ENERGY, ENTROPY, AND ARBITRAGE

SOUMIK PAL AND TING-KAM LEONARD WONG

Abstract. We introduce a framework to analyze the relative performance of a portfolio with respect to a benchmark market index. We show that this relative performance has three components: a term that can be interpreted as energy coming from the market fluctuations, a relative entropy term that measures “distance” between the portfolio holdings and the market capital distribution, and another entropy term that can be controlled by the trader by choosing a suitable strategy. The first aids growth in the portfolio value, and the second poses as relative risk of being too far from the market. We give several explicit controls of the third term that allows one to outperform a diverse volatile market in the long run. Named energy-entropy portfolios, these strategies work in both discrete and continuous time, and require essentially no probabilistic or structural assumptions. They are well-suited to analyze a hierarchical portfolio of portfolios and attribute relative risk and reward to different levels of the hierarchy. We also consider functionally generated portfolios (introduced by Fernholz) in the case of two assets and the binary tree model and give a novel explanation of their efficacy.

1. Introduction

Consider an equity market with several stocks. Consider two portfolios both starting with one dollar. In one, we always invest an equal proportion of the portfolio value in every stock. This is referred to as the equal-weighted portfolio. In the other, we divide one dollar initially in each stock proportional to its capitalization, buy whatever number of shares afforded by that amount, and hold them for all future time. This is the market portfolio.

It has been common wisdom in some quarters that, in a volatile market, the equal-weighted portfolio performs better than the market portfolio (“beats the market”) in the long run; see Plyakha, Uppal and Vilkov [PUV12] for a recent empirical study for the US stock market. It is also intuitively understood that this effect is due to “rebalancing”, i.e., buying a stock when its price is low, and selling it when its price is high. This effect is automatically captured in an equal-weighted portfolio. More generally, one can imagine a constant-weighted portfolio which, at every time point,
invests a fixed proportion of its total value in each stock. Such portfolios are also rebalancing and have been frequently shown in practice to outperform the market.

Several attempts have been made to make this idea mathematically precise. Fernholz and Shay [FS82] were one of the firsts to provide a mathematical framework via the concept of excess growth rate in a continuous time, continuous price set-up. Many other authors have put forwarded their analyses. For example, Dempster, Evstigneev, and Schenk-Hoppe [DESH07] explain rebalancing when returns are stationary ergodic, Platen and Rendek [PR11] argue that the equal-weighted portfolio approximates the numéraire portfolio in a large market, Chan and Ramkumar [CR11] advocate rebalancing to reduce tracking error, Kuhn and Luenberger [KL10] study how rebalancing affects the growth rate for a log-utility investor, while Bouchey et al. [BNPS12] and Chincarini and Kim [CK06] look at aspects of rebalancing from a practitioner’s point of view.

Moreover, it has been shown by Fernholz [Fer02] that one can generalize constant-weighted portfolios to what are called functionally generated portfolios. These are precisely formulated portfolios some of which can be proved to beat the market asymptotically in a continuous time, continuous price framework under the assumptions of diversity and in the presence of sufficient volatility. For precise statements of this result consult Fernholz and Karatzas [FK05], Fernholz, Karatzas and Kardaras [FKK05] and the fascinating survey by Fernholz and Karatzas [FK09, Chapter 3]. In short, prices are modeled as a multidimensional positive Itô process. The assumption of diversity encapsulates the idea that no stock is allowed to dominate the entire market, while the presence of sufficient volatility is a technical condition that requires a uniform positive lower bound for the eigenvalues of the volatility matrix at every time. The great benefit of such functionally generated portfolios is that the portfolio weights (i.e., the proportions invested in each stock) depend only on the current market weights. This can be used, say, to create enhanced indexing. Somewhat less satisfactorily, such strategies are specially designed and it is not clear if their outperformance is robust to slight changes in the formula. It is left as an open question [FK09, Remark 11.5] if there are portfolios, not functionally generated, that beat the market under similar assumptions. Our answer is a definitive yes.

Our primary contribution is the introduction of a framework to analyze the relative performance of any portfolio with respect to the market portfolio. This information-theoretic framework is valid in both discrete and continuous time and requires no stochastic modeling assumptions on the stock prices. The key quantities: free energy and relative entropy of our portfolio weights from the market weights can be calculated exactly by observing the stock prices. In the following section we will precisely define the italicized expressions above. This framework naturally begets a broad class of strategies, which includes the constant-weighted portfolios, that outperform the market portfolio in a diverse and sufficiently volatile market.
We stress that our results are true for an arbitrary deterministic process of stock prices. These portfolios are easy to implement and depend only on the current market weights and the portfolio weights at just the previous time period. They do not require any statistical estimation of parameters at all although, with additional information, they can be optimized and tailored to practical considerations. However, unless they are constant weighted, these are not functionally generated when interpreted in continuous time.

The rest of the article is divided in the following way.

(i) Section 2 describes the information-theoretic framework and the energy-entropy strategies.

(ii) Section 3 considers the problem of hierarchical portfolios, i.e., a portfolio of portfolios. It is desirable to have a method to attribute the gain and risk of the strategy separately for each level of the hierarchy. Fortunately energy-entropy portfolios are tailored for this purpose due to a chain-rule property that is well-known for the relative entropy and is proved here for the free energy.

(iii) Section 4 considers functionally generated portfolios in a discrete-time, two asset, binary tree set-up. We interpret both the risk and reward from rebalancing by combinatorial matchings. This sheds a new light on how such portfolios work.

(iv) We attempt to find optimal functionally generated portfolios in the two asset case in Section 5.

(v) We end with some data analysis that shows the performance of an energy-entropy strategy applied to historical data: (i) rebalancing between Apple and Starbucks between 1994 and 2012, and (ii) rebalancing between 18 emerging market economies between 2001 and 2013.

Information-theoretic analysis of portfolios has appeared before in the context of asymptotic log-optimal portfolios. See for example Kelly [Kel56] and Cover [Cov91]. Although similar mathematical ideas are used in this paper, our problem and analysis are very distinct from those cited above.

ACKNOWLEDGMENT

This work was mostly done during a summer spent at Parametric Portfolio Associates in Seattle, WA. We are grateful to Parametric for their hospitality and generous support. We particularly thank Paul Bouchey, Alex Paulsen, Mahesh Pritamani, Vassilii Nemtchinov, and David Stein for numerous comments, problem statements, and stimulating conversations. We are indebted to Robert Fernholz and Ioannis Karatzas for their detailed comments on an earlier draft which greatly improved the manuscript. Soumik would also like to thank Rik Sen from HKUST for many happy discussions.

2. Energy-entropy rebalancing

We start with the set-up that we have $n$ firms in a financial market. Although we restrict ourselves to equity markets, in fact, any kind of assets
can be used for the following analysis as long as we properly interpret the market portfolio.

The market capital distribution at any time is the vector of proportions of the total market capital that belongs to each firm. That is, if \( X_i(t) \) is the total capitalization of the \( i \)th firm at time \( t \), then the corresponding market weight is

\[
\mu_i(t) = \frac{X_i(t)}{\sum_{j=1}^{n} X_j(t)}, \quad i = 1, 2, \ldots, n.
\]

The market capital distribution is then given by the vector \( \mu = (\mu_1, \ldots, \mu_n) \).

We retain the convenient assumption that each stock has one outstanding share and any fraction of that share can be traded. We also assume zero transaction costs and dividend payments.

The primary concepts involved in this section are information-theoretic and the suggested reference is the book by Cover and Thomas [CT06]. Let \( \pi = (\pi_1, \ldots, \pi_n) \) be the weights of an all-long portfolio, i.e., \( \pi_i \) proportion of the total value is invested in asset \( i \). We think of \( \pi \) as a probability distribution on \( n \) atoms, so short sales are not allowed. The space of all portfolios/probability distributions is the unit simplex in dimension \( n \) defined by

\[
\mathcal{G}_n := \left\{ x = (x_1, \ldots, x_n) : x_i \geq 0 \text{ for all } i, \sum_{i=1}^{n} x_i = 1 \right\}.
\]

Let \( \mu \) and \( \nu \) be any two elements of \( \mathcal{G}_n \). Recall the relative entropy of \( \nu \) with respect to \( \mu \) is defined by

\[
H(\nu | \mu) := \sum_{i=1}^{n} \nu_i \log \left( \frac{\nu_i}{\mu_i} \right).
\]

The value is taken to be infinity whenever \( \nu_i > 0 \) for some \( i \) such that \( \mu_i = 0 \) (i.e., \( \nu \) is not absolutely continuous with respect to \( \mu \)).

It is well-known [CT06, Theorem 2.6.3] that the relative entropy is always nonnegative. Note that \( H \) is not symmetric in the two arguments. Nevertheless, the function \( H \) can be thought of as a “distance” of the distribution \( \nu \) from \( \mu \) and is known in statistics as the Kullback-Leibler divergence. This “distance” is well-behaved while \( \mu \) is away from the boundary of the unit simplex. However, as \( \mu \) approaches a boundary, some of its coordinates get very small. If those coordinates of \( \pi \) remain bounded away from zero, the relative entropy blows up.

To understand the role of the relative entropy we consider a discrete time model with no assumptions. Suppose \( \pi(t) = (\pi_1(t), \ldots, \pi_n(t)) \) is our portfolio at time \( t \). That is, at time \( t \), when we have observed prices until (and including) time \( t \), we invest \( \pi_i(t) \) proportion of our current portfolio value in stock \( i \). In other words, \( \pi \) is adapted to the underlying filtration of stock prices and possible extraneous events.
Due to fluctuations in prices at time $t+1$ our portfolio loses or gains value. Let $\bar{V}(t)$ denote the value of the portfolio at time $t$ starting with $\bar{V}(0) = 1$. When $\pi(t) = \mu(t)$, for all $t$, this portfolio is called the market portfolio. The value of this special portfolio will be denoted by $S$ ("the total market"). Let $V(t)$ denote the relative ratio $\bar{V}(t)/S(t)$. This is a change of numéraire procedure that will allow us to compare the relative performance of $\pi$ with respect to $\mu$. For example, when we say that the portfolio $\pi$ beats the market at time $t$ we mean that $V(t) > 1$.

We make the following definition for future use: for any process $A(t)$, $t = 0, 1, \ldots, T$, let $\Delta A(t)$ denote the discrete difference $A(t+1) - A(t)$.

**Lemma 2.1.** We have the following formula

$$V(t+1) = \sum_{i=1}^{n} \pi_i(t) \frac{\mu_i(t+1)}{\mu_i(t)}, \quad t = 0, 1, \ldots.$$  

**Proof.** Recall $X_i(t)$ is the capitalization of the $i$th stock. Then it is straightforward to check that

$$\Delta \bar{V}(t) = \sum_{i=1}^{n} \pi_i(t) \frac{\Delta X_i(t)}{X_i(t)}, \quad \text{or,}$$

$$\frac{\bar{V}(t+1)}{\bar{V}(t)} = \sum_{i=1}^{n} \pi_i(t) \frac{X_i(t+1)}{X_i(t)}.$$  

Multiply both sides of the above by the factor $S(t)/S(t+1)$ to obtain

$$V(t+1) = \sum_{i=1}^{n} \pi_i(t) \frac{X_i(t+1)}{X_i(t)}.$$  

This completes the proof of the lemma. \hfill $\square$

It follows from the previous lemma that

$$\Delta \log V(t) = \log \left( \sum_{i=1}^{n} \pi_i(t) e^{\Delta y_i(t)} \right), \quad \text{where we define } y_i(t) = \log \mu_i(t).$$  

**Definition 2.2.** Consider a random variable $Y$ such that

$$Y = \Delta y_i(t), \quad \text{with probability } \pi_i(t), \quad i = 1, 2, \ldots, n.$$  

Define the free energy at time $t$ by

$$\gamma^*_\pi(t) := \log \left( e^{Y - E(Y)} \right) = \Delta \log V(t) - \sum_{i=1}^{n} \pi_i(t) \Delta y_i(t).$$

The definition of free energy is a generalization of the one in Thermodynamics. It also coincides with the concept of excess growth rate of Fernholz [Fer02] in the continuous time stochastic calculus framework. This can be seen by a Taylor series approximation. For this reason we will often refer to $\gamma^*_\pi(t)$ as the excess growth rate in the following text.
Lemma 2.3. The free energy $\gamma^*(t)$ is always nonnegative, and is strictly positive unless $Y$ is a constant. Moreover, $\gamma^*(t)$ does not depend on the choice of the numéraire.

That is, free energy is strictly positive unless the market weights of all the assets over which our portfolio is spread move by the exact same ratio.

Proof. Since $\gamma^*_\pi(t) = \log E \left( e^{Y-E(Y)} \right)$, it follows from Jensen’s inequality that $\gamma^*_\pi(t)$ is strictly positive unless $Y = E(Y)$ with $\pi(t)$ probability one.

To see the numéraire invariance, consider first the case of changing the numéraire from dollar to the market $S$. Equation (4) shows that
$$\Delta \log V(t) = \log \left( \sum_{i=1}^{n} \pi_i(t) X_i(t+1) / X_i(t) \cdot S(t) / S(t+1) \right) = \Delta \log \tilde{V}(t) - \Delta \log S(t).$$

In the same way
$$\sum_{i=1}^{n} \pi_i(t) \log \frac{\mu_i(t+1)}{\mu_i(t)} = \sum_{i=1}^{n} \pi_i(t) \log \frac{X_i(t+1)}{X_i(t)} - \left( \sum_{i=1}^{n} \pi_i(t) \right) \Delta \log S(t)$$
$$= \sum_{i=1}^{n} \pi_i(t) \log \frac{X_i(t+1)}{X_i(t)} - \Delta \log S(t).$$

Taking difference between the last two displayed equations, we get
$$\gamma^*(t) = \Delta \log \tilde{V}(t) - \sum_{i=1}^{n} \pi_i(t) \log \frac{X_i(t+1)}{X_i(t)}.$$

This shows that the free energy remains the same whether we use dollar value or the market as the numéraire. An identical calculation proves the claim for an arbitrary choice of numéraire. □

Let $\Gamma^*_\pi(t)$ be the cumulative excess growth rate till time $t$, i.e., $\Gamma^*_\pi(t) = \sum_{u=0}^{t-1} \gamma^*_\pi(u)$. Now we observe:
$$\sum_{i=1}^{n} \pi_i(t) \Delta y_i(t) = \sum_{i=1}^{n} \pi_i(t) \log \left( \frac{\mu_i(t+1)}{\mu_i(t)} \right)$$
$$= \sum_{i=1}^{n} \pi_i(t) \log \left( \frac{\pi_i(t)}{\mu_i(t)} \right) - \sum_{i=1}^{n} \pi_i(t) \log \left( \frac{\pi_i(t)}{\mu_i(t+1)} \right)$$
$$= H(\pi(t) \mid \mu(t)) - H(\pi(t) \mid \mu(t+1)).$$

(7)
Since $V(0) = 1$, combining (5) and (7) we get
\[
\log \frac{V(T)}{V(0)} = \log V(T) = \sum_{t=0}^{T-1} \Delta \log V(t)
\]
\[
= \Gamma^\pi_\pi(T) + \sum_{t=0}^{T-1} \left[ H(\pi(t) \mid \mu(t)) - H(\pi(t) \mid \mu(t+1)) \right]
\]
\[
= \Gamma^\pi_\pi(T) + \sum_{t=0}^{T-1} \left[ H(\pi(t) \mid \mu(t)) - H(\pi(t+1) \mid \mu(t+1)) \right]
\]
\[
+ \sum_{t=0}^{T-1} \left[ H(\pi(t+1) \mid \mu(t+1)) - H(\pi(t) \mid \mu(t+1)) \right]
\]
\[
= \Gamma^\pi_\pi(T) + H(\pi(0) \mid \mu(0)) - H(\pi(T) \mid \mu(T))
\]
\[
+ \sum_{t=0}^{T-1} \left[ H(\pi(t+1) \mid \mu(t+1)) - H(\pi(t) \mid \mu(t+1)) \right].
\]

(8)

The following mnemonic is helpful to remember the previous decomposition:
\[
\Delta \log \text{relative value} = \Delta \text{energy} - \Delta \text{relative entropy} + \Delta \text{control}.
\]

The reason for calling the third term control will be immediately clear.

Notice what happens when we let $\pi$ be a constant-weighted portfolio. That is, for every $t$, we take $\pi(t) \equiv \pi(0) := \pi$. In that case every term inside the summand in the final line of (8) is zero, i.e., $\Delta \text{control} \equiv 0$. Thus
\[
\log V(T) = \Gamma^\pi_\pi(T) + H(\pi \mid \mu(0)) - H(\pi \mid \mu(T)).
\]

(9)

Now there are two ways that $\log V(T)$ can be large. A short to medium term effect can come from the relative entropy. For example, the market might be more concentrated at time zero and gradually gets more diverse at time $T$. In that case the equal weighted portfolio causes $H(\pi \mid \mu(0)) - H(\pi \mid \mu(T))$ to be a positive number. The reverse can happen if the market gradually becomes more concentrated with the passage of time.

However, long term growth comes from the accumulated energy $\Gamma^\pi(\cdot)$. If we assume that (i) $\mu$ does not approach the boundary of the simplex too rapidly (a notion to be defined precisely later), and (ii) $\Gamma^\pi(T)$ tends to infinity with $T$ (due to price fluctuations), sooner or later $\log V(T)$ is going to be forever positive. This can be interpreted as a long-term relative arbitrage opportunity with respect to the market. We would like to stress that we do not claim this to be \textit{free lunch}, but only as a dependable strategy whose risk and reward can be computed via the relative entropy and energy.
Remark 2.4. Consider the Shannon entropy function $H(\pi) = -\sum_{i=1}^{n} \pi_i \log \pi_i$. Notice that, in the case of constant-weighted portfolios

$$H(\pi | \mu) + H(\pi) = -\sum_{i=1}^{n} \pi_i \log \mu_i = -\log \prod_{i=1}^{n} \mu_i^{\pi_i}.$$  

The function $\mu \mapsto \prod_{i=1}^{n} \mu_i^{\pi_i}$ generates the constant-weighted portfolio according to the definition of Fernholz [Fer02].

If the portfolio is not constant-weighted, we can choose the weights such that the $\Delta$control term is either positive or negative. A positive $\Delta$control suggests moving away from the current market weights, often referred to as a contrarian strategy. However, in the future, if the market moves in an opposite direction this can affect the relative entropy risk. A less risky strategy will be to choose $\Delta$control $< 0$ so as to have $\pi(t+1)$ closer to $\mu(t+1)$ than $\pi(t)$. If we insist that $\Delta$control + $\Delta$energy remains positive, this can still lead to an outperformance of our portfolio relative to the market.

**Definition 2.5.** An energy-entropy portfolio is a portfolio $\pi(\cdot)$ which is adapted to the underlying filtration and satisfies for all time $t$:

$$\Delta D(t) := \gamma^{\ast}_\pi(t) + H(\pi(t+1) | \mu(t+1)) - H(\pi(t) | \mu(t+1)) \geq 0. \tag{10}$$  

That is, an energy-entropy portfolio allows a loss of relative entropy which is at most the gain in the free energy in the previous time period.

A greedy-entropy portfolio is an energy-entropy portfolio $\pi(\cdot)$ that satisfies

$$H(\pi(t+1) | \mu(t+1)) \geq H(\pi(t) | \mu(t+1)), \quad \text{for all } t. \tag{11}$$

In other words, we always choose to increase the relative entropy between our portfolio and the current market weights.

**Remark 2.6.** We have already seen that any constant-weighted portfolio is energy-entropy. The market portfolio itself is also energy-entropy. This is because

$$\Delta D(t) = \gamma^{\ast}_\mu(t) - H(\mu(t) | \mu(t+1)) = 0.$$  

That is, the market portfolio uses up all its energy in following the current market weights. In general, one should think of $\Gamma^{\ast}(t)$ as accumulated free energy that can be put to work to reduce relative entropy and draw the portfolio closer to the market.

Clearly, for an energy-entropy portfolio, equation (8) allows us to write an expression similar to (9):

$$\log V(T) = D(T) + H(\pi(0) | \mu(0)) - H(\pi(T) | \mu(T)). \tag{12}$$

The right side of the above expression can be thought of as drift plus risk. The drift term further decomposes as the sum of energy and control terms.
Figure 1. A greedy-entropy strategy chooses $\pi(t+1)$ so that the relative entropy increases.

Remark 2.7. “Risk” is a loaded term in finance; so we would like to clarify in what sense $H(\pi(0) \mid \mu(0)) - H(\pi(T) \mid \mu(T))$ poses as risk. As always, this term is only significant when it is negative, i.e., when $H(\pi(T) \mid \mu(T))$ is large. This term is large when our portfolio is far (in the relative entropy sense) from the current market weights. This can cause significant short term depreciation in the portfolio relative value. For example, this is caused when some stocks in our portfolio fail. But there is also a subtler reason. Suppose at time $T$ we stop active rebalancing, and let our portfolio drift. That is, from time $T$ onward we follow a buy-and-hold strategy. Then, in the future, the relative value of our portfolio can sink due to the underperformance of stocks which are overweighted at time $T$.

This brings us to the main theorem in this section which, we repeat, requires no stochastic modeling at all. It is a deterministic set-up that allows for shocks in prices and also allows stocks to fail.

**Theorem 2.8.** Suppose $\pi$ is an energy-entropy portfolio such that

$$
\lim_{T \to \infty} [D(T) - H(\pi(T) \mid \mu(T))] = \infty.
$$

Then $\pi$ eventually outperforms the market, i.e, $\lim_{T \to \infty} V(T) = \infty$.

In particular, suppose that there exists a compact subset of the open unit simplex in $n$ dimensions such that $\mu$ never escapes from that set. Then any energy-entropy portfolio for which the drift process $D(\cdot)$ grows unbounded eventually outperforms the market.

**Proof.** The proof immediately follows from decomposition (12). The significance of the compact set is that the function $\pi \mapsto H(\pi \mid \mu)$ is uniformly bounded as $\pi$ varies over $\mathcal{G}_n$ and $\mu$ varies over the compact set. $\square$

2.1. **Implementation and extensions to continuous time.** Theorem 2.2 brings the question of describing the set of all energy-entropy portfolios. In discrete time this class is too unwieldy. One can conceivably jump around the simplex and still keep the same relative entropy. This gives a lot of
flexibility to practitioners. For example, even within the same level surface of relative entropy, certain portfolios might be more attractive in terms of transaction costs. In continuous time, however, it seems natural to demand nicer continuity properties of $\pi$.

In this section we provide a classification of energy-entropy portfolios in terms of vector fields. Next we consider the continuous time framework in [FK09] and show that portfolios exist as finite variation strategies that are relative arbitrages in the set-up of Fernholz. In particular, other than the constant-weighted portfolios most of them are not functionally generated. In fact, the value of $\pi(t)$ depends on the entire history of the prices till time $t$.

We start with a classification of directions that lead to an increase of relative entropy. Fix a point $\mu \in G_n$ and consider the function $H_\mu : G_n \to [0, \infty)$ defined by $H_\mu(\pi) = H(\pi | \mu)$. This is a nonnegative strictly convex function which is zero at $\pi = \mu$. The level sets of this function are smooth and they grow towards the boundary of the simplex. At each point of a level set there is a tangent hyperplane and an outward normal of $H_\mu$ which induce a natural orthogonal coordinate system. We should point out that although the calculation of these objects require a specific local coordinate system, the geometric idea does not.

**Condition 2.9.** For every $\mu \in G_n$ consider a vector field $U_\mu$ that associates a direction vector $U_\mu(\pi)$ with each $\pi \in G_n$. Each such vector field is associated with a flow, i.e., a system of ordinary differential equations (ODE) that directs a particle along this field:

$$\pi'(u) = U_\mu(\pi(u)), \quad \pi(0) = \pi \in G_n.$$  

We assume the following properties on the vector field that must hold for every $\mu \in G_n$:

(i) $U_\mu(\pi) \perp 1$, for all $\pi \in G_n$, where $1$ is the vector of all ones.

(ii) $U_\mu(\pi) \cdot \nabla H_\mu(\pi) \geq 0$, for $\pi \in G_n$.

(iii) The flow, starting from any $\pi \in G_n$, never exits the unit simplex.

Condition (i) guarantees that the direction vector is parallel to $G_n$. Condition (ii) states that the direction vector points towards the outward normal of the level surface of the relative entropy at $\pi$. Condition (iii) guarantees that the portfolio defined by equation (15) below stays inside $G_n$.

The following few examples of flows will elucidate our approach. It is straightforward to verify that they all satisfy Condition 2.9. For any two vectors $a$ and $b$, the notation $a \propto b$ will denote that $a$ is a nonnegative multiple of $b$. This is necessary to meet our boundary condition.

(i) **(Gradient flow.)** Let

$$U_\mu \propto \nabla H_\mu = \left( \log \frac{\pi_i}{\mu_i} - \frac{1}{n} \sum_{j=1}^{n} \log \frac{\pi_j}{\mu_j}, \; i = 1, 2, \ldots, n \right)$$
This corresponds to the flow that follows the steepest gradient towards increasing relative entropy. In fact, the gradient can be calculated with respect to any metric on $\mathcal{G}_n$. A particularly important one in recent years has been the Wasserstein metric due to the seminal work of Jordan, Kinderlehrer, and Otto [JKO98] that relates heat flow with the gradient flow of the entropy.

Moving along the gradient has the physical intuition of doing the least amount of work to increase relative entropy. This can be interpreted as the trade with the least turnover that increases relative entropy.

(ii) **(Functionally generated gradient flow.)** For an increasing smooth function $R$ on $[0, \infty)$, a variation of the previous example will be to consider $U_\mu \propto \nabla R(H_\mu)$.

(iii) **(Directional flowing in.)** For every point $\pi$ we choose another point in $\mathcal{G}_n$ which has a higher entropy and we flow along the straight line joining the two.

For example, let $e_i$ be the vector that is 1 in the $i$th coordinate and zero elsewhere. These are the corners of the simplex. Now for each $\pi$ choose the corner that has a higher relative entropy than $\pi$ and move along a straight line path towards that corner.

(iv) **(Directional flowing out.)** A variation of the previous theme will be to join $\mu$ and $\pi$ by a straight line and flow in the direction opposite to $\mu$. This can be easily achieved by setting $U_\mu(\pi) \propto \pi - \mu$. The directional flows are not quite-well defined at $\pi = \mu$. But this can be done by fiat since it hardly matters in continuous time how we do.

Each of these vector fields induces a greedy-entropy portfolio by the following procedure in discrete time. Suppose we decide to choose the portfolio $\pi$ at discrete time points $\{t_0 = 0 < t_1 < \ldots < t_k-1 < t_k = T\}$, starting at $\pi(t_0) = \pi(0)$. We do this inductively. At time $t_{j+1}$ we observe $\mu(t_{j+1})$. Define

$$\pi(t_{j+1}) = \pi(t_j) + \int_{0}^{t_{j+1}-t_j} U_{\mu(t_{j+1})}(\pi(u)) \, du. \quad (15)$$

In other words, we start the flow at $\pi(t_j)$ and follow it till time $t_{j+1} - t_j$. The point where it ends is our new portfolio. It is clear from construction that this choice of portfolio is greedy-entropy as defined in (10).

The general case of energy-entropy portfolios can be now analyzed by choosing a direction of flow that can decrease entropy, but not more than a specified amount. This can be achieved by flowing opposite to the vector fields satisfying Condition 2.9.

Fix any $0 < \lambda < 1$. We will define an energy-entropy portfolio that uses at most $\lambda \times \Delta\text{energy}$ to reduce relative entropy by moving along certain directions.
For any \( \pi \) and \( \mu \in \mathcal{G}_n \) consider a vector field satisfying Condition 2.9. Define the reverse flow by the differential equation
\[
\pi'(u) = -U_\mu(\pi(u)), \quad \pi(0) = \pi.
\]
For any positive \( \alpha \), define the stopping time
\[
\varsigma(\alpha) = \inf \{ u \geq 0 : H(\pi(0) \mid | \mu) - H(\pi(u) \mid | \mu) > \alpha \}.
\]
By convention the infimum of an empty set is taken to be \( \infty \).

We can now define our portfolio as in (15)
\[
\pi(t_{j+1}) = \pi(t_j) - \int_0^{(t_{j+1} - t_j) \wedge \varsigma(\alpha)} U_{\mu(t_{j+1})}(\pi(u)) \, du, \quad \alpha = \lambda \gamma^*(t_j).
\]

In continuous time the above strategy results in a finite variation process \( \pi \) where at every infinitesimal time-interval we choose a direction \( d\pi \) aligned in the direction of the outward normal. The set-up and notations for the next result are taken from [FK09, Chapter 1] except their relative value process \( V_\pi(\cdot)/V_\mu(\cdot) \) is denoted by our usual notation \( V(\cdot) \).

**Theorem 2.10.** Suppose \( \pi \) is of finite variation such that, almost surely, one of the following two measures on the interval \([0, T]\) is nonnegative:

(i) **Greedy-entropy condition:**
\[
\sum_{i=1}^{n} \log \left( \frac{\pi_i(t)}{\mu_i(t)} \right) d\pi_i(t) \geq 0.
\]

(ii) **Energy-entropy condition:**
\[
\sum_{i=1}^{n} \log \left( \frac{\pi_i(t)}{\mu_i(t)} \right) d\pi_i(t) + \gamma^*(t) dt \geq 0.
\]

Then the following decomposition holds for the relative value process
\[
\log V(T) = H(\pi(0) \mid \mu(0)) - H(\pi(T) \mid \mu(T)) + D(T),
\]
where \( D(t) \) is a process increasing in time.
Proof. The proof follows by using the integration-by-parts formula from stochastic calculus. We often drop the time parameter for the sake of clarity.

\[ d \log V(t) - \gamma^*_\pi(t)dt = \sum_{i=1}^{n} \pi_i(t)d \log \mu_i(t), \quad \text{from [FK09 eqn. (3.4)]} \]

\[ = \sum_{i=1}^{n} \pi_i(t) \left( -d \log \frac{\pi_i(t)}{\mu_i(t)} + d \log \pi_i(t) \right) \]

\[ = -\sum_{i=1}^{n} \pi_i(t)d \log \frac{\pi_i(t)}{\mu_i(t)} + \sum_{i=1}^{n} \pi_i(t)d \pi_i(t) \]

\[ = -\sum_{i=1}^{n} \pi_i(t)d \log \frac{\pi_i(t)}{\mu_i(t)}, \quad \text{since } \sum_{i=1}^{n} \pi_i(t) \equiv 1, \]

\[ = \sum_{i=1}^{n} d \left[ -\pi_i(t) \log \frac{\pi_i(t)}{\mu_i(t)} \right] + \sum_{i=1}^{n} \log \frac{\pi_i(t)}{\mu_i(t)}d\pi_i(t) \]

\[ = -dH(\pi(t) \mid \mu(t)) + \sum_{i=1}^{n} \log \frac{\pi_i(t)}{\mu_i(t)}d\pi_i(t). \]

Therefore, by integrating over time interval \([0, T]\) and noticing that \(V(0) = 1\), we get

\[ \log V(T) = H(\pi(0) \mid \mu(0)) - H(\pi(T) \mid \mu(T)) + \int_{0}^{T} \left[ \sum_{i=1}^{n} \log \frac{\pi_i(t)}{\mu_i(t)}d\pi_i(t) + \gamma^*(t)dt \right] \]

\[ = H(\pi(0) \mid \mu(0)) - H(\pi(T) \mid \mu(T)) + D(T). \]

Clearly \(D\) is an increasing process in time under both conditions (i) and (ii). This completes the proof of the theorem. \(\square\)

Notice from [14] that \(\sum_{i=1}^{n} \log \left( \frac{\pi_i(t)}{\mu_i(t)} \right) d\pi_i(t) \geq 0\) is precisely the infinitesimal version of Condition 2.9(ii). Hence the existence of strategies satisfying the assumptions of Theorem 2.10 follows by (i) considering any vector field satisfying Condition 2.9 and is uniformly bounded in \(\mu\) and \(\pi\), (ii) discretize any interval \([0, T]\) in a fine partition of discrete intervals, (iii) run the inductive algorithm described in [15] and [16], and finally (iv) take a limit as the partition size goes to zero. A straightforward argument using the Arzela-Ascoli theorem shows the existence of a limiting continuous finite-variation portfolio that satisfies the assumptions of Theorem 2.10. However, we do not require \(\pi\) to be continuous for the theorem to hold.

Energy-entropy portfolios are typically not functionally generated in the sense of Fernholz [Fer02] since, barring constant-weighted portfolios, the latter are semimartingales, with a nontrivial martingale component that
can be computed by knowing the dynamics of the market weights. Energy-entropy portfolios are, however, of finite variation and typically depend on the entire history of the market weights.

With further assumptions one can quantify the time it takes to outperform the market, which necessarily accompanies a loss of generality of the result.

**Theorem 2.11.** Assume that there is a $\delta > 0$ such that, almost surely,
\begin{equation}
\inf_{t \geq 0} \min_{i} \mu_i(t) \geq \delta.
\end{equation}

Also assume that there is an absolutely continuous portfolio $\pi$ and an $\epsilon > 0$ such that, almost surely,
\begin{equation}
\sum_{i=1}^{n} \log \left( \frac{\pi_i(t)}{\mu_i(t)} \right) \frac{d\pi_i(t)}{dt} + \gamma^*(t) \geq \epsilon, \quad \text{for all } t \geq 0.
\end{equation}

Then, for any $r > 0$, $P \left( \tilde{V}(T) \geq rS(T) \right) = 1$ for all $T \geq (r - \log \delta)/\epsilon$.

**Proof.** Since the Shannon entropy for any $\pi \in \mathcal{G}_n$ is nonnegative, we have
\begin{equation*}
H(\pi(t) \mid \mu(t)) = \sum_{i=1}^{n} \pi_i(t) \log \pi_i(t) - \sum_{i=1}^{n} \pi_i(t) \log \mu_i(t) \leq -\log \delta.
\end{equation*}

Therefore, from (19), we get $\log V(T) \geq \epsilon T + \log \delta$. Thus to guarantee $V(T) \geq r$ we require $T \geq (r - \log \delta)/\epsilon$. This completes the proof. \qed

3. Hierarchical portfolios

Consider a trader in global markets. It is convenient to think of her portfolio in a hierarchical framework: (i) a portfolio that describes the proportion of money invested in each country, (ii) for each country with several sectors in its economy, a portfolio that describes the amount invested in each sector as a proportion of the total money invested in that country, and finally (iii), for each sector of every country, how the allocated amount is distributed among various stocks as proportions of the corresponding total.

The above is an example of a hierarchical portfolio. Similar structures result while investing in a fund of funds or combining performances of different managers. In these situations it is convenient to have both the construction of portfolios and the analysis of their benefit being done in the same hierarchical fashion. In particular, we would like to attribute the free energy and relative entropy risk separately for each level of the hierarchy.

Fortunately, energy-entropy portfolios are tailored to the stated attribution problem. We describe the solution for a two-step hierarchy (“sector” and “stocks”). But the analysis is similar in the case of multiple levels (“country”, “sector”, “stocks”, etc.).

Suppose there are $m$ sectors, and each sector has $n_i$ stocks, $i = 1, \ldots, m$. The “universe” thus consists of $n \leq n_1 + \cdots + n_m$ stocks, with equality when the collection of stocks considered in different sectors are disjoint. Our portfolio $\pi$ is a combination of “sector portfolios” $\pi_i$, $i = 1, \ldots, m$, whose
coordinates \((\pi_{ij}, j = 1, 2, \ldots, n_i)\) add up to one for each \(i\). It is helpful to think of each \(\pi_i\) as a vector of probabilities of length \(n_i\) simply by padding zeroes for stocks where the portfolio puts no weight.

Let the sector weights be \(\lambda = (\lambda_1, \ldots, \lambda_m)\), where \(\sum_{i=1}^{m} \lambda_i = 1\). So

\[
\pi = \sum_{i=1}^{m} \lambda_i \pi_i
\]

is our total portfolio over all stocks in every sector. We first note a fundamental property of both relative entropy and the free energy.

**Lemma 3.1.** Consider two pairs of sector portfolios and sector weights \((\pi_i, \lambda_i), i = 1, 2, \ldots, m\), and \((\nu_i, \alpha_i), i = 1, 2, \ldots, m\). Let \(\pi\) and \(\nu\) be the total portfolio as distributed over all stocks in all sectors. We have the following identities.

(i) Chain rule for relative entropy:

\[
H(\pi | \nu) = H(\lambda | \alpha) + \sum_{i=1}^{m} \lambda_i H(\pi_i | \nu_i).
\]

(ii) Chain rule for the free energy:

\[
\gamma^*_\pi(t) = \gamma^*_\pi,\text{sector}(t) + \sum_{i=1}^{m} \lambda_i(t) \gamma^*_\pi,\text{stock}(t).
\]

Here, \(\gamma^*_\pi\) is the total excess growth rate of \(\pi\) seen as a portfolio investing in all sectors and stocks. While \(\gamma^*_\pi,\text{stock}\) is the excess growth rate restricted to the \(i\)th sector; \(\gamma^*_\pi,\text{sector}\) is the excess growth rate where the basic assets are sectors themselves.

Hence, each of the relative entropy and the free energy of \(\pi\) over the entire universe of stocks equals the sum of the weighted averages of the corresponding quantities within the sectors, and the one coming from among the sectors.

**Proof.** The chain rule for relative entropy is a well-known fundamental result. For a proof see Theorem 2.5.3 on page 24 of [CT06]. Hence we only prove the chain rule for the free energy.

Recall that the growth rate of any portfolio \(\nu\) with dollar value \(\tilde{V}_\nu(\cdot)\) at time \(t\) is given by \(\Delta \log \tilde{V}_\nu(t)\). Also recall from Lemma 2.3 that the free energy does not depend on the numéraire, although the growth rate does. In the following argument we will use the dollar value as the numéraire.

Fix a time period \(t\) to \(t + 1\). This will be fixed throughout the argument and we will not explicitly refer to the time argument. Let \(\gamma^{\text{stock}}\) be the vector of the growth rates of all the stocks. Let \(\gamma_\pi\) be the growth rate of \(\pi\), and \(\gamma^\text{sector} = (\gamma_{\pi_1}, \ldots, \gamma_{\pi_m})\) be the vector of the sector growth rates. We now use the definition of the free energy from (3), repeatedly.
If we think of $\pi$ as a portfolio of the $n$ stocks, then we can write

\begin{equation}
\gamma_\pi = m \sum_{i=1}^{m} \sum_{j=1}^{n_i} \lambda_i \pi_{ij} \gamma_{\text{stock}}^{ij} + \gamma^*_{\pi}, \tag{25}
\end{equation}

where $\gamma^*_{\pi}$ is the free energy from the entire universe of stocks.

However we would like to think of $\pi$ as a mixture of sector portfolios, and each sector portfolio $\pi_i$ is a mixture of sector stocks. Hence, at the sector level we may also write

\begin{equation}
\gamma_\pi = \lambda_i \gamma_{\text{sector}}^{i} + \gamma^*_{\text{sector}}, \tag{26}
\end{equation}

Finally, for each sector portfolio, we have

\begin{equation}
\gamma_{\text{sector}}^{i} = \gamma_{\pi_i} = \sum_{j=1}^{n_i} \pi_{ij} \gamma_{\text{stock}}^{ij} + \gamma^*_{\text{stock}}. \tag{27}
\end{equation}

Let $a \cdot b$ denote the Euclidean inner product of two vectors $a$ and $b$. Putting (27) into (26), we get

\begin{align*}
\gamma_\pi &= \lambda_i \gamma_{\text{sector}}^{i} + \gamma^*_{\text{sector}} \\
&= \sum_{i=1}^{m} \lambda_i (\pi_i \cdot \gamma_{\text{stock}}^{i} + \gamma^*_{\text{stock}}) + \gamma_{\text{sector}} \\
&= \pi \cdot \gamma_{\text{stock}} + \sum_{i=1}^{m} \lambda_i \gamma^*_{\text{stock}} + \gamma^*_{\text{sector}}. \tag{28}
\end{align*}

Comparing (28) and (25), we have $\gamma^*_{\pi} = \sum_{i=1}^{m} \lambda_i \gamma^*_{\text{stock}}^{i} + \gamma^*_{\text{sector}}$. This completes the proof of the lemma. \qed

Formula (24) is a neat way of attributing the free energy of any portfolio across levels. For example, if the total free energy and individual sector free energies are known, one obtains the free energy from mixing the sectors by the simple linear formula.

We now ask the following natural question: if we run an energy-entropy portfolio within each sector and we run an energy-entropy portfolio among the sectors, is the total portfolio an energy-entropy portfolio? Lemma 3.1 gives precise conditions when this is indeed the case.

**Theorem 3.2.** Recall the notation from Lemma 3.1. Let $\pi = \sum_{i=1}^{m} \lambda_i \pi_i$ denote any portfolio and let $\mu = \sum_{i=1}^{m} \alpha_i \mu_i$ denote the market portfolio. Suppose each $\pi_i$ is an energy-entropy portfolio within sector $i$. Then $\pi$ is an energy-entropy portfolio over the entire universe of stocks if any of the following two conditions is satisfied.

(i) $\lambda$ is a constant-weighted portfolio.
(ii) \( \lambda \) is an energy-entropy portfolio that satisfies the following monotonicity condition for any pair of indices \((i, j)\):

\[
\frac{\lambda_i(t+1)}{\lambda_i(t)} \leq \frac{\lambda_j(t+1)}{\lambda_j(t)}, \quad \text{if } H(\pi_i(t+1) \mid \mu_i(t+1)) > H(\pi_j(t+1) \mid \mu_j(t+1)).
\]

(29)

Proof. We start at the decomposition (12). It suffices to consider the term \(D(\cdot)\) and show that it is increasing in time. Now

\[
\Delta D(t) = \gamma^{*}_\pi(t) + H(\pi(t+1) \mid \mu(t+1)) - H(\pi(t) \mid \mu(t+1)).
\]

(30)

We now use Lemma [3.1] to expand each term on the right side of the above equation.

\[
\gamma^{*}_\pi(t) = \gamma^{*}_\pi, \text{sector}(t) + \sum_{i=1}^{m} \lambda_i(t) \gamma^{*}_\pi, \text{stock}(t),
\]

\[
H(\pi(t+1) \mid \mu(t+1)) = H(\lambda(t+1) \mid \alpha(t+1)) + \sum_{i=1}^{m} \lambda_i(t+1) H(\pi_i(t+1) \mid \mu_i(t+1)),
\]

\[
H(\pi(t) \mid \mu(t+1)) = H(\lambda(t) \mid \alpha(t+1)) + \sum_{i=1}^{m} \lambda_i(t) H(\pi_i(t) \mid \mu_i(t+1)).
\]

When \( \lambda \) is constant-weighted, we have \( \lambda(t) = \lambda(t+1) \equiv \lambda \). This allows us to combine the three terms above and get

\[
\Delta D(t) = \Delta D^{\text{sector}}(t) + \sum_{i=1}^{m} \lambda_i \Delta D^{\text{stock}}_i(t),
\]

where the notations are self-explanatory. Since we have energy-entropy portfolios within each sector and a constant weighted portfolio among the sectors, each of the \( \Delta D(t) \) terms are non-negative. This shows that \( \Delta D(t) \geq 0 \) and proves that \( \pi \) is an energy-entropy portfolio.

In the general case

\[
\Delta D(t) = \Delta D^{\text{sector}}(t) + \sum_{i=1}^{m} \lambda_i(t+1) H(\pi_i(t+1) \mid \mu_i(t+1))
\]

\[
+ \sum_{i=1}^{m} \lambda_i(t) \left[ \gamma^{*}_\pi, \text{stock}(t) - H(\pi_i(t) \mid \mu_i(t+1)) \right]
\]

\[
= \Delta D^{\text{sector}}(t) + \sum_{i=1}^{m} \left[ \lambda_i(t+1) - \lambda_i(t) \right] H(\pi_i(t+1) \mid \mu_i(t+1))
\]

\[
+ \sum_{i=1}^{m} \lambda_i(t) \left[ H(\pi_i(t+1) \mid \mu_i(t+1)) + \gamma^{*}_\pi, \text{stock}(t) - H(\pi_i(t) \mid \mu_i(t+1)) \right].
\]
Because of our energy-entropy trading strategies within and among the sectors, the first and the third term in the final expression above is nonnegative. We now show how to control the middle term.

Consider a random integer \( I \) that takes value \( i \) with probability \( \lambda_i(t) \), for \( i = 1, 2, \ldots, m \). Consider two functions:

\[
 f(i) = \frac{\lambda_i(t+1)}{\lambda_i(t)}, \quad g(i) = H(\pi_i(t+1) | \mu_i(t+1)).
\]

Obviously \( Ef(I) = 1 \) and, hence

\[
 \sum_{i=1}^{m} [\lambda_i(t+1) - \lambda_i(t)] H(\pi_i(t+1) | \mu_i(t+1)) = \text{Cov}(f(I), g(I)).
\]

If \((I, I')\) are i.i.d. random variables, the above covariance is given by the symmetrized expression

\[
 \text{Cov}(f(I), g(I)) = \frac{1}{2} E \left[ (f(I) - f(I')) (g(I) - g(I')) \right].
\]

Thus, under the assumed monotonicity condition (29), for any pair of values of \((I, I')\) we must have

\[
 (f(I) - f(I')) (g(I) - g(I')) \geq 0.
\]

This, in turn, implies that \( \text{Cov}(f(I), g(I)) \geq 0 \). Combining everything we get \( \Delta D(t) \geq 0 \) and hence our result is proved. \( \square \)

**Remark 3.3.** Functionally generated portfolios are typically not amenable to such mixing. See, however, Example 11.1 and Example 4 below Remark 11.1 in the survey [PK09].

### 4. A binary tree model: Rebalancing between two assets

In this section we give a different picture of rebalancing for a two asset portfolio. The meaning of excess growth rate and its relation with functionally generated portfolios will be made clear.

#### 4.1. Discrete time model

We begin with a discrete model where again randomness plays no role. Suppose the two assets have nominal prices \( \tilde{X}_1(t) \) and \( \tilde{X}_2(t) \) at times \( t \in \{0, 1, 2, \ldots\} \). Apart from attempting to beat the two-asset market, we also ask the following natural question: is it more profitable to invest fully in asset 2, or should we invest in a constant-weighted portfolio between the two assets? For example, asset 2 could be a portfolio itself, and we decide on whether to include a new stock in the mix.

To answer these let \( \pi \) be a portfolio and \( \tilde{V}(t) \) be the value of the portfolio with an initial investment of $1. We use \( \tilde{X}_2(t) \) as the numéraire and define the relative prices by

\[
 X_1(t) = \frac{\tilde{X}_1(t)}{\tilde{X}_2(t)}, \quad X_2(t) = \frac{\tilde{X}_2(t)}{\tilde{X}_2(t)} \equiv 1, \quad \text{and} \quad V(t) = \frac{\tilde{V}_1(t)}{\tilde{X}_2(t)}.
\]
Since $X_2(t) \equiv 1$, we get

\begin{equation}
\frac{\Delta V(t)}{V(t)} = \pi_1 \frac{\Delta X_1(t)}{X_1(t)} + \pi_2 \frac{\Delta X_2(t)}{X_2(t)} = \pi_1 \frac{\Delta X_1(t)}{X_1(t)},
\end{equation}

where for any process $A$ we let $\Delta A(t) = A(t + 1) - A(t)$ as before.

Define the log-relative price

$$Y(t) = \log X_1(t) = \log \left( \frac{\tilde{X}_1(t)}{X_2(t)} \right).$$

Notice that $Y$ can be seen as a measure of diversity between the two assets.

Let $\tilde{S}(t) = X_1(t) + X_2(t)$ be the (relative) capitalization of the market. Clearly $\tilde{S}(t) = e^{Y(t)} + 1$. Hence, the (relative) value of the market portfolio at time $t$, with an initial investment of $\$1$, is given by $S(t) = \tilde{S}(t)/\tilde{S}(0)$.

We assume that $\Delta Y(t)$ only takes the values $\sigma$ and $-\sigma$, where $\sigma > 0$ is a fixed constant. Hence $Y$ follows a binary tree model where $\sigma$ is the step size.

It is clear, then, that $\sigma^2$ is the instantaneous volatility of the relative prices. In fact, if $\sigma_i$ is the volatility of asset $i$ and $\rho$ is the correlation between them, one can easily derive the formula

$$\sigma^2 = \sigma_1^2 + \sigma_2^2 - 2\rho \sigma_1 \sigma_2.$$

Thus, $\sigma$ increases as the correlation decreases to $-1$.

Consider a portfolio $\pi(t) = (q(t), 1 - q(t))$. From (31) we get

$$\frac{\Delta V(t)}{V(t)} = q(t) \frac{\Delta X_1(t)}{X_1(t)}$$

$$= q(t) \left( e^{Y(t) + \Delta Y(t)} - e^{Y(t)} \right) = q(t) \left( e^{\Delta Y(t)} - 1 \right).$$

It follows that

\begin{equation}
\frac{V(t + 1)}{V(t)} = 1 + q(t) \left( e^{\Delta Y(t)} - 1 \right) = \begin{cases} 
1 + q(t) (e^\sigma - 1), & \text{if } \Delta Y(t) = \sigma, \\
1 + q(t) (e^{-\sigma} - 1), & \text{if } \Delta Y(t) = -\sigma.
\end{cases}
\end{equation}

For any time $T$, we consider $V(T)/V(0)$ as a product of quantities as above.

We first analyze the performance of a constant-weighted portfolio $\pi = (q, 1 - q)$, where $0 < q < 1$. When $q = 1/2$ we recover the equal-weighted portfolio.

Clearly one can rearrange the order in which the factors coming from (32) get multiplied and get the same value $V(T)/V(0)$. We will do this rearrangement in a convenient way whereby one ‘up move’ of $Y$ from $k\sigma$ to $(k + 1)\sigma$ is matched with one ‘down move’ from $(k + 1)\sigma$ to $k\sigma$. Multiplying these two factors, we see that the contribution of a match at level $k$ is equal
to
\[(1 + q(e^\sigma - 1))(1 + q(e^{-\sigma} - 1)) = 1 + q(e^\sigma + e^{-\sigma} - 2)\]
\[+ q^2(e^\sigma - 1)(e^{-\sigma} - 1)\]
\[(33)\]
\[= 1 + q \left(\frac{e^{\sigma/2} - e^{-\sigma/2}}{e^\sigma - e^{-\sigma}}\right)^2 - q^2 \left(\frac{e^{\sigma/2} - e^{-\sigma/2}}{e^\sigma - e^{-\sigma}}\right)^2\]
\[= 1 + q(1 - q) \left(\frac{e^{\sigma/2} - e^{-\sigma/2}}{e^\sigma - e^{-\sigma}}\right)^2 > 1.\]

This is the exact gain from buying low when \(Y = k\sigma\) and selling high when \(Y = (k + 1)\sigma\).

Suppose there are \(N\) matches up to time \(T\). From Figure 2, it is clear that the number of unmatched moves equals exactly \(|Y(T) - Y(0)|/\sigma\). Since \(V(0) = 1\), we have the following decomposition:

\[\log V(T) = N \log \left(1 + q(1 - q) \left(\frac{e^{\sigma/2} - e^{-\sigma/2}}{e^\sigma - e^{-\sigma}}\right)^2\right) + \frac{|Y(T) - Y(0)|}{\sigma} \log F,\]

where \(F = 1 + q(e^{\pm\sigma} - 1)\) depending on the sign of \(Y(T) - Y(0)\). Finally, the performance relative to the market is

\[\log \frac{V(T)}{S(T)} = N \log \left(1 + q(1 - q) \left(\frac{e^{\sigma/2} - e^{-\sigma/2}}{e^\sigma - e^{-\sigma}}\right)^2\right)\]
\[+ \frac{|Y(T) - Y(0)|}{\sigma} \log F - \log \frac{e^{Y(t)} + 1}{e^{Y(0)} + 1}.\]

The first term of (34) is the accumulated gain from matchings; the last two terms capture the change in market concentration and represent the risk for the trader. From this decomposition and the one above, it is clear that the constant-weighted portfolio beats the market (or asset 2) if the number of matches is large relative to that of the unmatched. When there are plenty of matches, the constant-weighted portfolio will outperform even if concentration builds up slowly; see Section 6.2 for a numerical example.
Next consider a portfolio which is a function of $Y$. Then the gain or loss from a match depends on its level. We suppose $q : \sigma Z \to [0, 1]$ is a state dependent weight. Equation (32) continues to hold by replacing $q$ by $q(y)$ on the event $\{Y(t) = y\}$. However, now the pairing of up and down moves is slightly different. The contribution of any match straddled between height $k\sigma$ and $(k + 1)\sigma$ is

\[(35) \quad [1 + q(k\sigma)(e^\sigma - 1)] [1 + q((k + 1)\sigma)(e^{-\sigma} - 1)].\]

**Definition 4.1.** We will call $\pi = (q, 1 - q)$ a greedy-rebalancing portfolio if $q$ is a decreasing function of $Y = \log \tilde{X}_1(t)/\tilde{X}_2(t)$.

If $q$ is greedy-rebalancing, and hence decreasing in its argument, the contribution in (35) must be at least

\[ [1 + q(k\sigma)(e^\sigma - 1)] [1 + q(k\sigma)(e^{-\sigma} - 1)] > 1.\]

A greedy-rebalancing portfolio underweights more and more the outperforming asset, and is probably too aggressive for most investors. We begin with these portfolios because it is easy to see the effect of matching. In fact, we require much less to gain a premium from matching. To see this, write $q_k = q(k\sigma)$ and note that (35) equals

\[ (35) \quad [1 + q_k(e^\sigma - 1)] [1 + q_k(e^{-\sigma} - 1)] = 1 + q_k(1 - q_k) \left( e^{\frac{\sigma}{2}} - e^{-\frac{\sigma}{2}} \right)^2 - \Delta q_k(1 - e^{-\sigma}) - q_k\Delta q_k \left( e^{\frac{\sigma}{2}} - e^{-\frac{\sigma}{2}} \right)^2.\]

For this to be greater than 1, we need

\[ q_k(1 - q_k) \left( e^{\frac{\sigma}{2}} - e^{-\frac{\sigma}{2}} \right)^2 \geq \Delta q_k(1 - e^{-\sigma}) + q_k\Delta q_k \left( e^{\frac{\sigma}{2}} - e^{-\frac{\sigma}{2}} \right)^2.\]

Using Taylor approximation, for small $\sigma$ this inequality is approximately equivalent to

\[(36) \quad q_k(1 - q_k) \geq \frac{\Delta q_k}{\sigma} + q_k\Delta q_k.\]

Letting $\sigma \to 0$ and assuming $q$ is differentiable, we get the fundamental (continuous-time) inequality

\[(37) \quad q' \leq q(1 - q).\]

Note that if $q(y) = \frac{e^y}{1 + e^y}$ (i.e., when $q$ is the market portfolio), then $q' = q(1 - q)$ and equality holds. As long as $q$ increases at a rate slower than the market portfolio in the sense of (36) (discrete time) or (37) (continuous time), the portfolio will outperform the market over an excursion of $Y$, i.e., a time interval over which $Y$ reverts to the beginning value for the first time. The relation between (37) and functionally generated portfolios will be made clear in Section 4.2.
4.2. Extensions to continuous time. All ideas in the discrete time model extend immediately to continuous time. Let $X_1 = e^Y$ and $X_2 \equiv 1$ be the relative prices. We assume that $Y$ is a continuous semimartingale with respect to the underlying filtration. Let $L^y_t$ be the local time of $Y$ at $y$ up to time $t$ (see Chapter VI of the book by Revuz and Yor [RY99] for an introduction). We will make use of the celebrated occupation time formula:

$$
\int_0^t \Phi(Y_s)d\langle Y \rangle_s = \int_{-\infty}^\infty \Phi(y)L^y_t dy.
$$

Consider a portfolio $\pi(t) = (\pi_1(t), \pi_2(t)) = (q(t), 1 - q(t))$ with initial value $V(0) = 1$. The relative value $V(t)$ of the portfolio satisfies

$$
dV(t) = q(t) dX_1(t) / X_1(t).
$$

By Itô’s formula, we have

$$
d\log V(t) = q(t) \left( dY(t) + \frac{1}{2} d\langle Y \rangle_t \right) - \frac{q^2(t)}{2} d\langle Y \rangle_t
$$

$$
= q(t) dY(t) + \frac{1}{2} q(t)(1 - q(t)) d\langle Y \rangle_t.
$$

Hence, for any $t$ we have the following analogue of (34):

$$
\log V(t) = \int_0^t q(s) dY(s) + \frac{1}{2} \int_0^t q(s)(1 - q(s)) d\langle Y \rangle_s.
$$

This is equivalent to the usual logarithmic representation (see (1.11) of [FK09]), but here we think of the portfolio value as driven by the fluctuations of the relative price $Y$.

In fact, it is easy to show that the excess growth rate $\gamma^*_\pi(t)$ of the portfolio (in the sense of Fernholz) satisfies

$$
\gamma^*_\pi(t) dt = \frac{1}{2} q(t)(1 - q(t)) d\langle Y \rangle_t.
$$

Comparing (40) with (33), we see that for a two-asset portfolio, the excess growth rate can be thought of as the infinitesimal premium from matching up and down moves, which happens instantaneously in continuous time. If $\pi$ is constant-weighted the analogy is exact. When $q$ is a smooth function of $Y$, there is a correction term as shown in Proposition 4.3(ii).

As we have seen in the discrete time case, for constant-weighted portfolios the situation is clear.

**Proposition 4.2.** Consider a constant-weighted portfolio $\pi(t) = (q, 1 - q)$. Suppose that $[0, \tau]$ is an excursion of $Y$. Then

$$
\log \frac{V(\tau)}{V(0)} = \frac{1}{2} \int_0^\tau q(1 - q) d\langle Y \rangle_t = \frac{1}{2} q(1 - q) \langle Y \rangle_\tau.
$$
This quantity is maximized path by path when \( q = \frac{1}{2} \), i.e., when the portfolio is equally weighted.

**Proof.** Integrate (39) over \([0, \tau]\) and maximize over \( q \in (0, 1) \). \( \square \)

Next we consider the case of a portfolio depending continuously on \( Y(t) \).

**Proposition 4.3.**

(i) Suppose that \( q : \mathbb{R} \to [0, 1] \) is continuous and has finite variation. Let \( \pi \) be the portfolio defined by \( \pi(t) = (q(Y(t)), 1 - q(Y(t))) \) and \( F \) be an antiderivative of \( q \). Then for \( t \geq 0 \),

\[
\log \frac{V(t)}{V(0)} = F(Y(t)) - F(Y(0)) + \frac{1}{2} \left( \int_{-\infty}^{\infty} q(y)(1 - q(y)) L_y^t dy + \int_{-\infty}^{\infty} L_y^t d(-q(y)) \right),
\]

whenever the last (Lebesgue-Stieltjes) integral exists.

(ii) Suppose \( \pi(t) = (q(Y(t)), 1 - q(Y(t))) \) where \( q : [0, \infty) \to [0, 1] \) is continuously differentiable. Again let \( F \) be an antiderivative of \( q \). Then

\[
\log \frac{V(t)}{V(0)} = F(Y(t)) - F(Y(0)) + \frac{1}{2} \int_{-\infty}^{\infty} (-q'(y) + q(y)(1 - q(y))) L_y^t dy.
\]

**Proof.** To use (39), we need to evaluate the stochastic integral

\[
\int_0^t q(Y(s))dY(s).
\]

Since \( q \) has finite variation, we may write \( q = q_1 - q_2 \), where \( q_1 \) and \( q_2 \) are continuous and non-increasing. By linearity, it suffices to consider the case where \( q \) is non-increasing. Let \( F \) be an antiderivative of \( q \). Then \( F \) is concave since \( F'(y) = q(y) \) is is non-increasing. By the Itô-Tanaka formula, for all \( t \geq 0 \) we have

\[
F(Y(t)) = F(Y(0)) + \int_0^t q(Y(t))dY(t) + \frac{1}{2} \int_0^\infty L_y^t f''(dy),
\]

where \( f''(dy) = dq(y) \) is the second derivative of \( f \) in the sense of distribution. Then (i) is proved by substituting this into (39) and applying the occupation time formula. If \( q \) is continuously differentiable, as in (ii), we can apply Itô’s formula directly. \( \square \)

**Definition 4.4.** Let \( \pi(t) = (q(Y(t)), 1 - q(Y(t))) \) be a portfolio which depends only on the current market weights. We say that \( \pi \) is a reversion portfolio if for any continuous semimartingale \( Y \), \( V(\tau) \geq S(\tau) \) whenever \( Y(\tau) = Y(0) \).
Corollary 4.5. Let $\pi(t) = (q(Y(t)), 1 - q(Y(t)))$ be a portfolio where $q$ is continuously differentiable. Then $\pi$ is a reversion portfolio if and only if $q'(y) \leq q(y)(1 - q(y))$ for all $y \in \mathbb{R}$.

Proof. By Proposition 4.3, it is clear that if $q' \leq q(1 - q)$ then $\pi$ is a reversion portfolio. Suppose $q'(y_0) > q(y_0)(1 - q(y_0))$ for some $y_0 \in \mathbb{R}$. By continuity, $-q' + q(1 - q) < 0$ on some open interval $I$ containing $y_0$. We can then construct a semimartingale $Y$ which accumulates a lot of local time in $I$ before reverting to $Y(0)$. Then

$$\log \frac{V(\tau)}{V(0)} = \frac{1}{2} \int_{-\infty}^{\infty} (-q'(y) + q(y)(1 - q(y))) L_y^2 dy < 0.$$ 

□

It is well known that most portfolios, even if they are smooth functions of market weights, are not functionally generated. Nevertheless, it is worth pointing out that for a two-asset market, all such portfolios are functionally generated.

Lemma 4.6. Any portfolio $\pi = (q(y), 1 - q(y))$, where $q$ is a continuously differentiable function of $y$, is functionally generated. A generating function (unique up to a multiplicative constant) is

$$S(\mu_1, \mu_2) = \exp \left( F \left( \log \frac{\mu_1}{\mu_2} \right) - \log \frac{1}{\mu_2} \right),$$

where $F$ is an antiderivative of $q$.

Proof. Let $S = e^G$, where $G$ is a differentiable function of $y = \log \frac{\mu_1}{\mu_2}$. By (11.1) of [FK09], $S$ generates the portfolio $(\pi_1, \pi_2)$ where

$$\frac{\pi_1}{\mu_1} = D_1 \log S + 1 - \mu_1 D_1 \log S - \mu_2 D_2 \log S = \frac{G'(y) + \mu_1}{\mu_1}.$$

To generate the weights $(q, 1 - q)$, we require that

$$q(y) = G'(y) + \mu_1 = G'(y) + \frac{e^y}{1 + e^y}.$$ 

We may then pick $G(y) = F(y) - \int \frac{e^y}{1 + e^y} dy - F(y) - \log(1 + e^y).$ □

We remark that although $\pi$ is functionally generated, the function $S$ has no special interpretation other than being a function of the market weights. The following corollary completes the circle of ideas.

Corollary 4.7. Let $\pi(t) = (q(Y(t)), 1 - q(Y(t)))$ be a portfolio generated by a twice continuously differentiable function $S$. Then the inequality

$$q'(y) \leq q(y)(1 - q(y))$$

holds if and only if $S$ is concave. Thus $\pi$ is a reversion portfolio if and only if it is generated by a concave function.

Proof. Differentiate (12). □
5. Optimization over functionally generated portfolios

Suppose we hold a portfolio \( \pi(t) = (q(Y(t)), 1 - q(Y(t))) \) satisfying the conditions of Proposition 4.3(i). If we are confident that \( Y \) will fluctuate about a certain value, a natural investment strategy is to maximize the expected gain after we accumulate a certain amount of excursions about that value. Moreover, if we have statistical information about the dynamics of \( Y \), we can optimize the weight function \( q \) subject to various constraints. Mathematically, this amounts to picking a weight function \( w(y) \) representing the expected local time at each value of \( y \).

For concreteness, we consider excursions of \( Y \) about zero. If we model \( \{Y(t)\} \) as a one-dimensional diffusion process, Lemma 6.1 in the Appendix justifies consideration of the following variational problem:

\[
\sup_q \Lambda(q) = \sup_q \int_{-\infty}^{\infty} q(y)(1 - q(y))w(y)dy + \int_{-\infty}^{\infty} w(y)d(-q(y)).
\]

Here \( w : \mathbb{R} \to [0, \infty) \) is a given weight function and \( q : \mathbb{R} \to [0, 1] \) is continuous, has finite variation, and \( q(0) = \frac{1}{2} \) (\( \pi \) is equal-weighted whenever the market is). Additional constraints will be specified below. We assume that \( w \) is continuous and integrable: \( \int_{-\infty}^{\infty} w(y)dy < \infty \). As a benchmark we consider the equal-weighted portfolio where \( q(y) \equiv \frac{1}{2} \).

If \( w \) is continuously differentiable on \([0, \infty) \) and \((-\infty, 0]\), and \( \lim_{|y| \to \infty} w(y) = 0 \), integrating by parts gives

\[
\int_{-\infty}^{\infty} w(y)d(-q(y)) = 2w(0)q(0) + \int_{-\infty}^{\infty} q(y)w'(y)dy.
\]

Hence

\[
\Lambda(q) = 2w(0)q(0) + \int_{-\infty}^{\infty} (q(y)(1 - q(y))w(y) + q(y)w'(y))dy.
\]

We consider two special cases which are of particular interest.

(i) \( w(y) = e^{-\gamma|y|} \), where \( \gamma > 0 \). This is the case if \( Y \) is a \textit{Bang-bang process}. Then

\[
\Lambda(q) = 2q(0) + \int_{0}^{\infty} q(y)((1 - \gamma) - q(y))e^{-\gamma|y|}dy \\
+ \int_{-\infty}^{0} q(y)((1 + \gamma) - q(y))e^{-\gamma|y|}dy.
\]

For the equal-weighted portfolio we have \( \Lambda \left( \frac{1}{2} \right) = \frac{1}{2\gamma} \). For fixed \( y \), the integrand is a quadratic function in \( q(y) \), maximized when

\[
q(y) = \begin{cases} 
\frac{1}{2}(1 - \gamma)+, & \text{for } y > 0, \\
\left(\frac{1}{2}(1 + \gamma)\right) \wedge 1 = 1 - \frac{1}{2}(1 - \gamma)+, & \text{for } y < 0.
\end{cases}
\]
We use here the notation $x_+ = \max\{x, 0\}$. It follows that
\[
\sup_q \Lambda(q) = \begin{cases} 
\frac{(1+\gamma)^2}{2}, & \text{if } 0 \leq \gamma < 1 \\
2, & \text{if } \gamma \geq 1
\end{cases}
\]
and there is no optimal solution since we require $q$ to be continuous and $q(0) = \frac{1}{2}$. The supremum can be achieved asymptotically by interpolating $q$ continuously between $\frac{1}{2}$ and the optimal weights.

(ii) $w(y) = e^{-\gamma y^2}$, where $\gamma > 0$. This corresponds to the case where $Y$ is an Ornstein-Uhlenbeck process. Then
\[
\Lambda(q) = 2q(0) + \int_{-\infty}^{\infty} q(y)(1 - 2\gamma y - q(y))e^{-\gamma y^2} \, dy.
\]
In this case
\[
\Lambda\left(\frac{1}{2}\right) = 1 + \frac{\sqrt{\pi}}{4\sqrt{\gamma}}.
\]
For each $y$, the integrand is maximized when
\[
q(y) = \begin{cases} 
0, & \text{if } y > \frac{1}{2\pi} \\
\frac{1}{2}(1 - 2\gamma y), & \text{if } |y| \leq \frac{1}{2\pi} \\
1, & \text{if } y < -\frac{1}{2\pi}
\end{cases}
\]
Since this is continuous and $q(0) = \frac{1}{2}$, it is the optimal solution. It can be shown that the optimal value is
\[
\sup_q \Lambda(q) = 1 + \frac{1}{2} e^{-\frac{1}{\pi}} + \left(\frac{\sqrt{\pi}}{4\sqrt{\gamma}} + \frac{\sqrt{\pi a}}{2}\right) \text{erf}\left(\frac{1}{2\sqrt{a}}\right),
\]
where erf is the standard error function.

In general, suppose $w(y) = e^{\varphi(y)}$, where $\varphi$ is continuously differentiable. Then
\[
\Lambda(q) = \int_{-\infty}^{\infty} q(y)(1 + \varphi'(y) - q(y))e^{\varphi(y)} \, dy,
\]
and for each $y$ the optimal weight is $q(y) = \left(\frac{1}{2}(1 + \varphi'(y)) \vee 0\right) \wedge 1$ (since $\varphi$ is $C^1$, this is a continuous function of $y$). Additional constraints on $q(y)$ can be handled easily. For example, suppose we impose the constraint $q(y) \geq \delta$, i.e., we cannot underweight asset 1 too much. Then, for $y > 0$, the optimal solution for (i) is $q(y) = \frac{1}{2}(1 - \gamma)_+ \vee \delta$ (asymptotically), and for (ii), it is $q(y) = \frac{1}{2}(1 - 2\gamma y)_+ \vee \delta$.

As another example, suppose we impose the condition that the weight ratios are bounded above and below, i.e., for fixed constants $0 < A < 1 < B$ we require
\[
A \leq \frac{\pi_i}{\mu_i} \leq B, \quad i = 1, 2.
\]
Since $\mu_1 = \frac{e^y}{1+e^y}$, for $y > 0$ this is equivalent to
\[
q(y) \geq m(y) := \max\left\{ \frac{Ae^y}{1+e^y}, 1 - \frac{B}{1+e^y} \right\}.
\]
These constraints are adopted in practice to control tracking error and liquidity risk. Now the (asymptotic) optimal solution for (i) is
\[ q(y) = \frac{1}{2}(1 - \gamma) + m(y), \quad y > 0. \]
For (ii), the optimal solution is
\[ q(y) = \frac{1}{2}(1 - 2\gamma y) + m(y), \quad y > 0. \]

Remark 5.1. If \( Y \) is a diffusion, the portfolio \( (q, 1 - q) \) we found maximizes the growth rate of \( V \) subject to the weight constraints. To see this, suppose that \( Y \) satisfies the SDE
\[ dY(t) = d\log X_1(t) = a(t)dt + b(t)d\beta(t), \]
where \( \beta(t) \) is a Brownian motion. If we hold the portfolio \( \pi(t) = (q(t), 1 - q(t)) \), the growth rate of \( V \) is
\[ q(t)a(t) + \frac{1}{2}q(t)(1 - q(t))b^2(t). \]
This is maximized when
\[ q(t) = \frac{a(t) + \frac{1}{2}b^2(t)}{b^2(t)} = \frac{1}{2} \left( 1 + \frac{2a(t)}{b^2(t)} \right). \]
This is the same as the solution derived above without constraints.

6. **Empirical examples**

6.1. **A simple energy-entropy strategy.** As we have seen in Section 2, there is a lot of freedom in implementing energy-entropy portfolios. For the sake of illustration we consider the strategy corresponding to the vector field described in Example (iv) below Condition 2.9.

We define a family of strategies called “\( \lambda \)-strategy” indexed by a parameter \( \lambda \in [0, 1] \). The idea is the following. Suppose we hold \( \pi(t) \) from time \( t \) to time \( t + 1 \). At time \( t + 1 \), we observe \( \mu(t+1) \) and the free energy \( \gamma^* \) is then known. We move towards \( \mu(t+1) \) so that \( \pi(t+1) \) is a convex combination of \( \pi(t) \) and \( \mu(t+1) \), and the amount is chosen so that we only consume \( \lambda \) fraction of \( \gamma^* \).

The geometry of the strategy is illustrated in Figure 3. At time \( t + 1 \) before rebalancing, the portfolio weights have drifted to some new weights \( \tilde{\pi}(t+1) \). Then we rebalance to a convex combination of \( \pi(t) \) and \( \mu(t+1) \) chosen so that the control term is \( -\lambda \gamma^* \). As formulated in (16), the \( \lambda \)-strategy uses the vector field \( U_\mu(\pi) = \pi - \mu \) and stops the reverse flow at time \( (\lambda \gamma^*) \) when energy consumption equals \( \lambda \gamma^* \) or when the market portfolio is reached.

It is clear by construction that the portfolio \( \pi \) is an energy-entropy portfolio. For convenience, we use linear approximation and make adjustment so that \( \pi(t+1) \) is always a convex combination of \( \pi(t) \) and \( \mu(t+1) \) (and hence stays inside \( G_n \)). Explicitly, we construct the portfolio as follows.
(i) Fix any desirable starting weights \( \pi(0) \).

(ii) Suppose \( \pi(t) \) is chosen at time \( t \). When \( \mu(t + 1) \) is revealed, the free energy \( \gamma_\pi^*(t) \) can be computed. Then we define

\[
\pi(t + 1) = \pi(t) + s(\mu(t + 1) - \mu(t)),
\]

where

\[
s := \frac{\lambda \gamma_\pi^*(t) / |\nabla H(\pi(t) | \mu(t + 1)) \cdot v|}{|\mu(t + 1) - \pi(t)|} \land 1
\]

and \( v = \frac{\mu(t+1)-\pi(t)}{|\mu(t+1)-\pi(t)|} \).

Note that if \( \lambda = 0 \), then \( \pi(t + 1) = \pi(t) \) and \( \pi \) is a constant-weighted portfolio. If \( \lambda = 1 \) and \( \pi(0) = \mu(0) \), then (using exact calculations) \( \pi \equiv \mu \) is the market portfolio.

As a first example we consider the monthly stock prices of Apple and Starbucks from January 1994 to April 2012. The market consists of these two stocks and we normalize so that they are equally weighted in January 1994. This data set is also studied in Bouchey et al. [BNPS12]. Now we simulate the performance of \( \pi \) with \( \lambda = 0.3 \) relative to the market. We let \( \pi(0) = \mu(0) = (0.5, 0.5) \) be the starting weights and use monthly time steps. Figure 4 plots the energy-entropy decomposition

\[
\log V(t) = D(t) + H(\pi(0) | \mu(0)) - H(\pi(t) | \mu(t))
\]
as a function of time. Since \( \pi \) is an energy-entropy portfolio, the drift process is increasing by construction. From the figure, it is clear that the drift drives the long term outperformance of the portfolio. On the right we also plot the weight of Starbucks. We see that the portfolio moves towards the market slowly (approximating a finite variation process); it adjusts more rapidly when the market is volatile, i.e., when the free energy is large. The portfolio underperforms when one of the stocks dominates the market.

Next we turn to a more realistic example, where we consider monthly country returns (in US dollars) of 18 emerging countries from January 2001
to March 2013. More precisely, for each country we pick a country index, and the returns of that index are taken to be the country returns. The data is extracted from Factset. The market consists of these countries, where the starting market weights are proportional to the total capitalizations of the indices. For example, the beginning market weights of Brazil, Chile and China are respectively 0.138, 0.044 and 0.073.

Again we simulate the performance of the $\lambda$-strategy with $\lambda = 0.3$. We consider two cases where $\pi(0) = \mu(0)$ and $\pi(0) = (1/n, \ldots, 1/n)$. Figure 5 plots the energy-entropy decompositions of the two portfolios. In both cases $\pi$ outperforms the market and the drift process has a steady increasing trend. We note that although we use the same update rule in both cases, the terms represented by $H(\pi(0)|\mu(0)) - H(\pi(t)|\mu(t))$ are quite different because the portfolio depends on the entire history of portfolio and market weights.

6.2. Counting matched and unmatched edges. We consider again the Apple-Starbucks data in the context of Section 4. We let Apple be the numéraire ($X_2$) and consider the relative price $Y = \log X_1$. For simplicity, we discretize the series of $Y$ so that the step size is $\sigma = 0.1$. Figure 6 plots the series as well as the numbers of matched pairs and unmatched moves.
\[ \pi(0) = \mu(0) \]

\[ \pi(0) = (1/n, \ldots, 1/n) \]

**Figure 5.** Performance of $\pi$ relative to the hypothetical market of emerging countries. On the left the portfolio begins at the market weights, and on the right the portfolio is equal weighted initially.

**Appendix: Expected Local Time of One-Dimensional Diffusions**

Consider a one-dimensional diffusion

\[ dY(t) = b(Y(t))\,dt + \sigma(Y(t))\,d\beta(t), \]

where $b : \mathbb{R} \to \mathbb{R}$ is the drift function and $\sigma : \mathbb{R} \to (0, \infty)$ is the diffusion coefficient. These are assumed to be locally bounded. Define the scale function $s$ through its derivative

\[ s'(y) := \exp \left[ -\int_c^y \frac{2b(\xi)}{\sigma^2(\xi)}\,d\xi \right], \]

where $c$ is some arbitrary point on $\mathbb{R}$. It is then well known that $M(t) := s(Y(t))$ is a local martingale. The following lemma must be known in the literature but we cannot find a reference.

**Lemma 6.1.** Suppose $Y(0) = 0$. Let $L^y_Y$ denote the total local time of $Y$ accrued at $y$ until the first time its local time at zero hits one. We assume that the latter stopping time is finite almost surely. Then

\[ EL^y_Y = \frac{s'(0)}{s'(y)} = \exp \left[ \int_0^y \frac{2b(\xi)}{\sigma^2(\xi)}\,d\xi \right]. \]
Proof. Consider the Tanaka formula for $M$:

$$(45) \quad (M(t) - a)_+ = (M(0) - a)_+ + \int_0^t 1_{\{M(s) > a\}} dM(s) + \frac{1}{2} L^a_M(t),$$

where the final term is the local time of the semimartingale $M$ at a point $a$. The normalization of local time is chosen to keep it consistent with the occupation time formula (38).

Putting $a = s(0)$ in (45) and subtracting from a general $a$, we get

$$(M(t) - a)_+ - (M(t) - s(0))_+ = (M(0) - a)_+ - (M(0) - s(0))_+$$

$$+ N(t) + \frac{1}{2} L^a_M(t) - \frac{1}{2} L^{s(0)}_M(t),$$

where $N$ is a local martingale.

Assume without loss of generality $a > s(0)$, and let $M(0) = s(0)$. Let $\tau$ be the first time that the local time at $s(0)$ is $s'(0)$. Choose a sequence of localizing stopping times ($\sigma_k$, $k \geq 1$) for $N$ and apply the Optional Sampling Theorem to the resulting uniformly integrable stopped local martingales:

$$E (M(\tau \wedge \sigma_k) - a)_+ - E (M(\tau \wedge \sigma_k) - s(0))_+$$

$$= \frac{1}{2} EL^{a}_M (\tau \wedge \sigma_k) - \frac{1}{2} EL^{s(0)}_M (\tau \wedge \sigma_k).$$

The expression on the left is the expectation of a bounded function, and hence we are allowed to apply the Dominated Convergence Theorem as $k$ tends to infinity. On the right we have a difference in expectations of two
increasing functions and we are allowed to use the Monotone Convergence Theorem on each of them. Thus, letting \( k \) tend to infinity we get
\[
(46) \quad EL^n_M(\tau) = EL^{s(0)}_M(\tau) = s'(0).
\]

We now relate the expected local times of \( Y \) at \( y \) to that of \( M \) at \( s(y) \). The easiest way is to apply the occupation time formula (38) evaluated at the first time the local time of \( Y \) at zero hits one. As will be seen from the argument this time is exactly \( \tau \). We suppress the notation for time below for clarity. For any compactly supported bounded Borel function \( \Phi \) we get
\[
\int_{-\infty}^{\infty} \Phi(a) L^a_M \, da = \int_0^\tau \Phi(M(s)) \, d\langle M \rangle_s
\]
\[
= \int_0^\tau \Phi \circ s(Y(s)) (s'(Y(s))^2 \, d\langle Y \rangle_s
\]
\[
= \int_{-\infty}^{\infty} \Phi \circ s(y) (s'(y))^2 \, L^y_Y \, dy.
\]

By a change of variable to the above expression we get \( L^s_M(y) = s'(y)L^y_Y \).

Taking expectations on both sides at \( \tau \) and using (46) proves the lemma. \( \square \)

**Examples.** Consider the Bang-bang process
\[
dY(t) = -\alpha \text{sgn}(Y(t)) \, dt + \sigma d\beta(t).
\]
Here \( b(y) = -\text{sgn}(y) \) and \( \sigma(y) \equiv \sigma > 0 \). Hence, the expected local time at \( y \) is
\[
\exp \left[ -2 \frac{\alpha}{\sigma^2} \int_0^y \text{sgn}(\xi) d\xi \right] = \exp \left( -2\frac{\alpha |y|}{\sigma^2} \right).
\]

On the other hand if \( Y \) is an Ornstein-Uhlenbeck process
\[
dY(t) = -\alpha Y(t) dt + \sigma d\beta(t),
\]
we have \( b(y) = -\alpha y \) and \( \sigma(y) = \sigma \). Thus, the expected local time at \( y \) is
\[
\exp \left[ -2 \frac{\alpha}{\sigma^2} \int_0^y \xi d\xi \right] = \exp \left[ -\frac{\alpha y^2}{\sigma^2} \right].
\]

**References**

[BNPS12] P. Bouchey, V. Nemtchinov, A. Paulsen, and D. M. Stein, Volatility harvesting: Why does diversifying and rebalancing create portfolio growth?, Journal of Wealth Management 15 (2012), no. 2, 26–35.

[CK06] L. B. Chincarini and D. Kim, Quantitative equity portfolio management: An active approach to portfolio construction and management, McGraw-Hill Library of Investment and Finance, McGraw-Hill, 2006.

[Cov91] T. M. Cover, Universal portfolios, Mathematical Finance 1 (1991), no. 1, 1–29.

[CR11] L. J. Chan and S. R. Ramkumar, Tracking error rebalancing, Journal of Portfolio Management 37 (2011), no. 4, 54–66.

[CT06] T. M. Cover and J. A. Thomas, Elements of information theory, Wiley, 2006.

[DESH07] M. A. H. Dempster, I. V. Evstigneev, and K. R. Schenk-Hoppe, Volatility-induced financial growth, Quant. Finance 7 (2007), no. 2, 151–160.
[Fer02] E. R. Fernholz, *Stochastic portfolio theory*, Applications of Mathematics, Springer, 2002.

[FK05] R. Fernholz and I. Karatzas, *Relative arbitrage in volatility-stabilized markets*, Annals of Finance 1 (2005), no. 2, 149–177.

[FK09] R. Fernholz and I. Karatzas, *Stochastic portfolio theory: an overview*, Handbook of Numerical Analysis (P.G. Ciarlet, ed.), Handbook of Numerical Analysis, vol. 15, Elsevier, 2009, pp. 89 – 167.

[FKK05] R. Fernholz, I. Karatzas, and C. Kardaras, *Diversity and relative arbitrage in equity markets*, Finance and Stochastics 9 (2005), no. 1, 1–27.

[FS82] E. R. Fernholz and B. Shay, *Stochastic portfolio theory and stock market equilibrium*, J. Finance 37 (1982), 615–624.

[JKO98] R. Jordan, D. Kinderlehrer, and F. Otto, *The variational formula of the Fokker-Planck equation*, SIAM J. Math. Anal. 29 (1998), no. 1, 1–17.

[Kel56] J. Kelly, *A new interpretation of information rate information theory*, Information Theory, IRE Transactions on 2 (1956), no. 3, 185–189.

[KL10] D. Kuhn and D. G. Luenberger, *Analysis of the rebalancing frequency in log-optimal portfolio selection*, Quant. Finance 10 (2010), no. 2, 221–234.

[PR11] E. Platen and R. Rendi, *Approximating the numéraire portfolio by naive diversification*, Journal of Asset Management 13 (2011), no. 1, 34–50.

[PUV12] Y. Plyakha, R. Uppal, and G. Vilkov, *Why does an equal-weighted portfolio outperform value-and price-weighted portfolios?*, Available at http://ssrn.com/abstract=1787045, 2012.

[RY99] D. Revuz and M. Yor, *Continuous martingales and brownian motion*, Grundlehren der mathematischen Wissenschaften A series of comprehensive studies in mathematics, Springer, 1999.

Department of Mathematics, University of Washington, Seattle, USA  
E-mail address: soumikpal@gmail.com, wongting@uw.edu