{T} (1-\delta_t)^{-2} \rho_t^{-1}\Big)\Big] \tag{4}$$

$$\leq \sum_{t=1}^{T} 2\mathbb{E}\Big[-\log(1-\delta_t)\Big] - \mathbb{E}\Big[\log \rho_t\Big], \tag{5}$$

where (4) is by the inequality from (3), and (5) follows from linearity of expectation. We now will now use $\gamma = -\sum_{t=1}^{T} \mathbb{E}[\log \rho_t]$ to denote the "positioning error," and continue our analysis of the first term on the right hand side of the inequality.

$$\sum_{t=1}^{T} \mathbb{E}\Big[-\log(1-\delta_t)\Big] = \sum_{t=1}^{T} \int_0^\infty \Pr_{\tilde{x}_t}[-\log(1-\delta_t) \geq z]\, dz$$

$$= \sum_{t=1}^{T} \int_0^\infty \Pr_{\tilde{x}_t}[1-\delta_t \leq e^{-z}]\, dz,$$

$$= \sum_{t=1}^{T} \int_0^\infty 1 - \Pr_{\tilde{x}_t}[1-\delta_t > e^{-z}]\, dz,$$

$$= \sum_{t=1}^{T} \int_0^\infty 1 - \Pr_{\tilde{x}_t}[e^{-z} x_t \prec \tilde{x}_t \prec e^{z} x_t]\, dz,$$

where the last line above is obtained from applying the definition of $\delta_t$, giving us the bound on expected regret. □

Note that the quantity $\gamma$ in Theorem 2 captures the positioning error of our model arising from transaction costs. Hence, in the absence of transaction costs (that is, when $c = 0$), we have that $\gamma = 0$. In fact, we later prove in Section 3.3 that, in general, $\gamma = \Omega(T)$ for non-zero transaction costs (that is, when $c > 0$), by showing that there exists a sequence $x^T$ that yields an expected regret at least linear in $T$.

We also observe that $\gamma = 0$ in the weaker case when $x_t$ is a random variable that is independent of $x_{t-1}$ (hence, also independent of $b^*_{t-1}$ and $\hat{b}_{t-1}$), for all time steps $t \in [T]$, whereas Theorem 2 is stronger as it makes no assumption on how $x_t$ are chosen. This is because

$$\mathbb{E}[\log \theta(b^*_{t-1}, b^*_t, x_{t-1})] = \mathbb{E}[\log \theta(\hat{b}_{t-1}, \hat{b}_t, x_{t-1})],$$

intuitively meaning that the random choice of $x_t$ and $\tilde{x}_t$ are just as likely be favourable to $b^*_{t-1}$ as it is to $\hat{b}_{t-1}$. For example, suppose that we define $\tilde{x}_t = (1, ..., 1)$ and $x_t$ is drawn from some log-normal distribution with mean $\tilde{x}_t$. Then, this is equivalent to assuming that the returns $x_t$ follows a Geometric Brownian Motion and that the current price is the best prediction of the next time step's price; similar to the assumption surrounding much of the work in financial mathematics.

Finally, setting $\gamma$ aside, the result above gives us a good intuition on what the expected regret looks like. Namely, in each time step the regret can be thought of to be no larger than the sum of an integral of the tail probabilities. Having a small expected regret then hinges on bounding these tail probabilities.

### 3.2 Variance-of-Regret Bound

We can now prove a bound on the variance of regret, using much of the ideas from the proof of the bound on expected regret in Theorem 2.

**Theorem 3.** *The variance of regret of our portfolio model from Definition 1 can be bounded from above as*

$$\mathrm{Var}[R] \le \eta + 4 \sum_{t=1}^{T} \int_0^\infty \Pr_{\tilde{x}_t \sim \mathcal{D}_t(x_t)} [\tilde{x}_t \notin (e^{-\sqrt{z}} x_t, e^{\sqrt{z}} x_t)]\, dz\,,$$

*where $\eta$ accounts for the variance in the regret arising from the positioning error and the covariance of the single-time-step wealth ratios, defined as*

$$\eta = -\sum_{t=1}^{T} \mathrm{Var}\Big[\log \frac{\theta(\hat{b}_{t-1}, b^*_t, x_{t-1})}{\theta(b^*_{t-1}, b^*_t, x_{t-1})}\Big] + \sum_{t=1}^{T} \sum_{j \ne t} \mathrm{cov}\Big[\frac{b^*_t x_t \theta(b^*_{t-1}, b^*_t, x_{t-1})}{\hat{b}_t x_t \theta(\hat{b}_{t-1}, \hat{b}_t, x_{t-1})}, \frac{b^*_j x_j \theta(b^*_{j-1}, b^*_j, x_{j-1})}{\hat{b}_j x_j \theta(\hat{b}_{j-1}, \hat{b}_j, x_{j-1})}\Big].$$

*Proof.*

$$\begin{aligned}
\mathrm{Var}[R] &= \mathrm{Var}\Big[\max_{x^T} \log\Big(\frac{S_T^*}{\hat{S}_T}\Big)\Big] \\
&= \mathrm{Var}\Big[\max_{x^T} \log\Big(\prod_{t=1}^{T} \frac{b_t^* x_t \theta(b_{t-1}^*, b_t^*, x_{t-1})}{\hat{b}_t x_t \theta(\hat{b}_{t-1}, \hat{b}_t, x_{t-1})}\Big)\Big] \\
&\leq \mathrm{Var}\Big[\log\Big(\prod_{t=1}^{T} (1-\delta_t)^{-2} \rho_t^{-1}\Big)\Big] \\
&\leq \eta + 4 \sum_{t=1}^{T} \mathrm{Var}\Big[-\log(1-\delta_t)\Big],
\end{aligned}$$

where $\eta$ is the term representing the positioning errors and covariance terms, as described in the theorem statement. We continue to simplify the remaining part of the equation, making use of the inequality $\mathrm{Var}[R] \leq \mathbb{E}[R^2]$. Thus, we get

$$\begin{aligned}
\sum_{t=1}^{T} \mathrm{Var}\Big[-\log(1-\delta_t)\Big] &\leq \sum_{t=1}^{T} \mathbb{E}\Big[(-\log(1-\delta_t))^2\Big] \\
&= \sum_{t=1}^{T} \int_0^\infty \Pr_{\tilde{x}_t}[-\log(1-\delta_t) \geq \sqrt{z}]\, dz \\
&= \sum_{t=1}^{T} \int_0^\infty \Pr_{\tilde{x}_t}[1-\delta_t \leq e^{-\sqrt{z}}]\, dz, \\
&= \sum_{t=1}^{T} \int_0^\infty 1 - \Pr_{\tilde{x}_t}[1-\delta_t > e^{-\sqrt{z}}]\, dz, \\
&= \sum_{t=1}^{T} \int_0^\infty 1 - \Pr_{\tilde{x}_t}[e^{-\sqrt{z}} x_t \prec \tilde{x}_t \prec e^{\sqrt{z}} x_t]\, dz,
\end{aligned}$$

where the last line above is obtained from applying the definition of $\delta_t$ (as defined in the proof of Theorem 2), giving us the desired result. □

Similarly to the case for expected regret discussed in the previous section, we also have that $\eta = 0$ in the zero-transaction cost scenario (that is, $c = 0$) or $x_t$ is independently distributed from $x_{t-1}$ for $t \in [T]$.

### 3.3 Linear Expected Regret for Non-zero Transaction Costs

We will now show that for any class of non-trivial distributions $\mathcal{D}_t$, the expected-regret bound above will not be sublinear for non-zero transaction cost (in effect, showing that $\gamma$ is not necessarily sublinear for any $c > 0$). This is because there exists a sequence of returns $x_t$ for $t \in [T]$ that will favour $b_t^*$ position, hence, yielding a large enough long-term regret. Here, we define a *non-trivial* *distribution* as one where the preimage of the cumulative distribution function is non-empty at some value inside a constant interval around $\frac{1}{2}$. Note that any class of continuous distributions satisfies this criteria.

**Theorem 4.** *Given non-trivial $\mathcal{D}_t$, for all $t \in [T]$, $\mathbb{E}[R] = \Omega(T)$ when transaction cost $c$ is non-zero.*

*Proof.* To prove that the expected regret is not necessarily sublinear in the case of non-zero transaction cost, it is enough to come up with a sequence of $x_t$ that breaks this sub-linearity. Therefore, we will give a way to construct such $x_t$ for each $t \in [T]$ in the two-asset case ($m = 2$), where $b_t^*$ and $\hat{b}_t$ will always take the values of either $(0, 1)$ or $(1, 0)$ by our construction of the re-balancing scheme from Section 5.

For time step $t$, assume that $\hat{b}_{t-1} = (0, 1)$, without loss of generality, with $b_{t-1}^*$ is $(0, 1)$ or $(1, 0)$. We will calculate the single-time-step loss

$$\frac{b_t^* x_t \theta(b_{t-1}^*, b_t^*, x_{t-1})}{\hat{b}_t x_t \theta(\hat{b}_{t-1}, \hat{b}_t, x_{t-1})}$$

in these two cases separately.

**State 1 (Different)** $b_{t-1}^* = (1, 0)$

The adversary chooses $x_t = (1, 1 - c)$, resulting in a single-time-step loss of $\frac{1}{1-c}$, regardless of the choice $\tilde{x}_t \sim \mathcal{D}_t(x_t)$.

**State 2 (Same)** $b_{t-1}^* = (0, 1)$

The adversary chooses $x_t = (\xi_t, 1)$, where $\xi_t$ is chosen such that

$$\Pr_{\tilde{x}_t \sim \mathcal{D}_t((\xi_t, 1))}\left[\frac{\tilde{x}_t(1)}{\tilde{x}_t(2)} > \frac{1}{1-c}\right] = \frac{1}{2}.$$

Intuitively, this is the choice of price relative vector where the portfolio model (as represented by $\hat{b}_t$) has equal probabilities of shifting or staying put. This implies that $\Pr_{\tilde{x}_t \sim \mathcal{D}_t(x_t)}[\hat{b}_t = b_t^*] = \frac{1}{2}$, and the single-time-step loss may be as small as 1 in this case. Note that this choice of $\xi_t$ exists if the preimage of the CDF of $\mathcal{D}_t$ at $\frac{1}{2}$ is non-empty. One can easily extend this proof to cases where the preimage of the CDF is non-empty at some value inside a constant interval around $\frac{1}{2}$.

With this information, we can model the dynamics of the portfolio as a Markov chain with these two states (Different and Same). The transition probability matrix of that Markov chain, assuming worst-case, i.e., the lowest probability of staying in "different", is

$$\begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

which implies a limiting distribution $\pi = (\frac{1}{3}, \frac{2}{3})$. Using this, the expected regret (over all possible $x_t$) can be lower-bounded by the linear expected regret (over the particular choice of $x_t$, as described above).

$$\begin{aligned}
\mathbb{E}[R] &= \mathbb{E}\Big[\max_{x^T} \log\Big(\frac{S_T^*}{\hat{S}_T}\Big)\Big] \\
&\geq \mathbb{E}\Big[\log\Big(\frac{S_T^*}{\hat{S}_T}\Big)\Big] \\
&= \sum_{t=1}^{T} \mathbb{E}\Big[\log \frac{b_t^* x_t \theta(b_{t-1}^*, b_t^*, x_{t-1})}{\hat{b}_t x_t \theta(\hat{b}_{t-1}, \hat{b}_t, x_{t-1})}\Big] \\
&= -\frac{1}{3}\sum_{t=1}^{T} \log(1-c) = \Theta(T),
\end{aligned}$$

where the last line follows from the fact that the portfolio needs to shift all its wealth in one third of the steps in the long run (due to the limiting distribution of the Markov chain above), each of which incurs a loss factor of $1 - c$. □

So now we have established that we cannot hope for sublinear expected regret in the presence of transaction costs, no matter the choice of $\mathcal{D}_t$ (as long as it is non-trivial). However, we will later show in Section 4 that a few sensible choices for $\mathcal{D}_t$ will indeed yield sublinear expected regret (and variance of regret) in the case $c = 0$.

## 4 Special Cases for the Distributions of Predictions

Given the above results are for a generically distributed $\tilde{x}_t \sim \mathcal{D}_t(x_t)$, we will now look at some particular cases for $\mathcal{D}_t$ and compute the required quality of prediction in order to achieve sublinear expected regret. Herein we will assume that $c = 0$, as Theorem 4 shows that we cannot hope for sublinear expected regret in the presence of transaction costs.

Firstly, we shall assume that $\mathcal{D}_t$ is parametrised by two variables $\mu_t$ (mean) and $\sigma_t$ (standard deviation). We will look only at log-returns (rather than absolute returns); this is quite a standard notion in financial mathematics for a number of reasons [5, 12, 14]. In particular, we will say that the log-predicted returns ($\ln \tilde{x}_t$) are distributed around the mean (defined as the log-observed returns, $\ln x_t$) with some standard deviation $\sigma_t$. Formally, $\ln \tilde{x}_t \sim \mathcal{D}_{\ln x_t, \sigma_t^2}$ for some distribution $\mathcal{D}$, or simply $\tilde{x}_t \sim \ln \mathcal{D}_{\ln x_t, \sigma_t^2}$ for short-hand. As our portfolio vector is multi-dimensional, we will use $\sigma_t = (\sigma_t, ..., \sigma_t) \in \mathbb{R}_+^m$, apply the logarithm and distribution element-wise: that is,

$$\ln x_t = \ln(x_t(1), ..., x_t(m)) = (\ln x_t(1), ..., \ln x_t(m)),$$

and, thus,

$$\ln \mathcal{D}_{\ln x_t, \sigma_t^2} = \ln \mathcal{D}_{\ln x_t(1), \sigma_t^2} \times ... \times \ln \mathcal{D}_{\ln x_t(m), \sigma_t^2}.$$

Note that Chebyshev's inequality is too loose to obtain a reasonable bound for a generalised distribution $\mathcal{D}$:

$$\mathbb{E}[R] \leq 2\sum_{t=1}^{T}\int_{0}^{\infty} \Pr_{\tilde{x}_t\sim\mathcal{D}_t(x_t)}[\tilde{x}_t \notin (e^{-z}x_t, e^{z}x_t)]\, dz \leq 2\sum_{t=1}^{T}\int_{0}^{\infty}\frac{\sigma_t^2}{z^2}\, dz,$$

where the last inequality is due to Chebyshev's, which states that

$$Pr(|x-\mu| \geq z) \leq \sigma_t^2/z^2.$$

As a result, the last integral evaluates to $+\infty$. Therefore, the next three subsection looks at the required $\sigma_t$, for $t \in [T]$, to obtain sublinear expected regret for three particular cases of $\mathcal{D}$: uniform, linear, and normal.

### 4.1 Log-Uniformly Distributed Predictions

Suppose that $\tilde{x}_t \sim \ln\mathcal{U}_{\ln x_t, \sigma_t^2}$, where $\mathcal{U}$ is the uniform distribution on the log-returns between the range $[-\sigma_t, \sigma_t]$ with the following probability density function

$$f(y) = \begin{cases} \frac{1}{2\sigma_t} & \text{if } 0 \leq |y - \ln x_t| \leq \sigma_t, \\ 0 & \text{otherwise.} \end{cases}$$

In this case, applying Theorem 2 and Theorem 3 yields

$$\mathbb{E}[R] \leq 2\sum_{t=1}^{T}\int_{0}^{\sigma_t} 1 - \frac{z}{\sigma_t}\, dz = \sum_{t=1}^{T}\sigma_t,$$

$$\mathrm{Var}[R] \leq 4\sum_{t=1}^{T}\int_{0}^{\sigma_t} 1 - \frac{\sqrt{z}}{\sigma_t}\, dz = 4\sum_{t=1}^{T}\sigma_t - \frac{2}{3}\sqrt{\sigma_t}.$$

Thus, $\sigma_t \to 0$ at any speed will yield sublinear expected regret and variance of regret, hence, making no other restriction on the required rate of learning.

### 4.2 Log-Linearly Distributed Predictions

Suppose that $\tilde{x}_t \sim \ln\mathcal{L}_{\ln x_t, \sigma_t^2}$, where $\mathcal{L}$ is the linearly-decreasing distribution with largest density at the mean, $\ln x_t$. More precisely, it has the following probability density function

$$f(y) = \begin{cases} \frac{1}{\sigma_t} - \frac{|y-\ln x_t|}{\sigma_t^2} & \text{if } 0 \leq |y - \ln x_t| \leq \sigma_t, \\ 0 & \text{otherwise.} \end{cases}$$

In this case, applying Theorem 2 and Theorem 3 yields

$$\mathbb{E}[R] \leq 2\sum_{t=1}^{T}\int_{0}^{\sigma_t}(1 - 2\frac{z}{\sigma_t} + \frac{z^2}{\sigma_t^2})\, dz = 2\sum_{t=1}^{T}\frac{\sigma_t}{3} = \frac{2}{3}\sum_{t=1}^{T}\sigma_t,$$

$$\mathrm{Var}[R] \leq 4\sum_{t=1}^{T}\int_0^{\sigma_t}(1 - 2\frac{\sqrt{z}}{\sigma_t} + \frac{z}{\sigma_t^2})\,dz = 4\sum_{t=1}^{T}\sigma_t - \frac{4}{3}\sqrt{\sigma_t} + \frac{1}{2} = \Theta(T).$$

so $\sigma_t \to 0$ at any speed will yield sublinear expected regret, but the bound on the variance of regret is linear in $T$.

### 4.3 Log-Normally Distributed Predictions

We will now look at the particular case when $\mathcal{D}_t$ is log-normally distributed (analogous to Geometric Brownian Motion). Suppose that $\tilde{x}_t \sim \ln\mathcal{N}_{\ln x_t,\sigma_t^2}$, then

$$\mathbb{E}[R] \leq 4\sum_{t=1}^{T}\int_0^{\infty}\Pr_{y\sim\mathcal{N}_{0,1}}[y > z/\sigma_t]\,dz.$$

To achieve a sublinear expected regret then depends on the ability to obtain an appropriate sequence of predictions with $\sigma_t$ such that

$$\frac{1}{T}\sum_{t=1}^{T}\int_0^{\infty}\Pr_{y\sim\mathcal{N}_{0,1}}[y > z/\sigma_t]\,dz \to 0,$$

as $T \to \infty$. This has a very natural interpretation; the above condition can be viewed as an integral over the tail probabilities of the standard normal distribution, where the size of the tail is determined by $\sigma_t$.

Clearly, $\sigma_t = O(1)$ for all $t \in [T]$ is not a sufficient condition as the tail probabilities will not tend to zero for small values of $z$, so we must necessarily have that $\sigma_t \to 0$ as $t \to \infty$. However, it is unclear what rate of convergence would be required for this condition to hold. We suspect that $\sigma_t = O(1/\log t)$ suffices, but this remains to be shown and leaves an interesting open question. Similarly, the variance of regret in this case can be bounded as

$$\mathrm{Var}[R] \leq 8\sum_{t=1}^{T}\int_0^{\infty}\Pr_{y\sim\mathcal{N}_{0,1}}[y > \sqrt{z}/\sigma_t]\,dz.$$

## 5 Portfolio Computation

The $\theta$ function can be viewed as a variant of the earth mover's distance, which, in turn, can be formulated as a transportation or flow problem and solved using a linear program. Here, we present an LP for computing $\hat{b}$ (and, hence, for similarly computing $b^*$) by first computing $\theta$. The input to the computation is the original allocation vector $w = (w_1, \ldots, w_m)$ (corresponding to $K\hat{b}$, where $K$ is the total wealth before rebalancing and $b \in \mathcal{B}$) and the target portfolio vector given as $q = (q_1, \ldots, q_m)$ (with $\sum_i q_i = 1$). The variables of the LP are the wealth $W$ resulting after the rebalancing and $f_{ij}$, for $i, j \in [m]$, that corresponds to wealth that needs to be transferred from Asset $i$ to Asset $j$.

$$\max W$$
$$\text{subject to}$$
$$\sum_{j\in[m]} f_{ij} \leq w_i \qquad \forall i = 1, \dots, m \tag{6}$$
$$f_{jj} + (1-c) \cdot \sum_{\substack{i\in[m]\\ i\neq j}} f_{ij} \geq W \cdot q_j \qquad \forall j = 1, \dots, m \tag{7}$$
$$f_{ij} \geq 0 \qquad \forall i, j = 1, \dots, m \tag{8}$$

The constraints in (6) ensure that the wealth transferred out of each asset is bounded by the current wealth in that asset. The constraints in (7) ensure that the wealth that stays in each asset plus the wealth transferred into that asset, minus the incurred transaction costs, are sufficient to reach the target portfolio vector with a total wealth of $W$. Finally, the flow of wealth will always be positive by (8). Note that the sets of constraints in (6) and (7) will be satisfied tightly in an optimal solution. First of all, for any $i \in [m]$, total flow $\sum_{j\in[m]} f_{ij}$ out of Asset $i$ will be equal to $w_i$, because any increase in the total flow $\sum_{i,j} f_{ij}$ can be distributed over the assets according to $q$, creating slack in each constraint in (7) and allowing a strictly larger value for $W$. Similarly, if the flow into any Asset $j$, given as $f_{jj} + (1-c) \cdot \sum_{i\in[m],i\neq j} f_{ij}$, was strictly larger than $W \cdot q_j$, then this excess flow can be shifted to other assets to create slack in each constraint in (7), which, in turn, allows $W$ to be increased. The fact that the constraints in (6) and (7) are tight for an optimal solution shows that all the wealth in the previous time step is used during rebalancing and the resulting portfolio distribution adheres to $q$. Finally, by the maximisation of $W$, we get that the optimal solution to the LP gives the value of $\theta$, and also $\hat{b}$ (by summing up all of the flow in/out of each asset $f_{ij}$). In the case where there are multiple optimal solutions, we choose the one with the lowest $\sum_{j\in[m]} f_{ij}$, for $i = 1, ..., m$ sequentially; that is, we break ties by minimising the outflow from the smallest to the largest $i$.

## References


1. Agarwal, A., Hazan, E.: Efficient algorithms for online game playing and universal portfolio management. Electronic Colloquium on Computational Complexity 13(033) (2006), http://eccc.hpi-web.de/eccc-reports/2006/TR06-033/index.html
2. Agarwal, A., Hazan, E., Kale, S., Schapire, R.E.: Algorithms for portfolio management based on the Newton method. In: Cohen, W.W., Moore, A. (eds.) Machine Learning, Proceedings of the Twenty-Third International Conference (ICML 2006). ACM International Conference Proceeding Series, vol. 148, pp. 9–16. ACM (2006), http://doi.acm.org/10.1145/1143844.1143846
3. Bean, A.J., Singer, A.C.: Universal switching and side information portfolios under transaction costs using factor graphs. In: Proceedings of the IEEE International

Conference on Acoustics, Speech, and Signal Processing, ICASSP 2010. pp. 1986–1989. IEEE (2010), http://dx.doi.org/10.1109/ICASSP.2010.5495255
4. Bean, A.J., Singer, A.C.: Factor graph switching portfolios under transaction costs. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2011. pp. 5748–5751. IEEE (2011), http://dx.doi.org/10.1109/ICASSP.2011.5947666
5. Black, F., Scholes, M.S.: The Pricing of Options and Corporate Liabilities. Journal of Political Economy 81(3), 637–54 (May-June 1973), https://ideas.repec.org/a/ucp/jpolec/v81y1973i3p637-54.html
6. Blum, A., Kalai, A.: Universal portfolios with and without transaction costs. Machine Learning 35(3), 193–205 (1999), http://dx.doi.org/10.1023/A:1007530728748
7. Chiang, C.K., Yang, T., Lee, C.J., Mahdavi, M., Lu, C.J., Jin, R., Zhu, S.: Online optimization with gradual variations. In: COLT. pp. 6–1 (2012)
8. Cover, T.M.: Universal portfolios. Mathematical Finance 1(1), 1–29 (1991), http://dx.doi.org/10.1111/j.1467-9965.1991.tb00002.x
9. Cover, T.M., Ordentlich, E.: Universal portfolios with side information. IEEE Transactions on Information Theory 42(2), 348–363 (1996), http://dx.doi.org/10.1109/18.485708
10. Györfi, L., Walk, H.: Empirical portfolio selection strategies with proportional transaction costs. IEEE Transactions on Information Theory 58(10), 6320–6331 (2012), http://dx.doi.org/10.1109/TIT.2012.2205131
11. Hazan, E., Agarwal, A., Kale, S.: Logarithmic regret algorithms for online convex optimization. Machine Learning 69(2-3), 169–192 (2007), http://dx.doi.org/10.1007/s10994-007-5016-8
12. Karatzas, I., Shreve, S.E.: Brownian motion and stochastic calculus. Graduate texts in mathematics, Springer-Verlag, New York, Berlin, Heidelberg (1991), http://opac.inria.fr/record=b1079144, autres tirages corrigs : 1996, 1997, 1999, 2000, 2005
13. Kivinen, J., Warmuth, M.K.: Averaging expert predictions. In: Fischer, P., Simon, H. (eds.) Computational Learning Theory, 4th European Conference, EuroCOLT '99. Lecture Notes in Computer Science, vol. 1572, pp. 153–167. Springer (1999), http://dx.doi.org/10.1007/3-540-49097-3_13
14. Merton, R.C.: Optimum consumption and portfolio rules in a continuous-time model. Journal of Economic Theory 3(4), 373–413 (December 1971), https://ideas.repec.org/a/eee/jetheo/v3y1971i4p373-413.html
15. Ordentlich, E., Cover, T.M.: On-line portfolio selection. In: Proceedings of the Ninth Annual Conference on Computational Learning Theory. pp. 310–313. COLT '96, ACM, New York, NY, USA (1996), http://doi.acm.org/10.1145/238061.238161
16. Rakhlin, A., Sridharan, K.: Online learning with predictable sequences. In: COLT. pp. 993–1019 (2013)
17. Sharpe, W.F.: Capital Asset Prices: A Theory Of Market Equilibrium Under Conditions Of Risk. Journal of Finance 19(3), 425–442 (09 1964), https://ideas.repec.org/a/bla/jfinan/v19y1964i3p425-442.html
18. Stoltz, G., Lugosi, G.: Internal regret in on-line portfolio selection. Machine Learning 59(1-2), 125–159 (2005), http://dx.doi.org/10.1007/s10994-005-0465-4