Log-optimal currency portfolios and control Lyapunov exponents

L. Gerencsér, M. Rásonyi , Zs. Vágó

Abstract—P. Algoet and T. Cover have shown that a log-optimal portfolio in a stationary market gives the highest asymptotic growth rate and presented algorithms to achieve this rate. There is no analogous result for markets with friction, of which a currency market is a typical example. In this paper we restrict ourselves to simple static strategies. The problem is then reduced to the analysis of products of random matrices, the top-Lyapunov exponent giving the growth rate. New insights to products of random matrices will be given and an algorithm for optimizing top-Lyapunov exponents will be presented with partial analysis. Simulation results for a 2 dimensional process, which reduces to Cover’s algorithm in the frictionless case, will be given.

Let $X = (X_n)$ be a stationary process of $k \times k$ real-valued matrices, depending on some vector-valued parameter $\theta \in \mathbb{R}^p$, satisfying $E \log^+ ||X_0(\theta)|| < \infty$ for all $\theta$. The top-Lyapunov exponent of $X$ is defined as

$$\lambda(\theta) = \lim_{n \to \infty} \frac{1}{n} E \log ||X_n \cdot X_{n-1} \ldots \cdot X_0||.$$ 

We develop an iterative procedure for the optimization of $\lambda(\theta)$. In the case when $X$ is a Markov-process, the proposed procedure is formally within the class defined in [3]. However the analysis of the general case requires different techniques: an ODE method defined in terms of asymptotically stationary random fields. The verification of some standard technical conditions, such as a uniform law of large numbers for the error process is hard. For this we need some auxiliary results which are interesting in their own right.

I. INTRODUCTION

In this article we are interested in maximizing the long-term profit of an investor who is trading in a stock or currency market. As it will be explained below, a security which behaves “nicely” in the short term may well perform badly in the long run. Hence instead of maximizing the expected value of (some function of) short-term returns one needs to optimize the growth rate of the portfolio, which requires substantially different techniques.

This problem was studied in [4] for a simple stock market model where daily returns were supposed to be independent and identically distributed. An algorithm was presented to determine the optimal constant proportion of wealth held in the assets. A shortcoming of this algorithm is that the distribution of stock returns should be known in advance.

This work has then been generalized in various ways: [2] proved the existence and asymptotic optimality of “log-optimal” portfolios for stationary ergodic stock returns, [1] gave universal schemes producing an asymptotically optimal growth rate without a priori knowledge of the stock’s distribution, see also [5] and [8]. Existence of optimal portfolios was shown in [12] for certain models with transaction costs, though without an explicit construction or algorithm.

The present paper provides a framework which is pertinent to a wide range of market models with or without transaction costs. Parametrized families of investment strategies are considered and an algorithm for maximizing the logarithmic growth rate of portfolios is presented. The results apply whenever portfolio position at time $t+1$ is computable from that of time $t$ by multiplication with a random matrix $X_t(\theta)$. Here $\theta$ is a parametrization of our chosen class of strategies and the sequence $X_t(\theta)$ is supposed to be stationary in the strong sense, for each $\theta$. We remark that this algorithm is based on principles completely different from those of [1].

It is quite natural in the case of markets with transaction costs or currency markets to consider vector portfolios instead of one-dimensional value processes so that one could keep track of positions in each asset. These positions are no longer equivalent: due to costly transfers in the case of transaction costs, and for obvious reasons in the case of currencies. Such a “vector” view of portfolios was adopted in [10] and has become fairly standard since.

In section 2 we present a simple example to motivate our goal, section 3 provides applications of the algorithm which is described in sections 4, 5 and 6 in a general setting. Section 7 contains simulation results.

II. A MOTIVATING EXAMPLE

We start with a simple case of the model in [4] and [2]. Suppose that an agent may invest in a bond and in a stock. For simplicity we assume that a unit of bond is worth $\$1$ all the time (i.e. interest rate is 0), the price of one unit of stock at time $t$ is denoted by $S_t$, we have $S_0 = 1$ and $S_{t+1} = Y_{t+1} S_t$ where $Y_t$ is a strongly stationary ergodic sequence of positive random variables.
The investor wants to keep a fixed proportion $0 \leq \alpha \leq 1$ of his wealth in the stock and the rest in the bond. If his overall wealth equals $V_t$ at time $t$ then

$$V_{t+1} = (1 - \alpha)V_t + \alpha V_t Y_t = [1 - \alpha + \alpha Y_t]V_t.$$  

A “short-sighted” optimality criterion would be taking $\alpha^*$ maximizing

$$E[1 - \alpha + \alpha Y_t].$$

(1)

However, the logarithmic growth rate is

$$\lim_{t \to \infty} \frac{\log V_t}{t} = E \log(\alpha Y_t + 1 - \alpha),$$

by the stong law of large numbers. The maximizer $\alpha^*$ for this latter expression is rather different from the one for (1). An investor seeking long-term profit wants to maximize the rate (2) and our objective is to provide an algorithm to this end.

If the $X_t$ are i.i.d. with known distribution, then the algorithm of [4] does the job; in the stationary case without knowing the distributions the techniques of [1] can be applied, see also section 3. In this paper our purpose is to develop a method which finds the optimal long-term investment in similar, but possibly much more complicated, parametric models, involving e.g. transaction costs, see section 3.

III. GROWTH RATE MAXIMIZATION

Let us consider a currency portfolio $\phi = (\phi_n)$ consisting of $k$ currencies. Thus $\phi_n = (\phi_{i,n}), \ i = 1, \ldots, k$, where $\phi_{i,n}$ denotes the physical size of the portfolio held in the $i$-th currency at time $n$. At any time $n$ the exchange rates are collected in a $k \times k$ matrix $P_n$. For any fixed $P = P_n$ the entry $p_{ij}$ gives the amount of currency $i$ that can be purchased for 1 unit of currency $j$. It is reasonable to suppose $p_{ii} = 1$ for all $i$ and

$$p_{ij} p_{ji} \preceq p_{ii}$$

for all $i, j, l$.

A strategy at any time $n$ for currency $j$ is a vector $b_j = (b_{jr}), \ r = 1, \ldots, k$ such that

$$\sum_{j=1}^k b_{jr} = 1, \ b_{jr} \geq 0.$$ 

It gives the proportion of volume of currency $r$ that is converted into currency $j$. The total portfolio is then represented by a matrix

$$B = (b_{ij}).$$

If the current portfolio is $\phi = (\phi_1, \ldots, \phi_k)$ then the amount of currency $i$ at the next period will be

$$\phi_i^+ = \sum_{j=1}^k p_{ij} b_{ij} \phi_j,$$

(3)

Write for the matrix with elements $p_{ij} b_{ij}$

$$X = P \odot B.$$  

Then the dynamics for the portfolio is

$$\phi_{n+1} = (P_{n+1} \odot B_{n+1}) \phi_n =: X_{n+1} \phi_n$$

where $B_n$ is the strategy selected at time $n$.

$(X_n)$ will be assumed to be a strictly stationary process. We consider now parametric strategies $B = (B_n)\theta$), possibly random, such that $(X, B)$ is a stationary process in the strong sense. The simplest case is a constant strategy $B_n(\theta) = B(\theta)$ for all $n$. The wealth or the value of the portfolio at time $t$ expressed in currency $i$ will be obtained from a scalar product of the form

$$V_{n,i} = \sum_{j=1}^k p_{ij,n} \phi_{n,j}.$$ 

Then the growth rate of the wealth will be

$$\lambda = \lim_{n \to \infty} \frac{1}{n} \log V_n,$$  

which, under appropriate initialization, is equal to the top-Lyapunov exponent of $(X_n)$. Its maximization can be carried out by the procedure proposed in Section 3.

Example 1. Consider a currency market with assets 1 and 2 e.g. dollars and euros. The respective exchange rates are then given by the processes $p_{12,t}, \ p_{21,t}$ which we assume strongly stationary. The investor is supposed to act as follows: if the mean euro exchange rate at time $t$

$$\hat{p}_t := \frac{p_{12,t} + 1/p_{21,t}}{2}$$

goes above a threshold $\overline{h}$ then he or she buys euros for a proportion $0 < \alpha < 1$ of his dollars. If it goes below another threshold $\underline{h} < \overline{h}$ then he or she converts proportion $\alpha$ of his euros into dollars. Formally, set

$$b_{12,t}(\alpha) := \alpha I_{[\hat{p}_t < \underline{h}]}, \ b_{21,t}(\alpha) := \alpha I_{[\hat{p}_t > \overline{h}]},$$

$$b_{22,t}(\alpha) := 1 - b_{12,t}(\alpha), \ b_{11,t}(\alpha) := 1 - b_{21,t}(\alpha).$$

The $\alpha$ giving optimal growth rate can be determined by the algorithm of section 6.

Example 2. Remembering the notations of section 2, define:

$$X_{n}(\alpha) := \begin{pmatrix} 1 & 0 \\ 0 & Y_t \end{pmatrix} \begin{pmatrix} \alpha & \alpha \\ (1-\alpha) & (1-\alpha) \end{pmatrix}.$$ 

In this way the problem presented in section 2 can be solved by the results of sections 4 and 6.

Example 3. Finally, we include an adaptation of the previous example to the case of currency markets. (The case of a stock market with transaction costs can be treated in much
the same way.) Let $P_t$ be as in Example 1. Suppose that the investor wishes to keep a fixed proportion $\alpha$ of his or her wealth (computed in dollars) in euros and the rest in dollars. Suppose that his current portfolio (in currency units) is $(\phi_1, \phi_2)$. We have to distinguish two cases:

**Case 1.** If

$$\phi_1/\alpha > p_{12}\phi_2/(1 - \alpha),$$

then some wealth (say, $\beta$ dollars) must be transferred from dollars to euros. This $\beta$ should satisfy

$$\frac{\phi_1 - \beta}{\alpha} = \frac{\phi_2 p_{12} + p_{21}\beta p_{12}}{1 - \alpha}.$$

From here one can easily compute

$$\beta = \frac{(1 - \alpha)\phi_1 - \alpha p_{12}\phi_2}{(1 - \alpha) + \alpha p_{21} p_{12}},$$

and the new portfolio positions are

$$\phi_1^+ = \phi_1 - \beta = \frac{p_{12}\alpha p_{21}\phi_1 + \alpha p_{12}\phi_2}{(1 - \alpha) + \alpha p_{21} p_{12}},$$

$$\phi_2^+ = \frac{(1 - \alpha)\phi_1^+}{p_{12}\alpha} = \frac{(1 - \alpha)p_{21}\phi_1 + (1 - \alpha)\phi_2}{(1 - \alpha) + \alpha p_{21} p_{12}}.$$

**Case 2.** If

$$\phi_1/\alpha < p_{12}\phi_2/(1 - \alpha),$$

then some units (say, $\gamma$) of euros must be converted into dollars such that

$$\frac{\phi_1 + \gamma p_{12}}{\alpha} = \frac{\phi_2 p_{12} - \gamma p_{12}}{1 - \alpha}.$$

Now

$$\gamma = \frac{\alpha\phi_2 p_{12} - (1 - \alpha)\phi_1}{p_{12}},$$

and

$$\phi_1^+ = \phi_1 + \gamma p_{12} = \alpha\phi_1 + \alpha\phi_2 p_{12},$$

$$\phi_2^+ = \frac{(1 - \alpha)\phi_1^+}{p_{12}\alpha} = (1 - \alpha)\phi_2 + (1 - \alpha)\phi_1/p_{12}.$$

Putting together these two cases and taking $p_{12} := p_{12,t}$ and $p_{21} := p_{21,t}$ we may define the (stationary random) transformations $X_t(\alpha)$ by

$$X_t(\alpha)(\phi_1, \phi_2) := (\phi_1^+, \phi_2^+).$$

Note that these transformations are linear if and only if $p_{12} = 1/p_{21}$, which corresponds to the case of frictionless markets (i.e. no bid-ask spread).

With the notation at the beginning of this section we may write

$$X_t(\alpha) = P_t \circ B_t(\alpha),$$

where

$$B_t(\alpha) = \left(\begin{array}{cc} I_{A_\alpha}(1 - \beta/\phi_1) + I_{A_\alpha}^c & \frac{I_{A_\alpha}^c\gamma/\phi_2}{I_{A_\alpha}/\beta/\phi_1} + I_{A_\alpha} \frac{1 - \gamma/\phi_2}{1}\end{array}\right),$$

and

$$A_t = \{(\phi_1, \phi_2) \in \mathbb{R}^2 : \phi_1/\alpha > \phi_2 p_{12}/(1 - \alpha)\}.$$

In this way we arrive at a “semilinear” modification of the model in section 4, see section 5. Using our algorithm one may find the constant proportion $\alpha$ with highest asymptotic logarithmic growth rate.

**Example 4.** Now we look at the case of $k$ currencies, (the first is, say, dollar). Suppose that we would like to keep fixed proportions of our wealth (computed in dollars) in these assets, represented by $q = (q_1, \ldots, q_k)$, where $q_i \geq 0$ and $q_1 + \ldots + q_k = 1$. While in the two-asset case portfolio rebalancing was rather an obvious issue, in the present situation there may be several possibilities to reach a position with proportion $q$ from a given initial position, and these might result in different overall wealth, i.e. certain transfers may cost more than others. An investor will try to choose the optimal transfer and now we describe how this can be done.

Assume that the present portfolio position is $\phi = (\phi_1, \ldots, \phi_k)$. Then using strategy $B = (b_{ij})_{1 \leq i, j \leq k}$ the new position $\phi^+(B)$ is given by (3). We are looking for the “optimal” rebalancing $b_{ij}$ for which

$$\phi_1^+/q_1 = \phi_2^+/q_2 = \ldots = \phi_k^+/q_k,$$

and

$$\sum_{j=1}^k b_{ij} = 1, \ 1 \leq i \leq k,\ b_{ij} \geq 0, \ 1 \leq j, i \leq k,$$

hold and which attains

$$\max_B \phi_1^+(B).$$

The constraints (6) and (7), together with the maximization criterion (8) form a linear programme, which can be solved with standard techniques. Define $X_t(q)(\phi) := \phi^+$ when prices are given by $P_t$. This is again a generically non-linear transformation. Applying sections 5 and 6 we may find the long-term optimal parameter $q$.

**IV. RANDOM MATRIX-PRODUCTS**

Let $X = (X_n), n = 0, 1, \ldots$ be a stationary process of $k \times k$ real-valued matrices over some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, satisfying

$$\mathbb{E} \log^+ ||X_0|| < \infty$$

(9)
where \( \log^+ x \) denotes the positive part of \( \log x \). It is well-known (see [6]) that under the above condition

\[
\lambda = \lim_{n \to \infty} \frac{1}{n} E \log \|X_n \cdot X_{n-1} \cdots X_0\| \tag{10}
\]

exists. Here \( \lambda = -\infty \) is allowed. The following result is fundamental in multiplicative ergodic theory (see [6]):

**Theorem 1:** Assume that the process \( X = (X_n) \) described above satisfies (9) and in addition it is ergodic. Then \( P \)-almost surely

\[
\lambda = \lim_{n \to \infty} \frac{1}{n} \log \|X_n \cdot X_{n-1} \cdots X_0\|.
\]

The number \( \lambda \), the exponential growth rate of the product \( \|X_n \cdot X_{n-1} \cdots X_0\| \), is called the top Lyapunov-exponent of the process \( X = (X_n) \). Now we can ask what happens if we apply the above random matrix products to a fixed vector. An answer is given by Oseledec’s theorem (see [13] and [11]):

**Theorem 2:** Under the conditions of Theorem 1 there exists a set of random subspaces of fixed dimension

\[
\mathbb{R}^k = V_0 \supset V_1(\omega) \supset \ldots \supset V_{p-1}(\omega) \supset V_p = 0
\]

with strict inclusion such that for all \( \omega \in \Omega \) and \( v \in V_1(\omega) \setminus V_{i-1}(\omega) \) we have

\[
\lim_{n \to \infty} \frac{1}{n} \log \|X_n(\omega)X_{n-1}(\omega) \cdots X_1(\omega)v\| = \lambda_i.
\]

The numbers \( \lambda_i \) are called Lyapunov-exponents. The theorem above implies that for \( v \in V_1(\omega) \) we have

\[
\lim_{n \to \infty} \frac{1}{n} \log \|X_n(\omega)X_{n-1}(\omega) \cdots X_1(\omega)v\| = \lambda.
\]

Following [11], let us consider the singular-value decomposition of the random product \( X_n \cdot X_{n-1} \cdots X_0 \), and write

\[
X_n \cdot X_{n-1} \cdots X_0 = U_nD_nV_n^T
\]

where \( U_n, V_n \) are orthogonal \( k \times k \) matrices, and \( D_n = \text{diag}(d_{n1}) \) is a diagonal matrix, with \( d_{n1} \) being the singular values in decreasing order: \( d_{n1} \geq \ldots \geq d_{nk} \). In particular \( d_{n1} = \|X_n \cdot X_{n-1} \cdots X_0\| \). The following extension of the Fürstenberg-Kesten theorem holds:

**Theorem 3:** Assume that the process \( X = (X_n) \) described above satisfies (9) and in addition it is ergodic. Then \( P \)-almost surely the following limit exists:

\[
\lambda_i = \lim_{n \to \infty} \frac{1}{n} \log d_{ni}.
\]

To characterize the asymptotic behavior of the orthogonal matrices \( U_n, V_n^T \) let us decompose the set \( \{1, \ldots, k\} \) into disjoint “intervals” \( I_1, \ldots, I_r \) such that for \( i, i' \in I_m \) we have \( \lambda_i = \lambda_{i'} \) but for \( i \in I_m, i' \in I_{m'} \), with \( m \neq m' \), \( \lambda_i > \lambda_{i'} \). Let us now consider the corresponding decomposition of the column-indices of \( V_n^T \) and decompose \( V_n^T \) accordingly as

\[
V_n^T = (V_{n1}^T, \ldots, V_{nr}^T).
\]

**Theorem 4:** Assume that the process \( X = (X_n) \) described above satisfies (9) and, in addition, it is ergodic. Then the linear subspaces spanned by the column-vectors of \( V_{nm}^T, m = 1, \ldots, r \) converge \( P \)-almost surely in the gap-metric when \( n \) tends to \( \infty \).

Assume now that \( \lambda_1 > \lambda_2 \). Then the above result implies that the first column of \( V_{n1}^T \), denoted by \( v_{11}^T \) converges \( P \)-almost surely to some limit that will be denoted by \( v_1^T \).

A theoretical expression for \( \lambda \) is given in [6] as follows. Let us define the normalized products

\[
Z_k = X_k \cdots X_1 / \|X_k \cdots X_1\| = \pi_m(X_k \cdots X_1)
\]

where for a non-singular matrix \( A \) we set

\[
\pi_m(A) = A/\|A\|.
\]

Then it is shown in [6] that the distribution of the process \( (X_{k+1}, Z_k) \) converges weakly to a stationary distribution in an appropriate Cesaro-sense. Let \( \mu(dx, dz) \) denote the stationary joint distribution of \( (X_2, Z_1) \) on \( \mathbb{R}^k \). Then we have

\[
\lambda = \int \log \|xz\| \mu(dx, dz).
\]

Obviously this expression is not very useful for practical computations. Much more constructive results are available if there is a gap between the first and second Lyapunov-exponent, i.e. if the co-dimension of \( V_1 \) is 1.

**Theorem 5:** Assume that the process \( X = (X_n) \) described above satisfies (9), it is ergodic, and \( \lambda_1 > \lambda_2 \). Then for some \( \varepsilon > 0 \) we have for \( \omega \in \Omega \)

\[
\pi_m(X_k \cdots X_1) = u_k^e(v_1^T)^T + O(e^{-\varepsilon k}),
\]

where \( (u_k^e) \) is a strictly stationary sequence of unit vectors, \( v_1^T \) is a fixed random unit vector, and the error term is a random variable bounded by \( C(\omega)e^{-\varepsilon k} \) with some finite \( C(\omega) \) and \( \gamma > 0 \).

Let us now take random vector \( \xi \in \mathbb{R}^k \) such that \( \xi \in V_1(\omega) \) almost surely, say, for \( \omega \in \Omega' \). E.g. take \( \xi \) independently of \( (X_n) \), having uniform distribution over \( \mathbb{S}_k = \{v \in \mathbb{R}^k, \|v\| = 1\} \). Redefine \( \Omega' \) as \( \Omega' \cap \Omega'' \). Assume that \( \lambda > -\infty \). Then \( X_n \cdot X_{n-1} \cdots X_1 \xi \neq 0 \) for all \( n \) and \( \omega \in \Omega' \). Let us define an \( \mathbb{R}^k \)-valued process \( z = (z_n) \), \( n \geq 0 \) as follows: \( z_0 = \xi/\|\xi\| \), and for \( n \geq 1 \)

\[
z_n = \frac{X_n \cdot X_{n-1} \cdots X_1 \xi}{\|X_n \cdot X_{n-1} \cdots X_1 \xi\|}.
\]
Obviously, $z = (z_n)$ can be defined recursively as follows:

$$z_{n+1} = \frac{X_{n+1}z_n}{|X_{n+1}z_n|} = \Pi_v(X_{n+1}z_n)$$

(16)

with initial condition $z_0 = \xi/|\xi|$, where now

$$\Pi_v(y) = y/|y|.$$  

It is easily seen that

$$\log |X_n \cdot X_{n-1} \cdots X_1\xi| = \sum_{k=0}^{n-1} \log |X_{k+1}z_k| + \log |\xi|.$$ 

Thus Theorem 1 implies

$$\lambda = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \log |X_{k+1}z_k|,$$ 

for $\omega \in \Omega'$.  

The following result immediately follows from 5, but in fact, it is the statement below which should be proved first and then one can deduce Theorem 5.

**Theorem 6:** Assume that the process $X = (X_n)$ described above satisfies (9) and in addition it is ergodic, and $\lambda_1 > \lambda_2$. Let $\xi V_1(\omega)$ almost surely, say for, $\omega \in \Omega'$, where $\Omega'$ was defined above. Then for some $\varepsilon > 0$ we have for $\omega \in \Omega'$

$$\pi_m(X_k \cdots X_1\xi) = u^*_k + O(e^{-\varepsilon k}).$$

(18)

The above result indicates that for all initial conditions $\xi V_1(\omega)$ the process $(z_n, X_n)$ is asymptotically stationary.  

A stationary initialization can be constructed as follows:

**Proposition 1:** Assume that the process $X = (X_n)$ satisfies (9), it is ergodic, and $\lambda_1 > \lambda_2$. Let $V_1$ have co-dimension 1. Let $\xi$ be uniformly distributed over the unit sphere, let it be independent from $(X_n)$ and define

$$z^*_0 = \lim_n \Pi_v(X_0X_{-1} \cdots X_{-n}\xi).$$

Then $z^*_0$ is a stationary initialization for (16), i.e. defining

$$z^*_{n+1} = \Pi_v(X_{n+1}z^*_n), \quad n \geq 0,$$

the process $(X_n, z^*_n)$ is stationary. Moreover, we have

$$E \log |X_1 z^*_0| = \lambda_1.$$ 

The fact that the process $(z_n)$ forgets its initial condition exponentially fast can be expressed also in an infinitesimal form:

**Proposition 2:** Assume that the process $X = (X_n)$ satisfies (9), it is ergodic, and $\lambda_1 > \lambda_2$. Assume that $\xi = \zeta_0(\omega) V_1(\omega)$ for $\omega \in \Omega'$. Then

$$\left| \frac{\partial z_n}{\partial \xi} \right| \leq C(\omega, \xi) e^{-\gamma n}$$

where $C(\omega, \xi)$ is finite and $\gamma > 0$ is constant.

V. STATE DEPENDENT RANDOM PRODUCTS

Consider now the problem when the random matrix $X$ can be written in the form

$$X_n = X(P_n, \phi_{n-1}),$$

where $X(P, \phi)$ is a fixed function of $P$ and $\phi$, which is continuous in $\phi$, $(P_n)$ is an ergodic strongly stationary random matrix-process satisfying (9), and $\phi_n \in \mathbb{R}^p$ is a sequence of vectors computed recursively by

$$\phi_{n+1} = X(P_{n+1}, \phi_n)\phi_n.$$ 

A standard example we have in mind is

$$X = P \odot B(\phi)$$

where $B(\phi)$ is a redistribution matrix depending on the current portfolio $\phi$, see Example 3 of section 3. Assuming that all elements of $P_n$ are positive for all $n$, it follows that for any non-negative, non-zero initial portfolio $\phi_0$ the portfolios $\phi_n$ will be non-negative and non-zero for all $n$. Thus we can define the normalized portfolio sequence

$$z_n = \phi_n/|\phi_n|.$$ 

Note that $B(\phi)$ is scale-invariant, thus we can write $B(\phi_n) = B(z_n)$. This will be assumed in general, i.e.:

$$X(P, \phi) = X(P, z) \quad \text{with} \quad z = \phi/|\phi|.$$ 

(19)

The process $z_n$ satisfies the usual recursion

$$z_{n+1} = \frac{X_{n+1}z_n}{|X_{n+1}z_n|} = \Pi_v(X(P_{n+1}, z_n))z_n$$

(20)

with initial condition $z_0 = \xi/|\xi|$. It is also obvious that the growth-rate can be expressed as follows:

$$\log |X_n \cdot X_{n-1} \cdots X_1\xi| = \sum_{k=0}^{n-1} \log |X_{k+1}z_k| + \log |\xi|.$$ 

Note, however, we cannot directly apply the results of the previous section, since the sequence of matrices $X(P_{n+1}, \phi_n)$ is not necessarily a stationary sequence. However, by a basic result of Has’minskii (see [9]) we get:

**Proposition 3:** Let $X_n = P_n \odot B(\phi_{n-1})$, where the process $(P_n)$ is stationary, ergodic and satisfies

$$E \log^+ ||P_0|| < \infty,$$

and $B(\phi)$ is continuous in $\phi$. Then there exists an initialization $z_0$ such that for the resulting sequence $\phi_n$ the sequence $(P_{n+1}, z_n)$ is stationary.

Now the sequence $X_n = P_n \odot B(z_{n-1})$ is stationary, ergodic, and satisfies (9). Let the corresponding top Lyapunov
exponent be denoted by $\lambda$. We also see that $(X_n^*, z_{n-1}^*)$ is almost surely with some $\lambda' \leq \lambda$.

For any initialization $z_0 = \xi$ write $X_n = X_n(\xi)$ and $X(\xi) = (X_n(\xi))_{n \geq 1}$. Its probability distribution will be denoted by $P^X(\xi)$. Now we have

$$P^{X^n} = \int P^X(\xi) dP(z_0^n = \xi).$$

It follows that for almost all $\xi \in \text{supp } z_0^n$

$$\lim_{n \to \infty} \frac{1}{n} \log |X_n(\xi) \cdot X_{n-1}(\xi) \cdots X_1(\xi)| = \lambda'$$

almost surely. It is an intriguing question whether the present $z_0^n$ identical with the $z_0^n$ constructed in the previous section.

Let us assume that $\lambda_1 > \lambda_2$. Then by Theorem 5 we have for some $\varepsilon > 0$ for $\omega \in \Omega'$

$$\pi_n(X_k^* \cdots X_1^*) = u_k^* (v_1^*)^T + O(e^{-\varepsilon k}),$$

where $(u_k^*)$ is a strictly stationary sequence of unite vectors, and $v_1^*$ is a fixed random unit vector. Assuming that the support of the measure of $v_1^*$ is the whole unit sphere, i.e. $\text{supp } v_1^* = \mathbb{S}^p$, we have for any fixed initial portfolio $z_0 = \xi$

$$\lim_{n \to \infty} \frac{1}{n} \log |X_n(\xi) \cdot X_{n-1}(\xi) \cdots X_1(\xi)| = \lambda.$$

VI. MAXIMIZATION OF THE TOP-LYAPUNOV EXPONENT

Assume now that the matrices $X_n, n = 0, 1, \ldots$ depend on a common parameter, say $\theta$, where $\theta \in D \subset \mathbb{R}^p$, and $D$ is an open domain. $\theta$ is considered as a control-parameter, and the top Lyapunov-exponent $\lambda = \lambda(\theta)$ will be a function of $\theta$, and will be called a controlled Lyapunov-exponent. The problem we consider is:

$$\max_\theta \lambda(\theta).$$

(22)

To compute the gradient of $\lambda$ with respect to $\theta$ consider first a pair of smooth curves $(X(t), z(t), t \geq 0)$ in $\mathbb{R}^{k \times k}$ and $\mathbb{R}^k$, respectively, with $X(0) = X, z(0) = z$, such that $Xz \neq 0$. Then it is easy to see that

$$\frac{d}{dt} \log |X(t)z(t)| = \frac{1}{|Xz|^2} (z^T X^T Xz + z^T X^T Xz).$$

A shorthand notation will be

$$\frac{d}{dt} \log |X(t)z(t)| = \dot{H}(X, z, \dot{X}, \dot{z}).$$

Let us now consider the case when $X_n = X_n(\theta)$ is a smooth function of $\theta$ for $\theta \in D \subset \mathbb{R}^p$, where $D$ is an open domain. Let $\theta_i$ for some $i = 1, \ldots, p$ be a fixed coordinate and let us write

$$X_{\theta_i, n} = \frac{\partial}{\partial \theta_i} X_n(\theta) \quad z_{\theta_i, n} = \frac{\partial}{\partial \theta_i} z_n(\theta).$$

Similarly, we write for the gradients

$$X_{\theta, n} = \frac{\partial}{\partial \theta} X_n(\theta) \quad z_{\theta, n} = \frac{\partial}{\partial \theta} z_n(\theta).$$

Differentiating the $k$-th term of (17) and setting

$$H_k(X, z, \theta_k) = \dot{H}(X, z, \theta_k)$$

we get formally the following expression for the gradient of $\lambda$, denoted by $\lambda_{\theta}$:

$$\lambda_{\theta} = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} H(X_{k+1}, z_k, X_{\theta, k+1}, z_{\theta, k}).$$

(24)

It is assumed that the partial derivatives $X_{\theta_i, k+1}$ are available explicitly. On the other hand the partial derivatives $z_{\theta_i, k}$ will be computed recursively, taking into account the recursive definition of $z_n$ given in (16). For this purpose consider the mapping of $\mathbb{R}^{k \times k} \times \mathbb{R}^k$ into $\mathbb{R}^{k \times k}$ defined by

$$f(X, z) = Xz/|Xz|$$

assuming that $Xz \neq 0$. Consider first a pair of smooth curves $(X(t), z(t), t \geq 0)$ in $\mathbb{R}^{k \times k}$ and $\mathbb{R}^k$, respectively, with $X(0) = X, z(0) = z$. Then it is easy to see that

$$\frac{d}{dt} f(X(t), z(t)) = f_X \dot{X} + f_z \dot{z} = g(X, z, \dot{X}, \dot{z}),$$

where

$$f_X \dot{X} = \left( \dot{X} - X \frac{X^T z}{|Xz|^2} \right) z,$$

and

$$f_z \dot{z} = \frac{X}{|Xz|} \left( I - z z^T X^T \frac{1}{|Xz|^2} \right) \dot{z}.$$

Applying the above notations we can express the derivatives $z_{\theta_i, n}(\theta)$ in a recursive manner as follows:

$$z_{\theta_i, n+1} = g(X_{n+1}, z_n, X_{\theta_i, n+1}, z_{\theta_i, n}).$$

The iterative scheme. Assume, that at time $n$ we have already computed $\theta_n$ and also $X_n, X_{\theta, n}, z_n$ and $z_{\theta, n}$. Observe $X_{n+1} = X_{n+1}(\theta_n)$ and $X_{\theta, n+1} = X_{\theta, n+1}(\theta_n)$. Then set

$$z_{n+1} = X_{n+1} z_n/|X_{n+1} z_n|$$

$$z_{\theta, n+1} = g(X_{n+1}, z_n, X_{\theta, n+1}, z_{\theta, n})$$

$$H_n = H(X_{n+1}, X_{\theta, n+1}, z_n, z_{\theta, n})$$

$$\theta_{n+1} = \theta_n + \frac{1}{n} H_n.$$
We propose to enforce boundedness of the estimator process by resetting: if \( \theta_{n+1} \) would leave a fixed compact domain then we reset to its initial value.

While the above method works well in simulation, its convergence analysis is yet incomplete. The algorithm formally falls within the class of recursive estimation methods described in [3] if \( X \) is a Markov-process, but the application of the results of [3] is not straightforward. In particular, the verification of Conditions A4 and A5 of Section 1.2 Part II of [3] seems to be hard. Also, [3] does not consider the effect of resetting.

The convergence analysis requires completely different tools if \( X \) is not Markov. The first step is relatively straightforward: the extension of a version of the ODE-method given in [7], to recursive estimation processes defined in terms of asymptotically stationary and ergodic random processes depending on a parameter \( \theta \). The hard part is to establish uniform laws of large numbers for not exactly stationary processes.

VII. SIMULATION RESULTS

We take the model of Example 3 and suppose that \( p_{12}(t), p_{21}(t) \) satisfy

\[
p_{12}(t + 1) := p_{12}(t) \xi_{t+1} (1 - d \varepsilon_{t+1}),
\]

\[
p_{21}(t + 1) := \frac{1}{p_{12}(t) \xi_{t+1}} (1 - d \varepsilon_{t+1}),
\]

where \( \xi_t \) are independent and identically distributed random variables with distribution

\[
P(\xi_t = c) = P(\xi_t = 1/c) = 1/2,
\]

and \( (\varepsilon_t) \) are independent and uniformly distributed random variables on \([0, 1]\). Here \( C > 1 \) and \( 0 < d < 1 \) are arbitrary constants. The price evolution is supposed to be driven by \( \xi_t \) while the \( \varepsilon_t \) are responsible for the margin of a dealer, i.e. they are thought to represent transaction costs.

We remark that if we choose \( d = 0 \) (no transaction costs) then the model reduces to Example 2. Choosing \( C := 2 \) we know from the Example on p. 370 of [4] that the optimal value of \( \alpha \) is \( \alpha^* = 0.5 \) and this gives an asymptotic logarithmic growth rate \( \lambda(\alpha^*) = 0.0588 \).

In our simulations we took \( C := 2, d := 0.1 \) and found that in this case the optimal value of \( \alpha \) is \( \alpha^* = 0.59 \). The corresponding growth rate \( \lambda(\alpha^*) \) decreases to 0.0375 due to the presence of transaction costs. The growth rate \( \lambda \) as a function of \( \alpha \) is shown on Figure 1.

The thick line in Figure 2 shows the convergence of our algorithm for the model with the above parameters, starting from \( \alpha_0 := 0.1 \). The horizontal axis shows the number of iteration on \( \alpha \). It is worth noting that about 30 iterations already gave a fairly good approximation of the optimal value. The thin line shows the performance of an algorithm when we replaced (25) by

\[
\theta_{n+1} = \theta_n + 0.05 H_n,
\]

that is we used a fixed-gain approximation.

Finally, Figure 3 corresponds to the initialization \( \alpha_0 := 0.75 \).

We may conclude that the algorithm converges fairly fast in a model class which could not be treated by previous methods.

ACKNOWLEDGEMENT

The first author expresses his thanks to James C. Spall and John L. Maryak of the Applied Physics Laboratory of Johns Hopkins University for cooperating in this research. The support of the National Research Foundation of Hungary (OTKA) under Grants nos. T 047193 and F 049094 is gratefully acknowledged.

REFERENCES


Fig. 2. Convergence for $\alpha_0 := 0.1$. Fixed gain iteration also included.

Fig. 3. Convergence for $\alpha_0 = 0.75$.


