An Online Portfolio Selection Algorithm with Regret Logarithmic in Price Variation

Elad Hazan
Technion - Israel Inst. of Tech.
Haifa, 32000 Israel
ehazan@ie.technion.ac.il

Satyen Kale
IBM T.J. Watson Research Center
P.O. Box 218, Yorktown Heights, NY 10598
sckale@us.ibm.com

Abstract

We present a novel efficient algorithm for portfolio selection which theoretically attains two desirable properties:

1. Worst-case guarantee: the algorithm is universal in the sense that it asymptotically performs almost as well as the best constant rebalanced portfolio determined in hindsight from the realized market prices. Furthermore, it attains the tightest known bounds on the regret, or the log-wealth difference relative to the best constant rebalanced portfolio. We prove that the regret of algorithm is bounded by $O(\log Q)$, where Q is the quadratic variation of the stock prices. This is the first improvement upon Cover's [Cov91] seminal work that attains a regret bound of $O(\log T)$, where T is the number of trading iterations.
2. Average-case guarantee: in the Geometric Brownian Motion (GBM) model of stock prices, our algorithm attains tighter regret bounds, which are provably impossible in the worst-case. Hence, when the GBM model is a good approximation of the behavior of market, the new algorithm has an advantage over previous ones, albeit retaining worst-case guarantees.

We derive this algorithm as a special case of a novel and more general method for online convex optimization with exp-concave loss functions.[1]

1 Introduction

A widely used model in mathematical finance for stock prices is the Geometric Brownian Motion (GBM). This model has been successfully applied to study and price financial instruments, and is the basis of much investing in practice. However, while it

[1] An extended abstract of this work appeared in [HK09]

is often an acceptable approximation, the GBM model is not always valid empirically. This motivates a worst-case approach to investing, called universal portfolio management, where the objective is to maximize wealth relative to the wealth earned by the best fixed portfolio in hindsight. In this paper we tie the two approaches, and design an investment strategy which is universal in the worst-case, and yet capable of significantly improved performance when the market is closely approximated by the GBM model.

"Average-case" Investing: Much of mathematical finance theory is devoted to the modeling of stock prices and devising investment strategies that maximize wealth gain (while controlling risk). Typically, such investment strategies involve estimating fitting a parametric model of stock prices to observed data. Such strategies are geared to the *average-case market* (in the formal computer science sense), and are naturally susceptible to drastic deviations from the model, as witnessed in the recent stock market crash.

Even so, empirically the Geometric Brownian Motion (GBM) ([Osb59, Bac00]) has enjoyed great predictive success and every year significant investments are made assuming this model. One illustration of its wide acceptability is that Black and Scholes [BS73] used this same model in their Nobel prize winning work on pricing options on stocks.

"Worst-case" Investing: The fragility of average-case models in the face of rare but dramatic deviations led Cover [Cov91] to take a worst-case approach to investing in stocks. The performance of an online investment algorithm for *arbitrary* sequences of stock price returns is measured with respect to the best CRP (constant rebalanced portfolio, see [Cov91]) in hindsight. A universal portfolio selection algorithm is one that obtains sublinear (in the number of trading periods T) *regret*, which is the difference in the logarithms of the final wealths obtained by the two.

Cover [Cov91] gave the first universal portfolio selection algorithm with regret bounded by $O(\log T)$. There has been much follow-up work after Cover's seminal work, such as [HSSW96, MF93, KV03, BK99, HAK07], which focused on either obtaining alternate universal algorithms or improving the efficiency of Cover's algorithm. However, the best regret bound is still $O(\log T)$.

This dependence of the regret on the number of trading periods is not entirely satisfactory for two main reasons. First, *a priori* it is not clear why the online algorithm should have high regret (growing with the number of iterations) in an unchanging environment. As an extreme example, consider a setting with two stocks where one has an "upward drift" of 1% daily, whereas the second stock remains at the same price. One would expect to "figure out" this pattern quickly and focus on the first stock, thus attaining a constant fraction of the wealth of the best CRP in the long run, i.e. constant regret, unlike the worst-case bound of $O(\log T)$.

The second problem arises from trading frequency. Suppose we need to invest over a fixed period of time, say a year. Trading more frequently potentially leads to higher wealth gain, by capitalizing on short term stock movements. However, increasing trading frequency increases T, and thus one may expect more regret. The problem is actually even worse: since we measure regret as a difference of logarithms of the final wealths, a regret bound of $O(\log T)$ implies a polynomial in T factor ratio between the final wealths. In reality, however, experiments [AHKS06] show that some known online algorithms actually *improve* with increasing trading frequency.

Bridging Worst-case and Average-case Investing: Both these issues are resolved if one can show that the regret of a "good" online algorithm depends on total *variation* in the sequence of stock returns, rather than purely on the number of iterations. If the stock return sequence has low variation, we expect our algorithm to be able to perform better. If we trade more frequently, then the per iteration variation should go down correspondingly, so the total variation stays the same.

We analyze a portfolio selection algorithm and prove that its regret is bounded by $O(\log Q)$, where Q (formally defined in Section 1.2) is the sum of squared deviations of the returns from their mean. Since $Q \leq T$ (after appropriate normalization), we improve over previous regret bounds and retain the worst-case robustness. Furthermore, in an average-case model such as GBM, the variation can be tied very nicely to the volatility parameter, which explains the experimental observation the regret doesn't increase with increasing trading frequency. Our algorithm is efficient, and its implementation requires constant time per iteration (independent of the number of game iterations).

1.1 New Techniques and Comparison to Related Work

Cesa-Bianchi, Mansour and Stoltz [CBMS07] initiated work on relating worst case regret to the variation in the data for the related learning problem of prediction from expert advice, and conjectured that the optimal regret bounds should depend on the observed variation of the cost sequence. Recently, this conjectured was proved and regret bounds of $\tilde{O}(\sqrt{Q})$ were obtained in the full information and bandit linear optimization settings [HK10, HK11], where Q is the variation in the cost sequence. In this paper we give an exponential improvement in regret, viz. $O(\log Q)$, for the case of online *exp-concave* optimization, which includes portfolio selection as a special case.

Another approach to connecting worst-case to average-case investing was taken by Jamshidian [Jam92] and Cross and Barron [CB03]. They considered a model of "continuous trading", where there are T "trading intervals", and in each the online investor chooses a *fixed* portfolio which is rebalanced k times with $k \to \infty$. They prove familiar regret bounds of $O(\log T)$ (independent of k) in this model w.r.t. the best fixed portfolio which is rebalanced $T \times k$ times. In this model our algorithm attains the tighter regret bounds of $O(\log Q)$, although our algorithm has more flexibility. Furthermore their algorithms, being extensions of Cover's algorithm, may require exponential time in general[2].

Our bounds of $O(\log Q)$ regret require completely different techniques compared to the $\tilde{O}(\sqrt{Q})$ regret bounds of [HK10, HK11]. These previous bounds are based on first-order gradient descent methods which are too weak to obtain $O(\log Q)$ regret. Instead we have to use the second-order Newton step ideas based on [HAK07] (in particular, the Hessian of the cost functions).

[2]Cross and Barron give an efficient implementation for some interesting special cases, under assumptions on the variation in returns and bounds on the magnitude of the returns, and assuming $k \to \infty$. A truly efficient implementation of their algorithm can probably be obtained using the techniques of Kalai and Vempala.

The second-order techniques of [HAK07] are, however, not sensitive enough to obtain $O(\log Q)$ bounds. This is because progress was measured in terms of the distance between successive portfolios in the usual Euclidean norm, which is insensitive to variation in the cost sequence. In this paper, we introduce a different analysis technique, based on analyzing the distance between successive predictions using norms that keep changing from iteration to iteration and are actually sensitive to the variation.

A key technical step in the analysis is a lemma (Lemma 5) which bounds the sum of differences of successive Cesaro means of a sequence of vectors by the logarithm of its variation. This lemma, which may be useful in other contexts when variation bounds on the regret are desired, is proved using the Kahn-Karush-Tucker conditions, and also improves the regret bounds in previous papers.

1.2 The model and statement of results

Portfolio management. In the universal portfolio management model [Cov91], an online investor iteratively distributes her wealth over n assets before observing the change in asset price. In each iteration $t = 1, 2, \dots$ the investor commits to an n-dimensional distribution of her wealth, $x_t \in \Delta_n = \{\sum_i x_i = 1 \ , \ x \geq 0\}$. She then observes a *price relatives vector* $r_t \in \mathbb{R}^n_+$, where $r_t(i)$ is the ratio between the closing price of the i^{th} asset on trading period t and the opening price. In the t^{th} trading period, the wealth of the investor changes by a factor of $(r_t \cdot x_t)$. The overall change in wealth is thus $\prod_t (r_t \cdot x_t)$. Since in a typical market wealth grows at an exponential rate, we measure performance by the exponential growth rate, which is $\log \prod_t (r_t \cdot x_t) = \sum_t \log(r_t \cdot x_t)$. A *constant rebalanced portfolio* (CRP) is an investment strategy which rebalances the wealth in every iteration to keep a fixed distribution. Thus, for a CRP $x \in \Delta_n$, the change in wealth is $\prod_t (r_t \cdot x)$.

The *regret* of the investor is defined to be the difference between the exponential growth rate of her investment strategy and that of the best CRP strategy in hindsight, i.e.

$$\text{Regret} := \max_{x^* \in \Delta_n} \sum_t \log(r_t \cdot x^*) - \sum_t \log(r_t \cdot x_t)$$

Note that the regret doesn't change if we scale all the returns in any particular period by the same amount. So we assume w.l.o.g. that in all periods t, $\max_i r_t(i) = 1$. We assume that there is known parameter $r > 0$, such that for all periods t, $\min_{t,i} r_t(i) \geq r$. We call r the *market variability* parameter. This is the only restriction we put on the stock price returns; they could be chosen adversarially as long as they respect the market variability bound.

Online convex optimization. In the online convex optimization problem [Zin03], which generalizes universal portfolio management, the decision space is a closed, bounded, convex set $K \in \mathbb{R}^n$, and we are sequentially given a series of convex cost[3] functions $f_t : K \to \mathbb{R}$ for $t = 1, 2, \dots$. The algorithm iteratively produces a point $x_t \in K$ in every round t, without knowledge of f_t (but using the past sequence of cost

[3] Note the difference from the portfolio selection problem: here we have convex *cost* functions, rather than concave payoff functions. The portfolio selection problem is obtained by using $-\log$ as the cost function.

functions), and incurs the cost $f_t(x_t)$. The regret at time T is defined to be

$$\text{Regret} \ := \ \sum_{t=1}^{T} f_t(x_t) - \min_{x \in K} \sum_{t=1}^{T} f_t(x).$$

In this paper, we restrict our attention to convex cost functions which can be written as $f_t(x) = g(v_t \cdot x)$ for some univariate convex function g and a parameter vector $v_t \in \mathbb{R}^n$ (for example, in the portfolio management problem, $K = \Delta_n$, $f_t(x) = -\log(r_t \cdot x)$, $g = -\log$, and $v_t = r_t$).

Thus, the cost functions are *parametrized* by the vectors $v_1, v_2, \ldots, v_T$. Our bounds will be expressed as a function of the *quadratic variability* of the parameter vectors $v_1, v_2, \ldots, v_T$, defined as

$$Q(v_1, ..., v_T) \ := \ \min_{\mu} \sum_{t=1}^{T} \|v_t - \mu\|^2.$$

This expression is minimized at $\mu = \frac{1}{T}\sum_{t=1}^{T} v_t$, and thus the quadratic variation is just $T-1$ times the sample variance of the sequence of vectors $\{v_1, ..., v_t\}$. Note however that the sequence can be generated adversarially rather than by some stochastic process. We shall refer to this as simply Q if the vectors are clear from the context.
Main theorem. In the setup of the online convex optimization problem above, we have the following algorithmic result:

Theorem 1. *Let the cost functions be of the form $f_t(x) = g(v_t \cdot x)$. Assume that there are parameters $R, D, a, b > 0$ such that the following conditions hold:*

1. *for all t, $\|v_t\| \le R$,*
2. *for all $x \in K$, we have $\|x\| \le D$,*
3. *for all $x \in K$, and for all t, either $g'(v_t \cdot x) \in [0, a]$ or $g'(v_t \cdot x) \in [-a, 0]$, and*
4. *for all $x \in K$, and for all t, $g''(v_t \cdot x) \ge b$.*

Then there is an algorithm that guarantees the following regret bound:

$$\textit{Regret} = O((a^2 n/b)\log(1 + bQ + bR^2) + aRD\log(1 + Q/R^2) + D^2).$$

Now we apply Theorem 1 to the portfolio selection problem. First, we estimate the relevant parameters. We have $\|r_t\| \le \sqrt{n}$ since all $r_t(i) \le 1$, thus $R = \sqrt{n}$. For any $x \in \Delta_n$, $\|x\| \le 1$, so $D = 1$. $g'(v_t \cdot x) = -\frac{1}{(v_t \cdot x)}$, and thus $g'(v_t \cdot x) \in [-\frac{1}{r}, 0]$, so $a = \frac{1}{r}$. Finally, $g''(v_t \cdot x) = \frac{1}{(v_t \cdot x)^2} \ge 1$, so $b = 1$. Applying Theorem 1 we get the following corollary:

Corollary 2. *For the portfolio selection problem over n assets, there is an algorithm that attains the following regret bound:*

$$\textit{Regret} \quad = \quad O\left(\frac{n}{r^2}\log(Q + n)\right).$$

2 Bounding the Regret by the Observed Variation in Returns

2.1 Preliminaries

All matrices are assumed be real symmetric matrices in $\mathbb{R}^{n\times n}$, where n is the number of stocks. We use the notation $A \succeq B$ to say that $A - B$ is positive semidefinite. We require the notion of a norm of a vector x induced by a positive definite matrix M, defined as $\|x\|_M = \sqrt{x^\top M x}$. The following simple generalization of the Cauchy-Schwartz inequality is used in the analysis:

$$\forall x, y \in \mathbb{R}^n : \quad x \cdot y \ \le\ \|x\|_M \|y\|_{M^{-1}}.$$

We denote by $|A|$ the determinant of a matrix A, and by $A \bullet B = \mathbf{Tr}(AB) = \sum_{ij} A_{ij}B_{ij}$. As we are concerned with logarithmic regret bounds, potential functions which behave like harmonic series come into play. A generalization of harmonic series to high dimensions is the *vector-harmonic series*, which is a series of quadratic forms that can be expressed as (here $A \succ 0$ is a positive definite matrix, and $v_1, v_2, \ldots$ are vectors in $\mathbb{R}^n$):

$$v_1^\top (A + v_1 v_1^\top)^{-1} v_1,\ v_2^\top (A + v_1 v_1^\top + v_2 v_2^\top)^{-1} v_2, \ldots, v_t^\top (A + \textstyle\sum_{\tau=1}^t v_\tau v_\tau^\top)^{-1} v_t, \ldots$$

The following lemma is from [HAK07] (and proven in appendix A for completeness):

Lemma 3. *For a vector harmonic series given by an initial matrix A and vectors $v_1, v_2, \ldots, v_T$, we have*

$$\sum_{t=1}^T v_t^\top (A + \textstyle\sum_{\tau=1}^t v_\tau v_\tau^\top)^{-1} v_t \le \log \left[\dfrac{|A + \sum_{\tau=1}^T v_\tau v_\tau^\top|}{|A|} \right].$$

The reader can note that in one dimension, if all vectors $v_t = 1$ and $A = 1$, then the series above reduces exactly to the regular harmonic series whose sum is bounded, of course, by $\log(T + 1)$.

Henceforth we will denote by $\sum_t$ the summation $\sum_{t=1}^T$.

2.2 Algorithm and analysis

We analyze the following algorithm and prove that it attains logarithmic regret with respect to the observed variation (rather than number of iterations). The algorithm follows the generic algorithmic scheme of "Follow-The-Regularized-Leader" (FTRL) with squared Euclidean regularization.

Algorithm Exp-Concave-FTL. In iteration t, use the point x_t defined as:

$$x_t \triangleq \arg\min_{x \in \Delta_n} \left(\sum_{\tau=1}^{t-1} f_\tau(x) + \frac{1}{2}\|x\|^2 \right) \tag{1}$$

Note the mathematical program which the algorithm solves is convex, and can be solved in time polynomial in the dimension and number of iterations. The running time, however, for solving this convex program can be quite high. In section 4, for the specific problem of portfolio selection, where $f_t(x) = -\log(r_t \cdot x)$, we give a faster implementation whose per iteration running time is independent of the number of iterations.

We now proceed to prove the Theorem 1.

Proof. **[Theorem 1]** First, we note that the algorithm is running a "Follow-the-leader" procedure on the cost functions $f_0, f_1, f_2, \ldots$ where $f_0(x) = \frac{1}{2}\|x\|^2$ is a fictitious period 0 cost function. In other words, in each iteration, it chooses the point that would have minimized the total cost under all the observed functions so far (and, additionally, a fictitious initial cost function f_0). This point is referred to as the leader in that round.

The first step in analyzing such an algorithm is to use a stability lemma from [KV05], which bounds the regret of any Follow-the-leader algorithm by the difference in costs (under f_t) of the current prediction x_t and the next one x_{t+1}, plus an additional error term which comes from the regularization. Thus, we have (recall the notation $\sum_t \equiv \sum_{t=1}^T$)

$$\begin{aligned} \text{Regret} &\leq \sum_t f_t(x_t) - f_t(x_{t+1}) + \frac{1}{2}(\|x^*\|^2 - \|x_0\|^2) \\ &\leq \sum_t \nabla f_t(x_t) \cdot (x_t - x_{t+1}) + \frac{1}{2}D^2 \\ &= \sum_t g'(v_t \cdot x_t)[v_t \cdot (x_t - x_{t+1})] + \frac{1}{2}D^2 \qquad (2) \end{aligned}$$

The second inequality is because f_t is convex. The last equality follows because $\nabla f_t(x_t) = g'(x_t \cdot v_t)v_t$. Now, we need a handle on $x_t - x_{t+1}$. For this, define $F_t = \sum_{\tau=0}^{t-1} f_\tau$, and note that x_t minimizes F_t over K. Consider the difference in the gradients of F_{t+1} evaluated at x_{t+1} and x_t:

$$\begin{aligned} \nabla F_{t+1}(x_{t+1}) - \nabla F_{t+1}(x_t) &= \sum_{\tau=0}^{t} \nabla f_\tau(x_{t+1}) - \nabla f_\tau(x_t) \\ &= \sum_{\tau=1}^{t} [g'(v_\tau \cdot x_{t+1}) - g'(v_\tau \cdot x_t)]v_\tau + (x_{t+1} - x_t) \\ &= \sum_{\tau=1}^{t} [\nabla g'(v_\tau \cdot \zeta_\tau^t) \cdot (x_{t+1} - x_t)]v_\tau + (x_{t+1} - x_t) \qquad (3) \\ &= \sum_{\tau=1}^{t} g''(v_\tau \cdot \zeta_\tau^t)v_\tau v_\tau^\top (x_{t+1} - x_t) + (x_{t+1} - x_t). \qquad (4) \end{aligned}$$

Equation 3 follows by applying the Taylor expansion of the (multi-variate) function $g'(v_\tau \cdot x)$ at point x_t, for some point ζ_τ^t on the line segment joining x_t and x_{t+1}. The equation (4) follows from the observation that $\nabla g'(v_\tau \cdot x) = g''(v_\tau \cdot x)v_\tau$.

Define $A_t = \sum_{\tau=1}^{t} g''(v_\tau \cdot \zeta_\tau^t) v_\tau v_\tau^\top + I$, where I is the identity matrix, and $\Delta x_t = x_{t+1} - x_t$. Then equation (4) can be re-written as:

$$\nabla F_{t+1}(x_{t+1}) - \nabla F_t(x_t) - g'(v_t \cdot x_t) v_t \; = \; A_t \Delta x_t. \tag{5}$$

Now, since x_t minimizes the convex function F_t over the convex set K, a standard inequality of convex optimization (see [BV04]) states that for any point $y \in K$, we have $\nabla F_t(x_t) \cdot (y - x_t) \; \geq \; 0$. Thus, for $y = x_{t+1}$, we get that $\nabla F_t(x_t) \cdot (x_{t+1} - x_t) \geq 0$. Similarly, we get that $\nabla F_{t+1}(x_{t+1}) \cdot (x_t - x_{t+1}) \geq 0$. Putting these two inequalities together, we get that

$$(\nabla F_{t+1}(x_{t+1}) - \nabla F_t(x_t)) \cdot \Delta x_t \; \leq \; 0. \tag{6}$$

Thus, using the expression for $A_t \Delta x_t$ from (5) we have

$$\begin{aligned}
\|\Delta x_t\|_{A_t}^2 \; &= \; A_t \Delta x_t \cdot \Delta x_t \\
&= \; (\nabla F_{t+1}(x_{t+1}) - \nabla F_t(x_t) - g'(v_t \cdot x_t) v_t) \cdot \Delta x_t \\
&\leq \; g'(v_t \cdot x_t)[v_t \cdot (x_t - x_{t+1})] \qquad \text{(from (6))}
\end{aligned} \tag{7}$$

Assume that $g'(v_t \cdot x) \in [-a, 0]$ for all $x \in K$ and all t. The other case is handled similarly. Inequality (7) implies that $g'(v_t \cdot x_t)$ and $v_t \cdot (x_t - x_{t+1})$ have the same sign. Thus, we can upper bound

$$g'(v_t \cdot x_t)[v_t \cdot (x_t - x_{t+1})] \; \leq \; a(v_t \cdot \Delta x_t). \tag{8}$$

Define $\tilde{v}_t = v_t - \mu_t$, $\mu_t = \frac{1}{t+1} \sum_{\tau=1}^{t} v_\tau$. Then, we have

$$\sum_t v_t \cdot \Delta x_t \; = \; \sum_t \tilde{v}_t \cdot \Delta x_t + \sum_{t=2}^{T} x_t(\mu_{t-1} - \mu_t) - x_1 \mu_1 + x_{T+1} \mu_T, \tag{9}$$

Now, define $\rho = \rho(v_1, \ldots, v_T) = \sum_{t=1}^{T-1} \|\mu_{t+1} - \mu_t\|$. Then we bound

$$\begin{aligned}
&\sum_{t=2}^{T} x_t(\mu_{t-1} - \mu_t) - x_1 \mu_1 + x_{T+1} \mu_T \\
&\qquad \leq \; \sum_{t=2}^{T} \|x_t\| \|\mu_{t-1} - \mu_t\| + \|x_1\| \|\mu_1\| + \|x_{T+1}\| \|\mu_T\| \\
&\qquad \leq \; D\rho + 2DR.
\end{aligned} \tag{10}$$

We will bound ρ momentarily. For now, we turn to bounding the first term of (9) using the Cauchy-Schwartz generalization as follows:

$$\tilde{v}_t \cdot \Delta x_t \; \leq \|\tilde{v}_t\|_{A_t^{-1}} \|\Delta x_t\|_{A_t}. \tag{11}$$

By the usual Cauchy-Schwartz inequality,

$$\begin{aligned}
\sum_t \|\tilde{v}_t\|_{A_t^{-1}} \|\Delta x_t\|_{A_t} \; &\leq \; \sqrt{\sum_t \|\tilde{v}_t\|_{A_t^{-1}}^2} \cdot \sqrt{\sum_t \|\Delta x_t\|_{A_t}^2} \\
&\leq \; \sqrt{\sum_t \|\tilde{v}_t\|_{A_t^{-1}}^2} \cdot \sqrt{\sum_t a(v_t \cdot \Delta x_t)}.
\end{aligned}$$

from (7) and (8). We conclude, using (9), (10) and (11), that

$$\sum_t a(v_t \cdot \Delta x_t) \;\le\; a\sqrt{\sum_t \|\tilde{v}_t\|^2_{A_t^{-1}}} \cdot \sqrt{\sum_t a(v_t \cdot \Delta x_t)} + aD\rho + 2aDR.$$

This implies (using the AM-GM inequality applied to the first term on the RHS) that

$$\sum_t a(v_t \cdot \Delta x_t) \;\le\; a^2 \sum_t \|\tilde{v}_t\|^2_{A_t^{-1}} + 2aD\rho + 4aDR.$$

Plugging this into the regret bound (2) we obtain, via (8),

$$\text{Regret} \;\le\; a^2 \sum_t \|\tilde{v}_t\|^2_{A_t^{-1}} + 2aD\rho + 4aDR + \frac{1}{2}D^2.$$

The proof is completed by the following two lemmas (Lemmas 4 and 5) which bound the RHS. The first term is a vector harmonic series, and the second term can be bounded by a (regular) harmonic series. □

Lemma 4. $\sum_t \|\tilde{v}_t\|^2_{A_t^{-1}} \;\le\; \frac{5n}{b} \log\left[1 + bQ + bR^2\right]$.

Proof. We have $A_t = \sum_{\tau=1}^t g''(v_\tau \cdot \zeta_\tau^t) v_\tau v_\tau^\top + I$. Since $g''(v_t \cdot \zeta_\tau^t) \ge b$, we have $A_t \succeq I + b\sum_{\tau=1}^t v_\tau v_\tau^\top$. Using the fact that $\tilde{v}_t = v_t - \mu_t$ and $\mu_t = \frac{1}{t+1}\sum_{\tau \le t} v_\tau$ we get that

$$\sum_{\tau=1}^t \tilde{v}_\tau \tilde{v}_\tau^\top \;=\; \sum_{s=1}^t (v_s - \frac{1}{s+1}\sum_{\tau \le s} v_\tau)(v_s - \frac{1}{s+1}\sum_{\tau\le s} v_\tau)^\top$$
$$= \sum_{s=1}^t \left(1 + \sum_{\tau=s}^t \frac{1}{(\tau+1)^2} - \frac{2}{s+1}\right) v_s v_s^\top + \sum_{s=1}^t \sum_{r<s} \left(-\frac{1}{s+1} + \sum_{\tau=s}^t \frac{1}{(\tau+1)^2}\right)[v_r v_s^\top + v_s v_r^\top].$$

Since

$$\frac{1}{s+1} - \frac{1}{t+2} \;=\; \int_{s+1}^{t+2} \frac{1}{x^2}dx \;\le\; \sum_{\tau=s}^t \frac{1}{(\tau+1)^2} \;\le\; \int_s^{t+1} \frac{1}{x^2}dx \;=\; \frac{1}{s} - \frac{1}{t+1}, \tag{12}$$

we get that

$$\left|-\frac{1}{s+1} + \sum_{\tau=s}^t \frac{1}{(\tau+1)^2}\right| \;\le\; \frac{1}{s^2} + \frac{1}{t}. \tag{13}$$

Since $(v_r \pm v_s)(v_r \pm v_s)^\top \succeq 0$, we have $(v_r v_r^\top + v_s v_s^\top) \succeq \pm(v_r v_s^\top + v_s v_r^\top)$, and so

$$\left(-\frac{1}{s+1} + \sum_{\tau=s}^t \frac{1}{(\tau+1)^2}\right)[v_r v_s^\top + v_s v_r^\top] \;\preceq\; \left|-\frac{1}{s+1} + \sum_{\tau=s}^t \frac{1}{(\tau+1)^2}\right| [v_r v_r^\top + v_s v_s^\top]$$
$$\preceq\; \left(\frac{1}{s^2} + \frac{1}{t}\right)[v_r v_r^\top + v_s v_s^\top],$$

by (13). Also by (12), we have $(1 + \sum_{\tau=s}^{t} \frac{1}{(\tau+1)^2} - \frac{2}{s+1}) \leq 1 + \frac{1}{s} - \frac{1}{t+1} - \frac{2}{s+1} \leq 1$, so we have

$$
\begin{aligned}
\sum_{\tau=1}^{t} \tilde{v}_\tau \tilde{v}_\tau^\top &\preceq \sum_{s=1}^{t} v_s v_s^\top + \sum_{s=1}^{t} \sum_{r<s} \left(\frac{1}{s^2} + \frac{1}{t} \right) [v_r v_r^\top + v_s v_s^\top] \\
&\preceq 3 \sum_{s=1}^{t} v_s v_s^\top + \sum_{s=1}^{t} \left(\frac{1}{s^2} + \frac{1}{t} \right) \sum_{r<s} v_r v_r^\top \\
&\preceq 3 \sum_{s=1}^{t} v_s v_s^\top + \sum_{s=1}^{t} v_s v_s^\top \sum_{r=s+1}^{t} \left(\frac{1}{r^2} + \frac{1}{t} \right) \\
&\preceq 3 \sum_{s=1}^{t} v_s v_s^\top + \sum_{s=1}^{t} \left(1 + \frac{1}{s} \right) v_s v_s^\top \\
&\preceq 5 \sum_{s=1}^{t} v_s v_s^\top .
\end{aligned}
$$

Let $\tilde{A}_t = I + b \sum_\tau \tilde{v}_\tau \tilde{v}_\tau^\top$. Note that the inequality above shows that $\tilde{A}_t \preceq 5A_t$. Thus, using Lemma 3, we get

$$
\sum_t \|\tilde{v}_t\|_{A_t^{-1}}^2 \;=\; \sum_t \tilde{v}_t A_t^{-1} \tilde{v}_t \;\leq\; \frac{5}{b} \sum_t [\sqrt{b}\tilde{v}_t]^\top \tilde{A}_t^{-1} [\sqrt{b}\tilde{v}_t] \;\leq\; \frac{5}{b} \log \left[\frac{|\tilde{A}_T|}{|\tilde{A}_0|} \right]. \quad (14)
$$

To bound the latter quantity note that $|\tilde{A}_0| = |I| = 1$, and that

$$
|\tilde{A}_T| \;=\; |I + b \sum_t \tilde{v}_t \tilde{v}_t^\top| \;\leq\; (1 + b \sum_t \|\tilde{v}_t\|_2^2)^n \;=\; (1 + b\tilde{Q})^n
$$

where $\tilde{Q} = \sum_t \|\tilde{v}_t\|^2 = \sum_t \|v_t - \mu_t\|^2$. Lemma 6, shows that $\tilde{Q} \;\leq\; Q + R^2$. This implies that $|\tilde{A}_T| \leq (1 + bQ + bR^2)^n$ and the proof is completed by substituting this bound into (14). □

Lemma 5. $\rho(v_1, \ldots, v_T) \;\leq\; 4R[\log(1 + Q/R^2) + 2]$.

Proof. Define, for $\tau = 0, 1, 2, \ldots, T$, the vector $u_\tau = v_\tau - \mu$, where $\mu = \frac{1}{T} \sum_{t=1}^{T} v_t$. Let $v_0 = 0$ and $u_0 = -\mu$. We have

$$
\sum_{t=1}^{T} \|u_t\|^2 \;=\; \sum_{t=1}^{T} \|v_t - \mu\|^2 \;=\; Q.
$$

We have

$$\begin{aligned}
\|\mu_{t+1}-\mu_t\| &= \left\| \frac{1}{t+2}\sum_{\tau=0}^{t+1} v_\tau - \frac{1}{t+1}\sum_{\tau=0}^{t} v_\tau \right\| \\
&= \left\| \frac{1}{t+2}\sum_{\tau=0}^{t+1} u_\tau - \frac{1}{t+1}\sum_{\tau=0}^{t} u_\tau \right\| \\
&\le \frac{1}{(t+1)^2}\sum_{\tau=0}^{t} \|u_\tau\| + \frac{1}{t+1}\|u_{t+1}\|.
\end{aligned}$$

Summing up over all iterations,

$$\begin{aligned}
\rho \;=\; \sum_{t=1}^{T-1}\|\mu_{t+1}-\mu_t\| \;&\le\; \sum_{t=1}^{T-1}\left(\frac{1}{(t+1)^2}\sum_{\tau=0}^{t}\|u_\tau\| + \frac{1}{t+1}\|u_{t+1}\|\right) \\
&\le\; \|u_0\|\cdot\sum_{t=1}^{T-1}\frac{1}{(t+1)^2} + \sum_{t=1}^{T}\|u_t\|\cdot\left(\frac{1}{t}+\sum_{\tau=t}^{T-1}\frac{1}{(\tau+1)^2}\right) \\
&\le\; \|u_0\| + \sum_{t=1}^{T}\frac{2}{t}\|u_t\| \\
&\le\; 4R[\log(1+Q/R^2)+2].
\end{aligned}$$

The second inequality follows because $\sum_{\tau=t}^{\infty}\frac{1}{(\tau+1)^2} \le \int_{x=t}^{\infty}\frac{1}{x^2}dx = \frac{1}{t}$. The last inequality uses the following facts:

1. Since $v_0 = 0$, $u_0 = -\mu$, and hence $\|u_0\| = \|\mu\| \le R$ since for all t, $\|v_t\| \le R$.
2. Note $\|u_t\| = \|v_t - \mu\| \le \|v_t\| + \|\mu\| \le 2R$. Applying Lemma 7 below with $x_t = \|u_t\|/2R$ for $t = 1,2,\ldots,T$, and using the fact that $\sum_{t=1}^{T} x_t^2 = \sum_{t=1}^{T}\|u_t\|^2/4R^2 \le Q/R^2$, we get that

$$\sum_{t=1}^{T}\frac{2}{t}\|u_t\| \;\le\; 4R(\log(1+Q/R^2)+1).$$

□

Lemma 6. $\tilde{Q} \;\le\; Q + R^2$.

Proof. Consider the Be-The-Leader (BTL) algorithm played on the sequence of cost functions $c_t(x) = \|v_t - x\|^2$, for $t = 0,1,2,\ldots,T$, with $v_0 = 0$, when the convex domain is the ball of radius R. The BTL algorithm is as follows. On round t, this algorithm chooses the point that minimizes $\sum_{\tau=0}^{t} c_t(x)$ over the domain. It is easy to see that this point is exactly $\frac{1}{t+1}\sum_{\tau=0}^{t} v_t = \mu_t$. Thus, the cost of the algorithm is $\sum_{\tau=0}^{t}\|v_t - \mu_t\|^2 = \tilde{Q}$, since $\mu_0 = v_0 = 0$, and the first period cost is thus 0. The best fixed point in hindsight is μ_T. Thus, the cost of the best fixed point in hindsight, μ_T, is $\sum_{\tau=0}^{t}\|v_t - \mu_T\|^2 = Q + \|\mu_T\|^2$. Kalai and Vempala [KV05] prove that the BTL algorithm incurs 0 regret, i.e. $\tilde{Q} \le Q + \|\mu_T\|^2 \le Q + R^2$. □

Lemma 7. *Suppose that* $0 \le x_t \le 1$ *and* $\sum_t x_t^2 \le Q$. *Then*

$$\sum_{t=1}^{T} \frac{x_t}{t} \;\le\; \log(1+Q)+1.$$

Proof. By the Lemma 8 below, the values of x_t that maximize $\sum_{t=1}^{T} x_t/t$ must have the following structure: there is a k such that for all $t \le k$, we have $x_t = 1$, and for any index $t > k$, we have $x_{k+1}/x_t \ge (1/k)/(1/t)$, which implies that $x_t \le k/t$. We first note that $k \le Q$, since $Q \ge \sum_{t=1}^{k} x_t^2 = k$. Now, we can bound the value as follows:

$$\begin{aligned}
\sum_{t=1}^{T} \frac{x_t}{t} \;&\le\; \sum_{t=1}^{k} \frac{1}{t} + \sum_{t=k+1}^{T} \frac{k}{t^2} \\
&\le\; \log(k+1) + k \cdot \frac{1}{k} \\
&=\; \log(1+Q)+1.
\end{aligned}$$

□

Lemma 8. *Let* $a_1 \ge a_2 \ge \ldots a_n > 0$. *Then the optimal solution of*

$$\max \left\{ \sum_i a_i x_i : \; 0 \le x_i \le 1 \textit{ and } \sum_i x_i^2 \le Q \right\}$$

has the following properties: $x_1 \ge x_2 \ge \ldots x_n$, *and for any pair of indices* i, j, *with* $i < j$, *either* $x_i = 1$, $x_i = 0$ *or* $x_i/x_j \ge a_i/a_j$.

Proof. The fact that in the optimal solution $x_1 \ge x_2 \ge \ldots x_n$ is obvious, since otherwise we could permute the x_i's to be in decreasing order and increase the value.

The second fact follows by the Karush-Kuhn-Tucker (KKT) optimality conditions, which imply the existence of constants $\mu, \lambda_1, ..., \lambda_n, \rho_1, ..., \rho_n$ for which the optimal vector x satisfies (here e_i is the i^{th} standard basis vector, and $a = (a_1, a_2, \ldots, a_n)$):

$$-a + 2\mu x + \sum_i (\lambda_i + \rho_i) e_i = 0$$

By KKT theory, complementary slackness implies that the constants λ_i, ρ_i are equal to zero for all indices of the solution which satisfy $x_i \notin \{0, 1\}$. For these coordinates, the KKT equation is

$$a_i - 2\mu x_i = 0,$$

which implies the lemma. □

3 Implications in the Geometric Brownian Motion Model

We begin with a brief description of the model. The model assumes that stocks can be traded continuously, and that at any time, the fractional change in the stock price within an infinitesimal time interval is normally distributed, with mean and variance proportional to the length of the interval. The randomness is due to many infinitesimal trades that jar the price, much like particles in a physical medium are jarred about by other particles, leading to the classical Brownian motion.

Formally, the model is parameterized by two quantities, the *drift* μ, which is the long term trend of the stock prices, and *volatility* σ, which characterizes deviations from the long term trend. The parameter σ is typically specified as *annualized* volatility, i.e. the standard deviation of the stock's logarithmic returns in one year. Thus, a trading interval of $[0, 1]$ specifies 1 year. The model postulates that the stock price at time t, S_t, follows a geometric Brownian motion with drift μ and volatility σ:

$$dS_t \ = \ \mu S_t dt + \sigma S_t dW_t,$$

where W_t is a continuous-time stochastic process known as the Wiener process or simply Brownian motion. The Wiener process is characterized by three facts:

1. $W_0 = 0$,
2. W_t is almost surely continuous, and
3. for any two disjoint time intervals $[s_1, t_1]$ and $[s_2, t_2]$, the random variables $W_{t_1} - W_{s_1}$ and $W_{t_2} - W_{s_2}$ are independent zero mean Gaussian random variables with variance $t_1 - s_1$ and $t_2 - s_2$ respectively.

Using Itō's lemma (see, for example, [KS04]), it can be shown that the stock price at time t is given by

$$S_t \ = \ S_0 \exp((\mu - \sigma^2/2)t + \sigma W_t). \tag{15}$$

Now, we consider a situation where we have n stocks in the GBM model. Let $\mu = (\mu_1, \mu_2, \ldots, \mu_n)$ be the vector of drifts, and $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ be the vector of (annualized) volatilities. Suppose we trade for one year. We now study the effect of trading frequency on the quadratic variation of the stock price returns. For this, assume that the year-long trading interval is sub-divided into T equally sized intervals of length $1/T$, and we trade at the end of each such interval. Let $r_t = (r_t(1), r_t(2), \ldots, r_t(n))$ be the vector of stock returns in the t^{th} trading period. We assume that T is "large enough", which is taken to mean that it is larger than $\mu(i), \sigma(i), (\frac{\mu(i)}{\sigma(i)})^2$ for any i.

Then using the facts of the Wiener process stated above, we can prove the following lemma, which shows that the expected quadratic variation is the essentially the same regardless of trading frequency, while its variance decreases with trading frequency.

Lemma 9. *In the setup of trading n stocks in the GBM model over one year with T trading periods, where $T \gg \mu(i), \sigma(i)$, $\forall i$, there is a vector v such that*

$$\mathbf{E}\left[\sum_{t=1}^{T} \|r_t - v\|^2\right] \ \le \ \|\sigma\|^2(1 + O(\tfrac{1}{T}))$$

and

$$\mathit{VAR}\left[\sum_{t=1}^{T} \|r_t - v\|^2\right] \ \le \ \frac{3\|\sigma\|^4}{T}(1 + O(\tfrac{1}{T})),$$

regardless of how the stocks are correlated.

Proof. For every stock i, it follows from the GBM equation for the stock prices ($S_t = S_0 \exp((\mu - \sigma^2/2)t + \sigma W_t)$) that its return $r_t(i)$ in period t is given by

$$r_t(i) = \exp\left(\left(\mu(i) - \frac{\sigma(i)^2}{2}\right)\frac{1}{T} + \sigma(i)X_t(i)\right),$$

where $X_t(i) \sim \mathcal{N}(0, \frac{1}{T})$. Thus, for any given stock i, the returns $r_1(i), r_2(i), \ldots, r_T(i)$ are i.i.d. log-normal random variables, with parameters:

$$r_t(i) \sim \ln\mathcal{N}\left(\frac{\mu(i)}{T} - \frac{\sigma(i)^2}{2T}, \frac{\sigma(i)^2}{T}\right)$$

Recall that for a log-normal random variable $X \sim \ln\mathcal{N}(\mu, \sigma^2)$ the mean and variance are given by $\mathbf{E}[X] = e^{\mu+\sigma^2/2}$ and $\text{VAR}[X] = (e^{\sigma^2} - 1)e^{2\mu+\sigma^2}$.
Define the vector v as $v(i) = \mathbf{E}[r_t(i)] = e^{\mu(i)/T}$. Then, assuming that $T > \mu(i), \sigma(i)$, we have using the exponential approximation $e^x \leq 1 + x + x^2$ for $-1 \leq x \leq 1$:

$$\begin{aligned}
\mathbf{E}\left[(r_t(i) - v(i))^2\right] = \text{VAR}(r_t(i)) &= (e^{\sigma(i)^2/T} - 1)e^{2\mu(i)/T} \\
&\leq (\frac{\sigma(i)^2}{T} + O(\frac{1}{T^2})) \cdot e^{2\mu(i)/T} \\
&\leq \frac{\sigma(i)^2}{T}\left(1 + O\left(\tfrac{1}{T}\right)\right)
\end{aligned}$$

Where the last equality uses the Taylor approximation of the exponential and the fact that $\frac{2\mu(i)}{T} \ll 1$.

Summing up over all stocks i and all periods t, and using linearity of expectation, we get the first part of the Lemma:

$$\mathbf{E}\left[\textstyle\sum_{t=1}^T \|r_t - v\|^2\right] = \sum_{t=1}^T \sum_{i=1}^n \mathbf{E}[(r_t(i) - v(i))^2] \leq \|\sigma\|^2(1 + O(\tfrac{1}{T}))$$

As for the bound on the variance, we first bound the quantity $\mathbf{E}[(r_t(i) - v(i))^4]$, using the fact that if $X \sim \ln\mathcal{N}(\mu, \sigma^2)$, then $X^a \sim \ln\mathcal{N}(a\mu, a^2\sigma^2)$. We denote $\tilde{\mu} = \frac{\mu(i)}{T} - \frac{\sigma(i)^2}{2T}$, $\tilde{\sigma}^2 = \frac{\sigma(i)^2}{T}$.

$$\begin{aligned}
\mathbf{E}[(r_t(i) - v(i))^4] &= \mathbf{E}[r_t(i)^4 - 4r_t(i)^3 v(i) + 6r_t(i)^2 v(i)^2 - 4r_t(i)v(i)^3 + v(i)^4] \\
&= e^{4\tilde{\mu}+8\tilde{\sigma}^2} - 4e^{3\tilde{\mu}+\frac{9}{2}\tilde{\sigma}^2} e^{\tilde{\mu}+\tilde{\sigma}^2/2} + 6e^{2\tilde{\mu}+2\tilde{\sigma}^2} e^{2\tilde{\mu}+\tilde{\sigma}^2} \\
&- 4e^{\tilde{\mu}+\tilde{\sigma}^2/2} e^{3\tilde{\mu}+\frac{3}{2}\tilde{\sigma}^2} + e^{4\tilde{\mu}+2\tilde{\sigma}^2} \\
&= e^{4\tilde{\mu}+8\tilde{\sigma}^2} - 4e^{4\tilde{\mu}+5\tilde{\sigma}^2} + 6e^{4\tilde{\mu}+3\tilde{\sigma}^2} - 4e^{4\tilde{\mu}+2\tilde{\sigma}^2} + e^{4\tilde{\mu}+2\tilde{\sigma}^2} \\
&= e^{4\tilde{\mu}}(e^{8\tilde{\sigma}^2} - 4e^{5\tilde{\sigma}^2} + 6e^{3\tilde{\sigma}^2} - 3e^{2\tilde{\sigma}^2}) \\
&= (e^{\tilde{\sigma}^2} - 1)^2 e^{4\tilde{\mu}}(e^{6\tilde{\sigma}^2} + 2e^{5\tilde{\sigma}^2} + 3e^{4\tilde{\sigma}^2} - 3e^{2\tilde{\sigma}^2}) \\
&\leq (\frac{\sigma(i)^4}{T^2} + O(\frac{1}{T^3})) \cdot e^{4\tilde{\mu}}(e^{6\tilde{\sigma}^2} + 2e^{5\tilde{\sigma}^2} + 3e^{4\tilde{\sigma}^2} - 3e^{2\tilde{\sigma}^2}) \\
&= \frac{\sigma(i)^4}{T^2}(3 + O(\tfrac{1}{T})) = \frac{3\sigma(i)^4}{T^2}(1 + O(\tfrac{1}{T}))
\end{aligned}$$

Above the fifth equality follows from the polynomial factorization:

$$x^8 - 4x^5 + 6x^3 - 3x^2 = (x-1)^2(x^6 + 2x^5 + 3x^4 - 3x^2)$$

Thus for all stocks i and time periods t, we have:

$$\text{VAR}\left[(r_t(i) - v(i))^2\right] \;\leq\; \mathbf{E}[(r_t(i) - v(i))^4] \;\leq\; \frac{3\sigma(i)^4}{T^2}(1 + O(\tfrac{1}{T})) \tag{16}$$

Now, we use the following inequality which follows from the Cauchy-Schwarz inequality: for any random variables $X_1, X_2, \ldots, X_n$:

$$\text{VAR}[\textstyle\sum_{i=1}^n X_i] \;\leq\; (\sum_{i=1}^n \sqrt{\text{VAR}[X_i]})^2.$$

Hence, using inequality (16), we get:

$$\text{VAR}[\|r_t - v\|^2] \;\leq\; (\textstyle\sum_{i=1}^n \sqrt{\text{VAR}[(r_t(i) - v(i))^2]})^2 \;\leq\; \dfrac{3\|\sigma\|^4}{T^2}(1 + O(\tfrac{1}{T}))$$

Finally, using the fact that in the GBM model, the returns are independent between periods we get

$$\text{VAR}\left[\sum_{t=1}^T \|r_t - v\|^2\right] \;\leq\; \frac{3\|\sigma\|^4}{T}(1 + O(\tfrac{1}{T})).$$

□

Applying this bound in our algorithm, we obtain the following regret bound from Corollary 2.

Theorem 10. *In the setup of Lemma 9, for any $\delta > 0$, with probability at least $1-\delta$, we have*

$$\textit{Regret} \;=\; O(n \log((1 + \tfrac{3}{\sqrt{\delta T}})\|\sigma\|^2 + n)).$$

Proof. First, we show that the market variability parameter, r, is $\Omega(1)$ with high probability. Fix any stock i. The return $r_t(i)$ for stock i at time period t is a log-normal random variable: $r_t(i) = \exp(N_t(i))$ where $N_t(i) \sim \mathcal{N}((\mu(i) - \frac{\sigma(i)^2}{2})\frac{1}{T}, \frac{\sigma(i)^2}{T})$. Using standard tail bounds for normal variables (which is given in Lemma 14 in the Appendix for completeness), we have

$$\Pr\left[\left|N_t(i) - (\mu(i) - \frac{\sigma(i)^2}{2})\frac{1}{T}\right| > \frac{\sigma(i)}{\sqrt{T}}\sqrt{2\log(2nT/\delta)}\right] \;<\; e^{-\log(2nT/\delta)} \;=\; \frac{\delta}{2nT}.$$

Thus, assuming $T \gg \mu(i), \sigma(i)$, with probability at least $1 - \frac{\delta}{2nT}$, we have

$$|N_t(i)| = O(\tfrac{\sigma(i)}{\sqrt{T}}\sqrt{\log(nT/\delta)}).$$

Since $r_t(i) = \exp(N_t(i))$, we conclude that with probability at least $1 - \frac{\delta}{2nT}$, we have

$$|r_t(i) - 1| = O(\tfrac{\sigma(i)}{\sqrt{T}}\sqrt{\log(nT/\delta)}).$$

Applying a union bound over all stocks and all periods, we conclude that with probability at least $1-\frac{\delta}{2}$, the market variability parameter r is at least $1-O(\frac{\sigma}{\sqrt{T}}\sqrt{\log(nT/\delta)}) > 0.5$, if $T > \Omega(\sigma^2 \log(nT/\delta))$.

Now we bound the total variation. Applying Chebyshev's inequality to the random variable $\sum_{t=1}^{T} \|r_t - v\|^2$ and using Lemma 9, we have

$$\Pr\left[\sum_{t=1}^{T} \|r_t - v\|^2 > (1 + \tfrac{3}{\sqrt{\delta T}})\|\sigma\|^2\right] \;<\; \frac{\frac{3\|\sigma\|^4}{T}(1 + O(\frac{1}{T}))}{\frac{9\|\sigma\|^4}{\delta T}} \;<\; \frac{\delta}{2}.$$

The result now follows by a union bound applied to our regret bound of Corollary 2. □

Theorem 10 shows that one expects to achieve constant regret independent of the trading frequency, as long as the total trading period is fixed. This result is only useful if increasing trading frequency improves the performance of the best constant rebalanced portfolio. Indeed, this has been observed empirically (see e.g. [AHKS06] and Section 5).

To obtain a theoretical justification for increasing trading frequency, we consider an example where we have two stocks that follow independent Black-Scholes models with the same drifts, but different volatilities σ_1, σ_2. The same drift assumption is necessary because in the long run, the best CRP is the one that puts all its wealth on the stock with the greater drift. We normalize the drifts to be equal to 0, this doesn't change the performance in any qualitative manner.

Since the drift is 0, the expected return of either stock in any trading period is 1; and since the returns in each period are independent, the expected final change in wealth, which is the product of the returns, is also 1. Thus, in expectation, any CRP (indeed, any portfolio selection strategy) has overall return 1. We therefore turn to a different criterion for selecting a CRP. The risk of an investment strategy is measured by the variance of its payoff; thus, if different investment strategies have the same expected payoff, then the one to choose is the one with minimum variance. We therefore choose the CRP with the least variance.

Lemma 11. *In the setup where we trade two stocks with zero drift and volatilities σ_1, σ_2, the variance of the minimum variance CRP decreases as the trading frequency increases.*

Proof. To compute the minimum variance CRP, we first compute the variance of the CRP $(p, 1-p)$:

$$\begin{aligned}\mathrm{VAR}\left[\prod_t (pr_t(1) + (1-p)r_t(2))\right] &= \mathbf{E}\left[\prod_t (pr_t(1) + (1-p)r_t(2))^2\right] - 1 \\ &= \prod_t \mathbf{E}[(pr_t(1) + (1-p)r_t(2))^2] - 1.\end{aligned}$$

The second equality above follows from the independence of the randomness in each period. Thus, the minimum variance CRP is obtained by minimizing $\mathbf{E}[(pr_t(1) + (1 -$

$p)r_t(2))^2]$ over the range $p \in [0,1]$. We now note that in the Black-Scholes model, $r_t(1) = e^{X_1}$ where $X_1 \sim \mathcal{N}(-\frac{\sigma_1^2}{2T}, \frac{\sigma_1^2}{T})$, and $r_t(2) = e^{X_2}$ where $X_2 \sim \mathcal{N}(-\frac{\sigma_2^2}{2T}, \frac{\sigma_2^2}{T})$, and thus

$$\begin{aligned}
&\mathbf{E}[(pr_t(1) + (1-p)r_t(2))^2] \\
&= \mathbf{E}[p^2 e^{2X_1} + 2p(1-p)e^{X_1+X_2} + (1-p)^2 e^{2X_2}] \\
&= p^2 e^{\frac{\sigma_1^2}{T}} + 2p(1-p) + (1-p)^2 e^{\frac{\sigma_2^2}{T}}.
\end{aligned}$$

This is minimized at $p = \frac{\exp(\frac{\sigma_2^2}{T})-1}{\exp(\frac{\sigma_1^2}{T})+\exp(\frac{\sigma_2^2}{T})-2}$, and at this point, its value is $\frac{\exp(\frac{\sigma_1^2+\sigma_2^2}{T})-1}{\exp(\frac{\sigma_1^2}{T})+\exp(\frac{\sigma_2^2}{T})-2}$. Thus, the variance of the final payoff becomes

$$\left[\frac{\exp(\frac{\sigma_1^2+\sigma_2^2}{T}) - 1}{\exp(\frac{\sigma_1^2}{T}) + \exp(\frac{\sigma_2^2}{T}) - 2}\right]^T - 1.$$

This is a decreasing function of T. To prove this, by direct computation we have

$$\begin{aligned}
&\frac{d}{dT}\left[\left[\frac{\exp(\frac{\sigma_1^2+\sigma_2^2}{T}) - 1}{\exp(\frac{\sigma_1^2}{T}) + \exp(\frac{\sigma_2^2}{T}) - 2}\right]^T - 1\right] \\
&= \left[\frac{\exp(\frac{\sigma_1^2+\sigma_2^2}{T}) - 1}{\exp(\frac{\sigma_1^2}{T}) + \exp(\frac{\sigma_2^2}{T}) - 2}\right]^T \cdot \log\left[\frac{\exp(\frac{\sigma_1^2+\sigma_2^2}{T}) - 1}{\exp(\frac{\sigma_1^2}{T}) + \exp(\frac{\sigma_2^2}{T}) - 2}\right] \\
&\times \frac{-(\exp(\frac{\sigma_1^2}{T}) - 1)^2 \exp(\frac{\sigma_2^2}{T})\frac{\sigma_2^2}{T^2} - (\exp(\frac{\sigma_2^2}{T}) - 1)^2 \exp(\frac{\sigma_1^2}{T})\frac{\sigma_1^2}{T^2}}{(\exp(\frac{\sigma_1^2}{T}) + \exp(\frac{\sigma_2^2}{T}) - 2)^2} \\
&< 0.
\end{aligned}$$

The inequality above uses the fact that $\exp(\frac{\sigma_1^2+\sigma_2^2}{T}) - 1 > \exp(\frac{\sigma_1^2}{T}) + \exp(\frac{\sigma_2^2}{T}) - 2$ which can be verified by using the power series expansion of $\exp(x)$. $\square$

Thus, increasing the trading frequency decreases the variance of the minimum variance CRP, which implies that it gets less risky to trade more frequently; in other words, the more frequently we trade, the more likely the payoff will be close to the expected value. On the other hand, as we show in Theorem 10, the regret does not change even if we trade more often; thus, one expects to see improving performance of our algorithm as the trading frequency increases.

4 Faster Implementation

In this section we describe a more efficient algorithm compared to the one from the main body of the paper. The regret bound deteriorates slightly, though it is still logarithmic in the total quadratic variation. The algorithm is based on the online Newton method, introduced in [HAK07], and is described in the following figure. For simplicity, we focus on the portfolio management problem, although it is likely that similar ideas can work for general loss functions of the type we consider in this paper.

Algorithm 1 Faster quadratic-variation universal algorithm

for $t = 1$ to T **do**
  Use $x_t \triangleq \arg\min_{x \in \Delta_n} \left(\sum_{\tau=1}^{t-1} \tilde{f}_\tau(x) + \frac{1}{2}\|x\|^2 \right)$.
  Receive return vector r_t.
  Let $\tilde{f}_t(x) = -\log(r_t \cdot x_t) - \frac{r_t \cdot (x - x_t)}{(r_t \cdot x_t)} + \frac{r(r_t \cdot (x - x_t))^2}{8(r_t \cdot x_t)^2}$.
end for

The basic idea is to bound the cost functions by a paraboloid approximation in the FTRL algorithm. The paraboloid approximation only increases the regret, but since it is a simple quadratic function, the running time of the the FTRL algorithm is improved greatly. All we need to do is optimize the *sum* of quadratic cost functions, which has a compact representation (unlike the sum of log functions), over the n-dimensional simplex. This optimization can be carried out in time $O(n^{3.5})$ using interior point methods (assuming real number operations can be carried out in $O(1)$ time). Using observations made in [HAK07] it is possible to further speed up the algorithm and attain a running time proportional to $O(n^3)$.

We have the following regret bound for the algorithm:

Theorem 12. *For the portfolio selection problem, the regret of algorithm 1 is bounded by*

$$\textit{Regret} \ = \ O\left(\frac{n}{r^3}\log(Q+n)\right).$$

Proof. We first describe the paraboloid approximation to the cost functions that we use in the algorithm instead of actual cost functions. This approximation, based on a more general lemma from [HAK07], has the following property, for all x and y in the simplex and any return vector v with coordinates in $[r, 1]$:

$$-\log(v \cdot x) \ \geq \ -\log(v \cdot y) - \frac{v \cdot (x - y)}{(v \cdot y)} + \frac{r(v \cdot (x - y))^2}{8(v \cdot y)^2}.$$

Thus for any t, $\tilde{f}_t(x_t) = -\log(r_t \cdot x_t)$, and for any $x \in \Delta_n$, $\tilde{f}_t(x) \leq -\log(r_t \cdot x)$. Thus, if x^* is the best CRP in hindsight, we have the following bound on the regret of algorithm 1:

$$\begin{aligned} \text{Regret} \ &= \ \sum_t -\log(r_t \cdot x_t) - \sum_t -\log(r_t \cdot x^*) \\ &\leq \ \sum_t \tilde{f}_t(x_t) - \sum_t \tilde{f}_t(x^*). \end{aligned}$$

The RHS above is bounded by the regret of the algorithm 1 assuming that the cost functions are $\tilde{f}_t$. We therefore proceed to bound this regret. of the algorithm with cost functions $\tilde{f}_t$.

The cost functions $\tilde{f}_t$ can be written in terms of the univariate functions

$$g_t(y) = \frac{r}{8(r_t \cdot x_t)^2} \cdot y^2 - \frac{1}{r_t \cdot x_t}\left(1 + \frac{r}{4}\right) \cdot y + \frac{r}{8} + 1 - \log(r_t \cdot x_t)$$

as $\tilde{f}_t(x) = g_t(r_t \cdot x_t)$. Now we note that even though the statement of the main Theorem assumes that the cost functions can be written in terms of a *single* univariate function g for all t, the proof of the theorem is flexible enough to handle different functions g_t for different t, as long as conditions 3. and 4. in the main Theorem on the first and second derivatives of g hold uniformly with the same constants a and b for all functions g_t.

Furthermore, the proof only requires the bound a on the magnitude of the first derivatives at the points x_t which the algorithm produces. Thus, we can now estimate the a and b parameters for the g_t functions as follows: $g_t'(r_t \cdot x_t) = -\frac{1}{(r_t \cdot x_t)} \in [-\frac{1}{r}, 0]$, so we choose $a = \frac{1}{r}$. For any portfolio $x \in \Delta_n$, and $g_t''(r_t \cdot x) = \frac{r}{4(r_t \cdot x)^2} \geq \frac{r}{4}$, thus we choose $b = \frac{r}{4}$. The regret bound is now obtained via the bound of the main theorem. $\square$

5 Experiments

We tested the performance of the best CRP and Algorithm Exp-Concave-FTL as well as its regret on stock market data. The following graph in Figure 1 was generated using real NYSE quotes for the 1000 trading days from 2001 and 2005 obtained form Yahoo! finance. We randomly chose twenty S&P500 stocks, and computed the performance of the best CRP and Algorithm Exp-Concave-FTL on the relevant trading period, varying the trading frequency from daily to every 50 trading days (only integer periods which divide 1000 were tested). The regret, i.e difference between the log wealth of both methods is also depicted.

For various choices of stocks, the fact that the regret remains pretty much constant was consistent, as predicted by the theoretical arguments in the paper.

The change in performance of the best CRP with respect to trading frequency was not conclusive, at times agreeing with theory and at times not. In one sense, however, this change does agree with theory: in the previous work of [AHKS06], Figure 6 depicts the performance of several online portfolio management as varies with the trading period. These more comprehensive experiments, which measure the average APY (annual percentage yield) in an experiment of sampling a set of stocks from the S&P 500 (the average is taken over the different samples of random stocks from S&P 500). These experiments clearly show that on an average, the performance of many online algorithms, as well as that of the best CRP, improves with trading frequency.

6 Conclusions

We have presented an efficient algorithm for regret minimization with exp-concave loss functions whose regret strictly improves upon the state of the art. For the problem of portfolio selection, the regret is bounded in terms of the observed variation in stock returns rather than the number of iterations. This is the first theoretical improvement

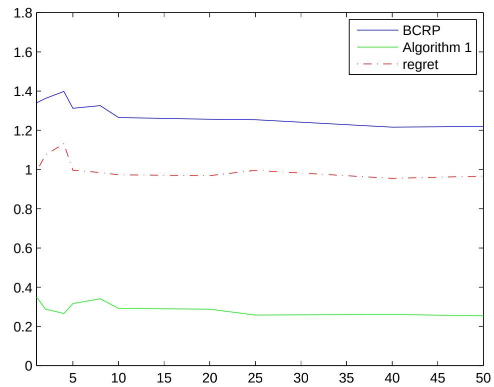

Figure 1: Best CRP and Algorithm Exp-Concave-FTL vs. Trading Period. The x-axis denotes the trading period in days, and the y-axis is the log-wealth factor (i.e. a real number).

in regret bounds for universal portfolio selection algorithms since the work of Cover [Cov91].

We show how this fact implies that in the standard Geometric Brownian Motion model for stock prices the regret does not increase with trading frequency, hence giving the first universal portfolio selection algorithm whose performance improves when the underlying assets are close to GBM. This serves as a bridge between universal portfolio theory and stochastic portfolio theory.

Open questions: It remains an intriguing open question to improve the dependence of the regret in portfolio selection in terms of other important parameters aside from the quadratic variability: our dependence on the number of stocks in the portfolio (previously denoted n) is linear. In contrast [HSSW96] obtain a logarithmic dependence in this parameter. Our regret bounds also depend on the market variability parameter (denoted r), whereas Cover's original algorithm does not have this dependence at all. Is it possible to obtain a $O(\log Q)$ regert bound, via an efficient algorithm, that behaves better with respect to n, r?

References

[AHKS06] Amit Agarwal, Elad Hazan, Satyen Kale, and Robert E. Schapire. Algorithms for portfolio management based on the newton method. In *Proceedings of the 23rd international conference on Machine learning*, ICML '06, pages 9–16, New York, NY, USA, 2006. ACM.

[Bac00] Louis Bachelier. Théorie de la spéculation. *Annales Scientifiques de l'École Normale Supérieure*, 3(17):21–86, 1900.

[BK99] Avrim Blum and Adam Kalai. Universal portfolios with and without transaction costs. *Machine Learning*, 35(3):193–205, 1999.

[BS73] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

[BV04] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004.

[CB03] Jason E Cross and Andrew R Barron. Efficient universal portfolios for past dependent target classes. *Mathematical Finance*, 13(2):245–276, 2003.

[CBMS07] Nicolò Cesa-Bianchi, Yishay Mansour, and Gilles Stoltz. Improved second-order bounds for prediction with expert advice. *Machine Learning*, 66(2-3):321–352, 2007.

[Cov91] Thomas Cover. Universal portfolios. *Mathematical Finance*, 1:1–19, 1991.

[HAK07] Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Mach. Learn.*, 69(2-3):169–192, December 2007.

[HK09] Elad Hazan and Satyen Kale. On stochastic and worst-case models for investing. In Y. Bengio, D. Schuurmans, J. Lafferty, C. K. I. Williams, and A. Culotta, editors, *Advances in Neural Information Processing Systems 22*, pages 709–717. 2009.

[HK10] Elad Hazan and Satyen Kale. Extracting certainty from uncertainty: regret bounded by variation in costs. *Machine Learning*, 80:165–188, 2010.

[HK11] Elad Hazan and Satyen Kale. Better algorithms for benign bandits. *Journal of Machine Learning Research*, 12:1287–1311, 2011.

[HSSW96] David P. Helmbold, Robert E. Schapire, Yoram Singer, and Manfred K. Warmuth. On-line portfolio selection using multiplicative updates. In *ICML*, pages 243–251, 1996.

[Jam92] Farshid Jamshidian. Asymptotically optimal portfolios. *Mathematical Finance*, 2:131–150, 1992.

[KS04] Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer Verlag, New York, NY, USA, 2004.

[KV03] Adam Kalai and Santosh Vempala. Efficient algorithms for universal portfolios. *Journal of Machine Learning Research*, 3:423–440, 2003.

[KV05] Adam Kalai and Santosh Vempala. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 71(3):291–307, 2005.

[MF93] Neri Merhav and Meir Feder. Universal schemes for sequential decision from individual data sequences. *IEEE Transactions on Information Theory*, 39:1280–1292, 1993.

[Osb59] M. F. M. Osborne. Brownian motion in the stock market. *Operations Research*, 2:145–173, 1959.

[Zin03] Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In Tom Fawcett and Nina Mishra, editors, *ICML*, pages 928–936. AAAI Press, 2003.

A Proof of the Vector Harmonic Series Lemma

The proof is based on the following fact, whose one-dimensional analogue is an easy consequence of the Taylor expansion of the logarithm.

Lemma 13. *Let $A \succeq B \succ 0$ be positive definite matrices. Then*

$$A^{-1} \bullet (A - B) \;\le\; \log \frac{|A|}{|B|}$$

where $|A|$ denotes the determinant of matrix A.

Proof. For any positive definite matrix C, denote by $\lambda_1(C), \lambda_2(C), \ldots, \lambda_n(C)$ its (positive) eigenvalues. Denote by $\mathbf{Tr}(C)$ the trace of the matrix, which is equal to the sum of the diagonal entries of C, and also to the sum of its eigenvalues.

Note that for the matrix product $A \bullet B = \sum_{i,j=1}^{n} A_{ij} B_{ij}$ defined earlier, we have $A \bullet B = \mathbf{Tr}(AB)$ (where AB is the standard matrix multiplication), since the trace is equal to the sum of the diagonal entries. Therefore,

$$
\begin{aligned}
A^{-1} \bullet (A-B) &= \mathbf{Tr}(A^{-1}(A-B)) \\
&= \mathbf{Tr}(A^{-1/2}(A-B)A^{-1/2}) \\
&= \mathbf{Tr}(I - A^{-1/2}BA^{-1/2}) \\
&= \sum_{i=1}^{n} \left[1 - \lambda_i(A^{-1/2}BA^{-1/2})\right] \\
&\quad (\because \mathbf{Tr}(C) = \sum_{i=1}^{n} \lambda_i(C)) \\
&\leq -\sum_{i=1}^{n} \log \left[\lambda_i(A^{-1/2}BA^{-1/2})\right] \\
&\quad (\because 1 - x \leq -\log(x)) \\
&= -\log \left[\prod_{i=1}^{n} \lambda_i(A^{-1/2}BA^{-1/2})\right] \\
&= -\log |A^{-1/2}BA^{-1/2}| \\
&= \log \left[\frac{|A|}{|B|}\right].
\end{aligned}
$$

In the last equality we use the following facts about the determinant of matrices: $|A| = \prod_{i=1}^{n} \lambda_i(A)$, $|AB| = |A||B|$ and $|A^{-1}| = \frac{1}{|A|}$. $\square$

We can now prove the vector harmonic series Lemma 3 :

Lemma 3.

$$
\sum_{t=1}^{T} v_t^\top (A + \textstyle\sum_{\tau=1}^{t} v_\tau v_\tau^\top)^{-1} v_t \leq \log \left[\frac{|A + \sum_{\tau=1}^{T} v_\tau v_\tau^\top|}{|A|}\right].
$$

Proof. Let $A_t := A + \sum_{\tau=1}^{t} v_\tau v_\tau^\top$ and denote $A_0 = A$. Now by the lemma above

$$
\begin{aligned}
\sum_{t=1}^{T} v_t A_t^{-1} v_t &= \sum_t A_t^{-1} \bullet v_t v_t^\top \\
&= \sum_t A_t^{-1} \bullet (A_t - A_{t-1}) \\
&\leq \sum_i \log \frac{|A_t|}{|A_{t-1}|} \\
&= \log \left[\frac{|A_T|}{|A_0|}\right].
\end{aligned}
$$

$\square$

B Tail bounds for Normal random variables

The following standard bound on the tail of the Normal distribution is given here for completeness.

Lemma 14. *Let $X \sim N(\mu, \sigma^2)$, then*

$$\Pr[|X - \mu| > x] \leq \frac{\sigma}{x} \exp\left(-\frac{x^2}{2\sigma^2}\right).$$

Proof. First, consider the case $\mu = 0, \sigma = 1$. Then by definition

$$\Pr[X > x] = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp(-t^2/2)\, dt$$

Since the exponent is a convex function, we can upper bound it by the first derivative at the point $t = x$ and obtain:

$$\int_x^\infty \frac{1}{\sqrt{2\pi}} \exp(-t^2/2)\, dt \leq \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp(-x^2/2 - x(t-x))\, dt$$

here the new exponent is the linearization of $-t^2/2$ at $t = x$. Then pull out factors which don't depend on t to get

$$\frac{\exp(x^2/2)}{\sqrt{2\pi}} \int_x^\infty \exp(-xt)\, dt$$

and doing that last integral gives the bound:

$$\Pr[X > x] \ \leq \frac{1}{\sqrt{2\pi} x} \exp(-x^2/2)$$

Hence, by symmetry of the distribution, we get for $X \sim N(0, 1)$:

$$\Pr[|X| > x] \leq \frac{\sqrt{2}}{\sqrt{\pi} x} \exp(-x^2/2) \leq \frac{1}{x} \exp(-x^2/2)$$

Next, consider $Y \sim N(\mu, \sigma^2) = \mu + \sigma N(0, 1)$. Hence,

$$\Pr[|Y - \mu| > x] = \Pr\left[|X| > \frac{x}{\sigma}\right] \leq \frac{\sigma}{x} \exp\left(-\frac{x^2}{2\sigma^2}\right).$$

□