



# Following a trend with an exponential moving average: Analytical results for a Gaussian model



Denis S. Grebenkov<sup>a,\*</sup>, Jeremy Serror<sup>b</sup>

<sup>a</sup> Laboratoire de Physique de la Matière Condensée, CNRS – Ecole Polytechnique, 91128 Palaiseau, France

<sup>b</sup> John Locke Investment, 38 Avenue Franklin Roosevelt, 77210 Fontainebleau-Avon, France

## HIGHLIGHTS

- How stock prices are transformed into profit-and-loss of a trend following strategy?
- We derive the probability distribution of P&L for a Gaussian model.
- We compute the net annualized risk adjusted P&L and the strategy turnover.
- We deduce the trend following optimal timescale and study its behavior.
- TF strategies admit large losses at short time but ensure small losses at longer time.

## ARTICLE INFO

### Article history:

Received 22 August 2013

Received in revised form 30 September 2013

Available online 12 October 2013

### Keywords:

Trend following

Systematic trading

Distribution of profits and losses

Correlated returns

Gaussian market model

Quadratic forms

## ABSTRACT

We investigate how price variations of a stock are transformed into profits and losses (P&Ls) of a trend following strategy. In the frame of a Gaussian model, we derive the probability distribution of P&Ls and analyze its moments (mean, variance, skewness and kurtosis) and asymptotic behavior (quantiles). We show that the asymmetry of the distribution (with often small losses and less frequent but significant profits) is reminiscent to trend following strategies and less dependent on peculiarities of price variations. At short times, trend following strategies admit larger losses than one may anticipate from standard Gaussian estimates, while smaller losses are ensured at longer times. Simple explicit formulas characterizing the distribution of P&Ls illustrate the basic mechanisms of momentum trading, while general matrix representations can be applied to arbitrary Gaussian models. We also compute explicitly annualized risk adjusted P&L and strategy turnover to account for transaction costs. We deduce the trend following optimal timescale and its dependence on both auto-correlation level and transaction costs. Theoretical results are illustrated on the Dow Jones index.

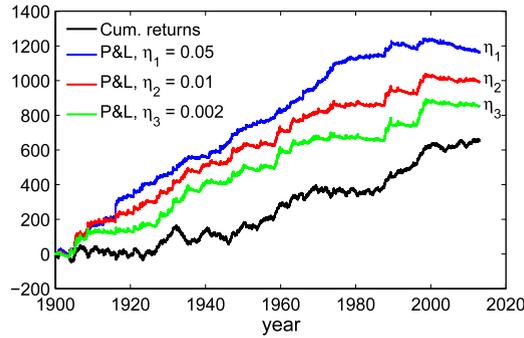
© 2013 Elsevier B.V. All rights reserved.

## 1. Introduction

Systematic trading has grown as an industry in finance, allowing to take rapid trading decisions for multiple stocks [1–7]. A strategy relies on price time series in the past in order to forecast price variations in near future and update accordingly its positions. Although the market complexity, variability and stochasticity damn such forecasting to fail in nearly half cases, even a tiny excess of successful forecasts is enhanced by a very large number of trades into statistically relevant profits. Many trading strategies attempt to detect an eventual trend in price series, i.e., a sequence of positively auto-correlated price variations which may be caused, e.g., by a news release or common activity of multiple traders. From a practical point

\* Corresponding author. Tel.: +33 619774327; fax: +33 169334799.

E-mail addresses: [denis.grebenkov@polytechnique.edu](mailto:denis.grebenkov@polytechnique.edu) (D.S. Grebenkov), [jeremy.serror@jl-investments.com](mailto:jeremy.serror@jl-investments.com) (J. Serror).



**Fig. 1.** Cumulative standardized logarithmic returns (normalized by realized volatility) of the Dow Jones index, from 1900 to 2012 (black curve), and cumulative P&Ls of trend following strategies (defined in Section 2) with timescales  $\eta_1 = 0.05$  (blue),  $\eta_2 = 0.01$  (red), and  $\eta_3 = 0.002$  (green), applied to this index.

of view, a strategy transforms the known past information into a signal for buying or selling a number of shares. From a mathematical point of view, systematic trading can be seen as a transformation of price time series into profit-and-loss (P&L) time series of the strategy, as illustrated in Fig. 1. For instance, the passive (long) strategy of buying and holding a stock corresponds to the identity transformation. The choice for the optimal strategy depends on the imposed risk–reward criteria.

In this paper, we study the transformation of price variations into P&Ls of a trend following strategy based on an exponential moving average (EMA). This archetypical strategy turns out to be the basis for many systematic trading platforms [1–7], while other methods such as the detrending moving average analysis or higher-order moving averages can also be employed [8–11]. A trend following strategy is known to skew the probability distribution of P&Ls [12,13], as we illustrate in Fig. 2. This figure shows how empirically computed quantiles of price variations<sup>1</sup> are transformed into quantiles of P&Ls for the Dow Jones index (1900–2012). Even for such a long sample with 30 733 daily returns, accurate estimation of quantiles remains problematic. Moreover, the basic mechanisms of this transformation remain poorly understood. For these reasons, we will study a simple model in which standardized logarithmic returns are Gaussian random variables [14] whose auto-correlations reflect random trends. Even though heavy tailed asymptotic distribution of returns and some other stylized facts are ignored [15–24], the Gaussian hypothesis will allow us to derive analytical results that can be later confronted to empirical market data. We will compute the probability distribution of P&Ls of a trend following strategy in order to understand how the Gaussian distribution of price variations is transformed by systematic trading. The respective roles of the market (positive auto-correlations) and of the strategy itself, onto profits and losses, will therefore be disentangled.

The paper is organized as follows. In Section 2, we introduce matrix notations, a market model and a trend following strategy. Main results about the probability distribution and moments of P&Ls are presented in Section 3. Discussion, conclusion and perspectives are summarized in Section 4.

## 2. Market model and trading strategy

### 2.1. Exponential moving average

The exponential moving average (EMA) is broadly employed in signal processing and data analysis [25–28]. The EMA can be defined as a linear transformation of a time series  $\{x_t\}$  to a smoother time series  $\{\tilde{x}_t\}$  according to

$$\tilde{x}_t = \lambda \sum_{k=0}^{\infty} (1 - \lambda)^k x_{t-k}, \quad (1)$$

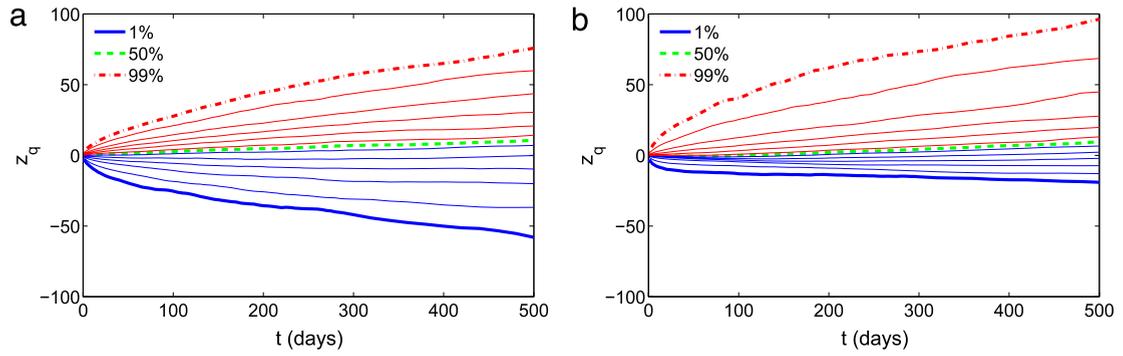
where  $0 < \lambda \leq 1$  is the (inverse of) timescale. When  $\lambda = 1$ , the EMA is the identity transformation:  $\tilde{x}_t = x_t$ ; in contrast, many terms  $x_{t-k}$  effectively contribute to  $\tilde{x}_t$  when  $\lambda \ll 1$ . The EMA is often preferred to simple moving average over a window of fixed length because it yields smoother results. In practice, it can be computed in real time according to a recurrent formula:

$$\tilde{x}_t = (1 - \lambda)\tilde{x}_{t-1} + \lambda x_t.$$

When a time series starts from  $t = 1$ , the non-existing elements  $x_0, x_{-1}, x_{-2}, \dots$  are set to 0. This is equivalent to setting the upper limit in Eq. (1) to  $t - 1$ . In the analysis of a finite sample of length  $T$ , the EMA can be written in a matrix form as

$$\begin{pmatrix} \tilde{x}_1 \\ \dots \\ \tilde{x}_T \end{pmatrix} = \lambda \mathbf{E}_{1-\lambda} \begin{pmatrix} x_1 \\ \dots \\ x_T \end{pmatrix},$$

<sup>1</sup> Here, by price variations we mean cumulative standardized logarithmic returns (normalized by realized volatility), to get closer to the Gaussian hypothesis [14].



**Fig. 2.** Comparison between quantiles of buy-and-hold strategy and trend following strategy. (a) Quantiles of cumulative standardized logarithmic returns (normalized by realized volatility) of the Dow Jones index (1900–2012) as functions of lag time  $t$  (in days). Lines from the bottom to the top show quantiles with 1%, 5%, 15%, 25%, 35%, 45% (blue color), 50% (green color), 55%, 65%, 75%, 85%, 95%, and 99% (red color), respectively. Thick lines highlight 1%, 50% and 99% quantiles. (b) Quantiles of cumulative P&Ls of a trend following strategy with the timescale  $\eta = 0.01$ , applied to the Dow Jones index (same notations).

where  $\mathbf{E}_q$  is the matrix of size  $T \times T$ , whose elements are

$$(\mathbf{E}_q)_{jk} = \begin{cases} q^{j-k-1} & (j > k), \\ 0 & (j \leq k). \end{cases} \quad (2)$$

## 2.2. Gaussian market models

The first Gaussian market model was introduced by Bachelier in 1900 and since that time, numerous models have been developed. For instance, the class of ARMA (Auto-Regressive Moving Average) models and their extensions were thoroughly employed in finance [25]. During decades, these models were getting more and more elaborate in order to account for various empirical features of markets. Our purpose is the opposite: we aim at understanding the basic mechanisms of trend following strategies, and we expect that qualitatively, these mechanisms weakly depend on market peculiarities. In turn, the quantitative behavior of P&Ls may of course be sensitive to particular features. For this reason, we choose a simple model exhibiting random trends, in order to be able to derive analytical results. At the same time, general matrix formulas used in this paper (see the beginning of Section 3) can be applied to arbitrary Gaussian market model. In this light, our methodology can be used for studying more elaborate models, though results will be less explicit.

In this paper, we consider a simple model of daily price variations, or returns,<sup>2</sup>  $r_t$ , in which random trends are induced by a discrete Ornstein–Uhlenbeck process, while short-time fluctuations are modeled by i.i.d. Gaussian variables  $\varepsilon_k \in \mathcal{N}(0, 1)$  with zero mean and unit variance:

$$r_t = \varepsilon_t + \beta \sum_{k=1}^{t-1} (1 - \lambda)^{t-1-k} \xi_k, \quad (3)$$

where  $\lambda$  and  $\beta$  are two parameters of the market model describing the characteristic timescale and the strength of the trend contribution, and  $\xi_k \in \mathcal{N}(0, 1)$  are i.i.d. Gaussian variables (independent of  $\varepsilon_k$ ). This is a model of stochastic trends which are induced by a persistent process generated by exogeneous random variables  $\xi_k$  which are independent from the short-time fluctuations  $\varepsilon_k$  [29]. In a matrix form, one writes

$$\mathbf{r} = \varepsilon + \beta \mathbf{E}_{1-\lambda} \xi, \quad (4)$$

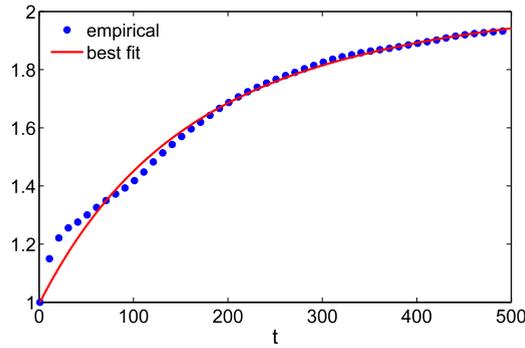
where  $\mathbf{r} = (r_1, \dots, r_T)^\dagger$  is the vector of returns (the superscript  $\dagger$  denoting the transpose), and  $\varepsilon$  and  $\xi$  are two vectors of i.i.d. Gaussian variables. As a consequence,  $\mathbf{r}$  is a Gaussian vector with zero mean,  $\langle r_j \rangle = 0$ , and the covariance matrix  $\mathbf{C}_{j,k} = \langle r_j r_k \rangle$ , for which

$$\mathbf{C} = \mathbf{I} + \beta^2 \mathbf{E}_{1-\lambda} \mathbf{E}_{1-\lambda}^\dagger, \quad (5)$$

where  $\mathbf{I}$  stands for the identity matrix. The elements of this matrix are

$$\mathbf{C}_{j,k} = \delta_{j,k} + \frac{\beta^2}{\lambda(2-\lambda)} \left[ (1-\lambda)^{|j-k|} - (1-\lambda)^{j+k-2} \right]. \quad (6)$$

<sup>2</sup> Throughout this paper, daily price variations will be called “returns” for the sake of simplicity. Rigorously speaking, we consider additive standardized logarithmic returns normalized by realized volatility. Such a resizing, which is a common practice on futures markets [13], allows one to reduce, to some extent, the impact of changes in volatility and its correlations [23,24], and to get closer to the Gaussian hypothesis of returns [14].



**Fig. 3.** Variogram of standardized logarithmic daily returns (normalized by realized volatility) of the Dow Jones index (full circles) and its best fit to Eq. (9) (red line) with  $\lambda = 0.011$  and  $\beta_0 = 0.08$ .

The second term is the covariance of a discrete Ornstein–Uhlenbeck process. The diagonal elements  $\mathbf{C}_{t,t}$  approach the constant  $\sigma_\infty^2 = 1 + \frac{\beta^2}{\lambda(2-\lambda)}$  as  $t \rightarrow \infty$ , i.e., auto-correlations increase the variance of returns. It is convenient to make the limiting variance  $\sigma_\infty^2$  independent of the timescale  $\lambda$  by rescaling the parameter  $\beta$  as

$$\beta = \beta_0 \sqrt{\lambda(2-\lambda)}, \tag{7}$$

so that  $\sigma_\infty^2 = 1 + \beta_0^2$ , independently of  $\lambda$ . In other words, the new parameter  $\beta_0^2$  is the asymptotic excess variance of returns due to their auto-correlations. This parameter can be calibrated from empirical price series. For this purpose, we consider the variogram of returns over the lag time  $t$

$$V_{t,t_0} \equiv \frac{\text{var}\{r_{t_0+1} + \dots + r_{t_0+t}\}}{\text{var}\{r_{t_0+1}\} + \dots + \text{var}\{r_{t_0+t}\}}, \tag{8}$$

where an initiation period of duration  $t_0$  can be ignored to approach the stationary regime. The variogram would be equal to 1 for i.i.d. random variables, while its deviations from 1 characterize auto-correlations between variables. Expressing the variogram through the covariance matrix in Eq. (5), one gets in the stationary limit  $t_0 \rightarrow \infty$ :

$$\lim_{t_0 \rightarrow \infty} \text{var}\{r_{t_0+1} + \dots + r_{t_0+t}\} = t \left[ 1 + \beta_0^2 \frac{2-\lambda}{\lambda} \right] - \frac{2(1-\lambda)\beta_0^2}{\lambda^2} (1 - (1-\lambda)^t)$$

and

$$V_{t,\infty} = 1 + \frac{2(1-\lambda)\beta_0^2}{\lambda(1+\beta_0^2)} \left( 1 - \frac{1-(1-\lambda)^t}{\lambda t} \right). \tag{9}$$

Fig. 3 shows the empirical variogram of returns obtained from the Dow Jones index (1900–2012), and its fit according to Eq. (9). Although the model fails to reproduce a steep growth of auto-corrections at short times, it captures correctly the behavior of the variogram at longer times and allows us to get realistic values for the parameters  $\lambda$  and  $\beta_0$  of the model:  $\lambda = 0.011$  and  $\beta_0 = 0.08$ . At the same time, these values are market dependent and, in general, difficult to calibrate. In what follows, the representative values  $\lambda = 0.01$  and  $\beta_0 = 0.1$  will be used for illustrative purposes.

For comparison, we also consider an auto-regressive model with exponential weights,

$$r_t = \varepsilon_t + \beta \sum_{k=1}^{t-1} (1-\lambda)^{t-1-k} r_k, \tag{10}$$

where  $\varepsilon_t \in \mathcal{N}(0, 1)$  are i.i.d. Gaussian variables. This is a model of autoregressive trends which are induced through auto-correlations with earlier returns. Writing Eq. (10) in a matrix form,  $\mathbf{r} = \varepsilon + \beta \mathbf{E}_{1-\lambda} \mathbf{r}$  and inverting this relation yields

$$\mathbf{r} = (\mathbf{I} - \beta \mathbf{E}_{1-\lambda})^{-1} \varepsilon = (\mathbf{I} + \beta \mathbf{E}_{1-\lambda+\beta}) \varepsilon, \tag{11}$$

where the explicit matrix inversion was possible due to the specific triangular structure of the matrix  $\mathbf{E}_q$ . As a consequence,  $\mathbf{r}$  is a Gaussian vector with zero mean and the covariance matrix

$$\mathbf{C} = (\mathbf{I} + \beta \mathbf{E}_{\tilde{q}})(\mathbf{I} + \beta \mathbf{E}_{\tilde{q}})^\dagger = \mathbf{I} + \beta(\mathbf{E}_{\tilde{q}} + \mathbf{E}_{\tilde{q}}^\dagger) + \beta^2 \mathbf{E}_{\tilde{q}} \mathbf{E}_{\tilde{q}}^\dagger, \tag{12}$$

where  $\tilde{q} = 1 - \lambda + \beta$ . Comparing this relation to Eq. (5), one notes the effective timescale  $\lambda - \beta$  (instead of  $\lambda$ ), and an additional term  $\beta(\mathbf{E}_{\tilde{q}} + \mathbf{E}_{\tilde{q}}^\dagger)$ . Although both models exhibit many similar features, they are not identical due to the presence

of this term. For the sake of simplicity, we focus on the stochastic trend model (defined by Eq. (3)), while similar results for the auto-regressive trend model are derived and discussed (see Appendix D). Qualitative conclusions of the paper do not depend on this choice.

It is worth noting that we focus on trends that are induced by auto-correlations, while mean returns are zero. While an extension to the case  $\langle r_t \rangle \neq 0$  is relatively straightforward, the studied situation with  $\langle r_t \rangle = 0$  allows us to easier illustrate the role of trend following strategy because the passive holding strategy is profitless in this case.

### 2.3. Trading strategy

Trading strategies often rely on an EMA of returns in order to detect eventual trends in price time series [25–28]. We consider the simple situation, in which the position of a trend following strategy,  $s_t$  (e.g., the number of shares), is proportional to the EMA of returns:

$$s_t = \gamma \sum_{k=1}^{t-1} (1-\eta)^{t-1-k} r_k, \quad (13)$$

where  $\eta$  and  $\gamma$  are two parameters of the strategy (in what follows, we will relate  $\gamma$  to  $\eta$ , the latter remaining the only parameter of the strategy). In other words, the position is regularly updated according to the indicator  $s_t$  in order to increase the number of shares during strong trends (large EMA) and decrease it for weak trends (small EMA). It is crucial that the position at time  $t$  is determined by *earlier* returns  $r_{t-1}, r_{t-2}, \dots$ , and does not rely on unavailable information on the present return  $r_t$ . In a matrix form, Eq. (13) reads as

$$\mathbf{s} = \gamma \mathbf{E}_{1-\eta} \mathbf{r}. \quad (14)$$

The profit-and-loss of a trend following strategy that holds  $s_t$  shares at step  $t$ , is  $s_t r_t$ , given that the return  $r_t$  is the (normalized) variation of price. The cumulative P&L after  $t$  steps is then

$$\mathcal{P}_{t,t_0} \equiv \sum_{k=t_0+1}^{t+t_0} r_k s_k = (\mathbf{r}^\dagger \mathbf{O}^{(t,t_0)} \mathbf{s}) = \gamma (\mathbf{r}^\dagger \mathbf{O}^{(t,t_0)} \mathbf{E}_{1-\eta} \mathbf{r}) = \frac{1}{2} (\mathbf{r}^\dagger \mathbf{M}_{1-\eta}^{(t,t_0)} \mathbf{r}), \quad (15)$$

where  $t_0$  is the duration of an initiation period,

$$\mathbf{M}_{1-\eta}^{(t,t_0)} \equiv \gamma [\mathbf{O}^{(t,t_0)} \mathbf{E}_{1-\eta} + \mathbf{E}_{1-\eta}^\dagger \mathbf{O}^{(t,t_0)}] \quad (16)$$

is a symmetric matrix, and  $\mathbf{O}^{(t,t_0)}$  is the matrix which has 1 in the diagonal positions between  $t_0 + 1$  and  $t_0 + t$ , and 0 elsewhere. The cumulative P&L in Eq. (15) is written as a quadratic form of the Gaussian vector  $\mathbf{r}$ . Similarly, an incremental P&L reads as

$$\delta \mathcal{P}_t \equiv \delta \mathcal{P}_{t,t_0} \equiv \mathcal{P}_{t,t_0} - \mathcal{P}_{t-1,t_0} = r_{t+t_0} s_{t+t_0} = \frac{1}{2} (\mathbf{r}^\dagger \mathbf{M}_{1-\eta}^{(1,\bar{t}-1)} \mathbf{r}), \quad (17)$$

where  $\bar{t} = t + t_0$  is a shortcut notation for  $t + t_0$ .

### 3. Profit-and-loss of trend following strategy

The representations (15), (17) of cumulative and incremental P&Ls as quadratic forms of Gaussian vectors allow one to investigate their properties. For a discrete Gaussian process  $\mathbf{r}$  with mean zero and covariance matrix  $\mathbf{C}$ , the quadratic form  $\chi = \frac{1}{2} (\mathbf{r}^\dagger \mathbf{M} \mathbf{r})$  defined by a symmetric matrix  $\mathbf{M}$  is a random variable whose moments and probability distribution are well known [30]. In fact, a matrix representation of the characteristic function of  $\chi$ ,

$$\phi(k) \equiv \langle \exp(ik\chi) \rangle = \frac{1}{\sqrt{\det(\mathbf{I} - ik\mathbf{M}\mathbf{C})}}, \quad (18)$$

yields the probability density  $p(z)$  of  $\chi$  through the inverse Fourier transform:

$$p(z) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{-ikz} \phi(k). \quad (19)$$

As described in Ref. [30], the determinant  $\det(\mathbf{I} - ik\mathbf{M}\mathbf{C})$  can be expressed through the eigenvalues of the matrix  $\mathbf{M}\mathbf{C}$  that speeds up numerical computations. Moreover, the smallest and the largest eigenvalues,  $\mu_-$  and  $\mu_+$ , essentially determine the asymptotic behavior of the probability density  $p(z)$ :

$$p(z) \propto A_{\pm} z^{\nu_{\pm}} \exp(-z/\mu_{\pm}) \quad (z \rightarrow \pm\infty) \quad (20)$$

(note that  $\mu_- < 0$  to ensure the decay of the density as  $z \rightarrow -\infty$ ).

Finally, the cumulant moments  $\kappa_m$  of the quadratic form  $\chi$  are

$$\kappa_m = \frac{(m - 1)!}{2} \text{tr}((\mathbf{MC})^m), \tag{21}$$

where  $\text{tr}$  denotes the trace. In particular,  $\kappa_1$  and  $\kappa_2$  are the mean and variance of  $\chi$ , while higher-order cumulant moments determine the skewness ( $\kappa_3/\kappa_2^{3/2}$ ) and kurtosis ( $\kappa_4/\kappa_2^2$ ).

### 3.1. Mean incremental P&L

We first consider the incremental P&L,  $\delta \mathcal{P}_{\bar{t}}$ , for which the matrix  $\mathbf{M}_{1-\eta}^{(1, \bar{t}-1)}$  from Eq. (16) has a particularly simple structure, with nonzero contributions only at  $\bar{t}$ -th row and column. The product of this matrix with the covariance matrix  $\mathbf{C}$  can be written explicitly, e.g.,

$$\kappa_1 = \frac{1}{2} \text{tr}(\mathbf{MC}) = \gamma \sum_{k=1}^{\bar{t}-1} p^{\bar{t}-k-1} \mathbf{C}_{\bar{t},k}, \tag{22}$$

where  $p = 1 - \eta$ . Substituting Eq. (5) into Eq. (22) yields the mean incremental P&L

$$\langle \delta \mathcal{P}_{\bar{t}} \rangle = \gamma \beta_0^2 \left[ q \frac{1 - (pq)^{\bar{t}-1}}{1 - pq} - q^{\bar{t}-1} \frac{p^{\bar{t}-1} - q^{\bar{t}-1}}{p - q} \right], \tag{23}$$

where  $q = 1 - \lambda$ . In the special case  $\eta = \lambda$ , this expression reduces to

$$\langle \delta \mathcal{P}_{\bar{t}} \rangle = \gamma \beta_0^2 \left[ q \frac{1 - q^{2(\bar{t}-1)}}{1 - q^2} - (\bar{t} - 1) q^{2\bar{t}-3} \right]. \tag{24}$$

In the stationary limit  $t_0 \rightarrow \infty$  (we recall that  $\bar{t} = t + t_0$ ), Eq. (23) yields the mean stationary daily P&L:

$$\langle \delta \mathcal{P}_{\infty} \rangle = \gamma \beta_0^2 \frac{1 - \lambda}{1 - (1 - \eta)(1 - \lambda)}. \tag{25}$$

The mean cumulative P&L,  $\langle \mathcal{P}_{t,t_0} \rangle$ , can be obtained by summing contributions in Eq. (23). In the stationary limit  $t_0 \rightarrow \infty$ , the mean cumulative P&L is simply proportional to  $t$ :

$$\langle \mathcal{P}_{t,\infty} \rangle = t \langle \delta \mathcal{P}_{\infty} \rangle.$$

### 3.2. Variance of incremental P&L

The variance  $v_{\bar{t}}$  of the incremental P&L,  $\delta \mathcal{P}_{\bar{t}}$ , is

$$v_{\bar{t}} \equiv \langle r_{\bar{t}}^2 s_{\bar{t}}^2 \rangle - \langle r_{\bar{t}} s_{\bar{t}} \rangle^2 = \gamma^2 \sum_{j,k=1}^{\bar{t}-1} [\mathbf{E}_{1-\eta}]_{\bar{t},j} [\mathbf{E}_{1-\eta}]_{\bar{t},k} [\langle r_{\bar{t}}^2 r_j r_k \rangle - \langle r_{\bar{t}} r_j \rangle \langle r_{\bar{t}} r_k \rangle].$$

Since  $\mathbf{r}$  is a Gaussian vector, Wick's theorem allows one to express the fourth-order correlation of Gaussian variables through the covariance matrix:

$$\langle r_{\bar{t}}^2 r_j r_k \rangle = \langle r_{\bar{t}}^2 \rangle \langle r_j r_k \rangle + 2 \langle r_{\bar{t}} r_j \rangle \langle r_{\bar{t}} r_k \rangle = \mathbf{C}_{\bar{t},\bar{t}} \mathbf{C}_{j,k} + 2 \mathbf{C}_{\bar{t},j} \mathbf{C}_{\bar{t},k},$$

from which

$$v_{\bar{t}} = \gamma^2 \left[ \mathbf{C}_{\bar{t},\bar{t}} (\mathbf{E}_{1-\eta} \mathbf{C} \mathbf{E}_{1-\eta}^\dagger)_{\bar{t},\bar{t}} + [(\mathbf{E}_{1-\eta} \mathbf{C})_{\bar{t},\bar{t}}]^2 \right]. \tag{26}$$

Substituting Eq. (5) yields

$$v_{\bar{t}} = \gamma^2 \left\{ \left( 1 + \beta_0^2 (1 - q^{2\bar{t}-2}) \right) \left( \frac{1 - p^{2\bar{t}-2}}{1 - p^2} + \frac{\beta_0^2 (1 - q^2)}{(1 - pq)(p - q)} \left[ \frac{p(1 - p^{2\bar{t}-2})}{1 - p^2} - \frac{q(1 - q^{2\bar{t}-2})}{1 - q^2} - \frac{(p^{\bar{t}-1} - q^{\bar{t}-1})^2}{p - q} \right] \right) + \frac{\beta_0^4 (1 - q^2)^2}{(p - q)^2} \left[ \frac{1 - (pq)^{\bar{t}-1}}{1 - pq} - \frac{1 - q^{2\bar{t}-2}}{1 - q^2} \right]^2 \right\}. \tag{27}$$

In the special case  $\eta = \lambda$ , one gets

$$v_{\bar{t}} = \frac{\gamma^2}{1 - q^2} \left\{ (1 + \beta_0^2(1 - q^{2\bar{t}-2})) \left( (1 - q^{2\bar{t}-2}) + \frac{\beta_0^2}{1 - q^2} \left[ 1 + q^2 - q^{2(\bar{t}-1)} \left( 1 + q^2 [1 + (\bar{t} - 1)(q^{-2} - 1)]^2 \right) \right] \right) \right. \\ \left. + \frac{\beta_0^4 q^2}{1 - q^2} \left[ 1 - q^{2(\bar{t}-1)} [1 + (\bar{t} - 1)(q^{-2} - 1)] \right]^2 \right\},$$

In the stationary limit  $t_0 \rightarrow \infty$ , Eq. (27) reduces to

$$v_{\infty} = \frac{\gamma^2}{1 - p^2} \left[ 1 + \frac{2\beta_0^2}{1 - pq} + \frac{\beta_0^4(1 + q^2 - 2p^2q^2)}{(1 - pq)^2} \right]. \tag{28}$$

Setting the parameter  $\gamma$  of the strategy to

$$\gamma^2 = 1 - p^2 = \eta(2 - \eta) \tag{29}$$

ensures the unit variance of the incremental P&L for the case of independent returns (i.e., when  $\beta_0 = 0$ ). The condition allows one to properly compare trend following and passive (long) strategies. When  $\lambda \ll 1$ ,  $\eta \ll 1$  and  $\beta_0^2 \ll 1$ , one gets  $v_{\infty} \approx 1 + \beta_0^2 \frac{2}{\lambda + \eta}$ , while the stationary variance of returns was  $\sigma_{\infty}^2 = 1 + \beta_0^2$ . In other words, the correction term  $\beta_0^2$  is enhanced by the large factor  $\frac{2}{\lambda + \eta}$ .

One can also consider the variogram of incremental P&Ls for which we derive in Appendix B the exact formula in the stationary limit  $t_0 \rightarrow \infty$ . Interestingly, the variogram of incremental P&L can be larger or smaller than the variogram of returns, depending on the timescale  $\eta$  of the strategy. It is worth noting that the variogram of incremental P&Ls is equal to 1 for the case of independent returns.

It is instructive to consider the net risk adjusted P&L of the strategy,  $\frac{\langle \delta \mathcal{P}_{\infty} \rangle - \langle \mathcal{T}_{\infty} \rangle}{\sqrt{v_{\infty}}}$ , in which the mean turnover  $\langle \mathcal{T}_{\infty} \rangle$  is included to account for transaction costs. In Appendix A, we derive the exact formula for the mean daily turnover  $\langle \mathcal{T}_t \rangle$  and its stationary limit  $\langle \mathcal{T}_{\infty} \rangle$ . For  $\lambda \ll 1$ ,  $\eta \ll 1$ , and linear transaction costs (i.e.,  $\alpha = 1$  in Eq. (A.1)), Eq. (A.7) becomes  $\langle \mathcal{T}_{\infty} \rangle \approx \frac{2}{\sqrt{\pi}} \theta \sqrt{\eta}$ , where  $\theta$  represents transaction cost. Using this approximate relation and approximations of Eqs. (25), (28) for  $\lambda \ll 1$ ,  $\eta \ll 1$ , and  $\beta_0^2 \ll 1$ , we obtain

$$\frac{\langle \delta \mathcal{P}_{\infty} \rangle - \langle \mathcal{T}_{\infty} \rangle}{\sqrt{v_{\infty}}} \approx \frac{\beta_0^2 \sqrt{2\eta} - \frac{2}{\sqrt{\pi}} \theta \sqrt{\eta} (\lambda + \eta)}{\sqrt{(\lambda + \eta)^2 + 2\beta_0^2 (\lambda + \eta)}}. \tag{30}$$

When there is no transaction cost (i.e.,  $\theta = 0$ ), this function is maximized at  $\eta_{\text{opt}} = \lambda \sqrt{1 + 2\beta_0^2/\lambda}$ , as illustrated in Fig. 4 (solid curve). For  $\lambda = 0.01$  and  $\beta_0 = 0.1$ , the position of the maximum is around  $\lambda \sqrt{3}$ , while the maximum level 0.8 of the annualized risk adjusted P&L (given by Eq. (30) conventionally multiplied by  $\sqrt{255}$ ) is a typical value for systematic trading. Interestingly, the optimal timescale of the strategy is not equal to the timescale  $\lambda$  of the market model but it is enhanced by the factor  $\sqrt{1 + 2\beta_0^2/\lambda}$  due to auto-correlations of returns. When transaction costs are included, an explicit expression for the optimal timescale  $\eta_{\text{opt}}$  is too lengthy.<sup>3</sup> As expected, an increase of the transaction cost  $\theta$  reduces the risk adjusted P&L but also shifts the position of the maximum to smaller  $\eta$  in order to get smoother position and thus reduce transactions. This behavior is illustrated in Fig. 4. Note that general formulas in Appendix A are also applicable to nonlinear transaction costs. Interestingly, the optimal timescale depends on  $\lambda$  and  $\beta_0$  through the ratio  $\beta_0^2/\lambda$  which is of the order of unity.

Finally, the strategy is profitable only if the net risk adjusted P&L is positive, i.e.,  $\langle \delta \mathcal{P}_{\infty} \rangle \geq \langle \mathcal{T}_{\infty} \rangle$ , from which one gets a simple condition on transaction costs

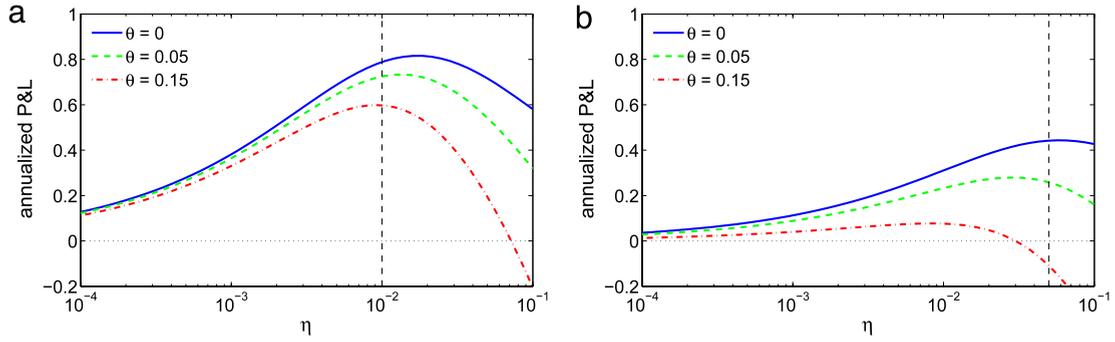
$$\theta \leq \sqrt{\pi/2} \frac{\beta_0^2}{\lambda + \eta}. \tag{31}$$

The inequality (31) can be seen as a limitation either on the maximal transaction cost  $\theta$ , or on the minimal level of auto-correlations  $\beta_0$ , or on the maximal timescale  $\eta$  of the strategy.

<sup>3</sup> In fact,  $\eta_{\text{opt}} = \lambda z$ , where  $z$  is the positive root of the cubic polynomial

$$\theta' z^3 + (c + \theta'(4c + 3))z^2 + 3\theta'(1 + 2c)z - (1 + 2c)(c - \theta') = 0,$$

which determines zeros of the derivative of Eq. (30) (here,  $\theta' = \theta \sqrt{2/\pi}$  and  $c = \beta_0^2/\lambda$ ). Although an exact solution can be written, the formula is too lengthy for further theoretical analysis. In turn, this formula can be used for numerical computation of the optimal timescale.



**Fig. 4.** Net annualized risk adjusted P&L,  $\sqrt{255} \frac{(\delta \mathcal{P}_\infty) - \langle \mathcal{P}_\infty \rangle}{\sqrt{v_\infty}}$ , as a function of the strategy timescale  $\eta$ , for the market with  $\beta_0 = 0.1$ , and (a) long-term correlations ( $\lambda = 0.01$ ) and (b) short-time correlation ( $\lambda = 0.05$ ). Three curves correspond to different transaction costs  $\theta$ : 0 (solid blue), 0.05 (dashed green) and 0.15 (dash-dotted red). Vertical black line indicates the market timescale  $\lambda$ .

### 3.3. Skewness and kurtosis

In principle, one can compute explicitly the other cumulant moments and access skewness and kurtosis of the cumulative P&L. However, these expressions become too lengthy for practical use. In turn, the general matrix formula (21) allows for rapid numerical computation of these quantities. Fig. 5 shows skewness ( $\kappa_3/\kappa_2^{3/2}$ ) and kurtosis ( $\kappa_4/\kappa_2^2$ ) of the cumulative P&L,  $\mathcal{P}_{t,t_0}$ , as functions of the lag time  $t$ . Both quantities exhibit a maximum at  $t \approx 1/\eta$ , i.e., the timescale of trend following strategy. In other words, the strategy induces auto-correlations of P&Ls that are significant up to time  $1/\eta$  and then slowly decay. In fact, if incremental P&Ls,  $\delta \mathcal{P}_{t_0+1}, \dots, \delta \mathcal{P}_{t_0+t}$ , were independent and identically distributed, the skewness and kurtosis of their sum,  $\mathcal{P}_{t,t_0}$ , would decay as  $1/\sqrt{t}$  and  $1/t$ , respectively. We emphasize that this behavior of  $\delta \mathcal{P}_t$  is induced by the trend following strategy itself, irrespectively of auto-correlations of returns. This is confirmed by the fact that both skewness and kurtosis behave similarly for independent and auto-correlated returns.

### 3.4. Distribution of incremental P&L

According to Eq. (17), the incremental P&L,  $\delta \mathcal{P}_t$ , is the quadratic form defined by the symmetric matrix  $\mathbf{M}_{1-\eta}^{(1,\bar{t}-1)}$  from Eq. (16). The probability distribution of  $\delta \mathcal{P}_t$  can therefore be determined through the inverse Fourier transform (19).

#### 3.4.1. Independent returns

We first consider the case of independent returns ( $\beta_0 = 0$ ), for which the covariance matrix is trivial:  $\mathbf{C} = \mathbf{I}$ . In that case, there are only two nonzero eigenvalues of the matrix  $\mathbf{M}\mathbf{C} = \mathbf{M}_{1-\eta}^{(1,\bar{t}-1)}$ ,

$$\mu_{\pm} = \pm \gamma \sqrt{\frac{1 - (1 - \eta)^{2(\bar{t}-1)}}{1 - (1 - \eta)^2}} = \pm \sqrt{1 - (1 - \eta)^{2(\bar{t}-1)}}, \quad (32)$$

where Eq. (29) was used in the last relation. As a consequence, the characteristic function of the incremental P&L is  $\phi(k) = (1 + k^2 \mu_+^2)^{-1/2}$ , from which the inverse Fourier transform yields

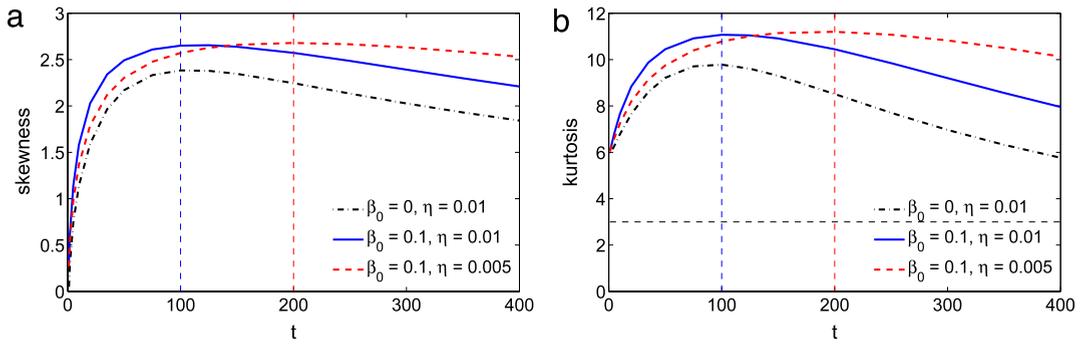
$$p(z) = \frac{K_0(|z|/\mu_+)}{\pi \mu_+}, \quad (33)$$

where  $K_0(x)$  is the modified Bessel function of the second kind. For large  $|z|$ , the asymptotic behavior is

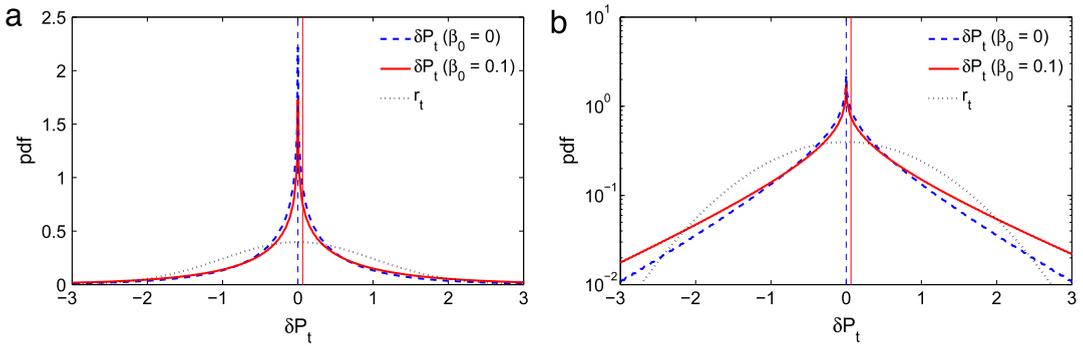
$$p(z) \simeq \frac{\exp(-|z|/\mu_+)}{\sqrt{2\pi \mu_+ |z|}} (1 + O(1/|z|)) \quad (|z| \rightarrow \infty). \quad (34)$$

Note that  $\mu_+^2$  is the variance of the incremental P&L. The skewness and kurtosis are 0 and 6, respectively (see Fig. 5, in which an incremental P&L corresponds to  $t = 1$ ). In the stationary limit  $t_0 \rightarrow \infty$ , one gets  $\mu_+ = 1$ .

Fig. 6 compares the probability density  $p(z)$  of the incremental P&L,  $\delta \mathcal{P}_t$ , with the Gaussian density  $(2\pi)^{-1/2} \exp(-z^2/2)$  of a single return  $r_t$  (with the unit variance). Although the mean and variance of these two distributions are identical, their overall behaviors are drastically different. The incremental P&L is peaked at 0 (in fact,  $K_0(z)$  logarithmically diverges at 0), while the tail decay is much slower than for returns. This transformation from a Gaussian density to  $p(z)$  is the effect of a trend following strategy.



**Fig. 5.** (a) Skewness,  $\kappa_3/\kappa_2^{3/2}$ , and (b) kurtosis,  $\kappa_4/\kappa_2^2$ , of the cumulative P&L,  $\mathcal{P}_{t,t_0}$  (with  $t_0 = 200$  and two timescales,  $\eta_1 = 0.01$  and  $\eta_2 = 0.005$ ), for auto-correlated returns ( $\beta_0 = 0.1$  and  $\lambda = 0.01$ ) and independent returns ( $\beta_0 = 0$ ). Two vertical dashed lines indicate  $1/\eta_1$  and  $1/\eta_2$ , while horizontal dashed line locates the level of kurtosis (equal to 3) of Gaussian distribution.



**Fig. 6.** Probability distribution of the incremental P&L,  $\delta\mathcal{P}_t$  ( $\bar{t} = 200$  and  $\eta = 0.01$ ), for independent returns (dashed blue line,  $\beta_0 = 0$ , Eq. (33)) and for auto-correlated returns (solid red line,  $\beta_0 = 0.1$ ,  $\lambda = 0.01$ ); (a) linear scale, (b) semi-logarithmic scale. The Gaussian distribution of returns  $r_t$  is plotted for comparison (dotted black line). Vertical lines indicate the mean value ( $\langle\delta\mathcal{P}_t\rangle$ ) for both cases (note that the variance of  $\delta\mathcal{P}_t$  is equal to 1 for independent returns and to 1.01 for auto-correlated returns).

### 3.4.2. Auto-correlated returns

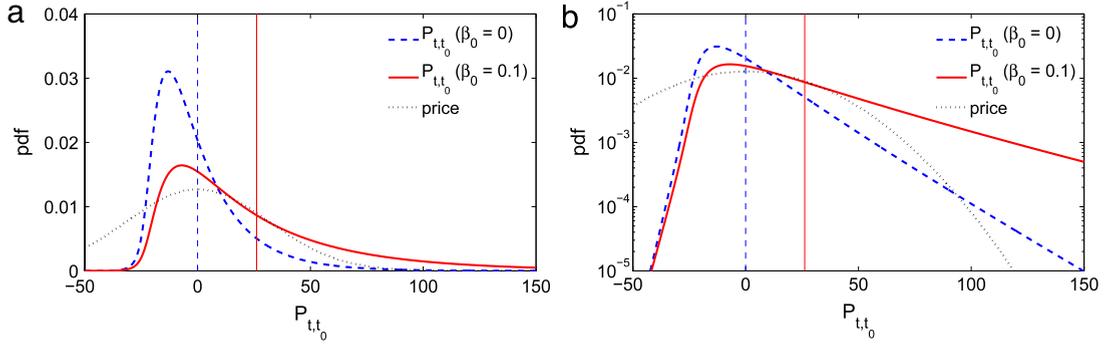
For auto-correlated returns ( $\beta_0 \neq 0$ ), the diagonalization of the matrix  $\mathbf{MC}$  and computation of the probability density  $p(z)$  can be performed numerically. As illustrated in Fig. 6, small auto-correlations of returns ( $\beta_0 = 0.1$ ) slightly modify the probability distribution (33) by shifting the mean ( $\langle\delta\mathcal{P}_t\rangle$ ) from zero to a small positive value and by increasing the probability of extreme values of  $\delta\mathcal{P}_t$  (both positive and negative tails).

### 3.5. Distribution of cumulative P&L

The probability distribution of the cumulative P&L,  $\mathcal{P}_{t,t_0}$ , can be obtained numerically through the inverse Fourier transform (19). Fig. 7 shows the probability density  $p(z)$  of  $\mathcal{P}_{t,t_0}$  for independent returns ( $\beta_0 = 0$ ) and for auto-correlated returns (with  $\beta_0 = 0.1$  and  $\lambda = 0.01$ ). The initiation period of  $t_0 = 200$  points was ignored to achieve stationary properties. In sharp contrast to symmetric (or almost symmetric) distributions of the incremental P&L from Fig. 6, the distribution of the cumulative P&L is strongly skewed and asymmetric, even for independent returns, in agreement with earlier observations [12,13]. The most probable P&L is negative, while the mean ( $\langle\mathcal{P}_{t,t_0}\rangle$ ) is nonnegative (it is 0 for  $\beta_0 = 0$  and strictly positive for  $\beta_0 \neq 0$ ). The positive mean P&L for  $\beta_0 \neq 0$  is ensured by relatively large probability for getting large positive profits. This result agrees with earlier observations that trend followers experience often small losses, waiting for a trend that may lead to considerable profits [12]. We emphasize again that skewness of  $\mathcal{P}_{t,t_0}$  emerges due to the trading strategy itself, irrespectively of auto-correlations of returns.

### 3.6. Quantiles

Inspecting the distribution in Fig. 7, one can clearly observe an exponential decay of both positive and negative tails, in agreement with the expected asymptotic behavior (20). Importantly, the decay of the probability of negative P&Ls is much steeper than that of positive P&Ls. These extreme events can be characterized by quantiles. For this purpose, one first integrates the density  $p(z)$  to get the cumulative probability distribution  $F(z)$  and then solves the equation  $F(z_q) = q$  with  $0 < q < 1$  that determines the  $q$ -quantile  $z_q$  of the distribution. In a first approximation, power law corrections in Eq. (20)



**Fig. 7.** Probability distribution of the cumulative P&L,  $\mathcal{P}_{t,t_0}$  ( $t = 300$ ,  $t_0 = 200$ , and  $\eta = 0.01$ ), for independent returns (dashed blue line,  $\beta_0 = 0$ ) and for auto-correlated returns (solid red line,  $\beta_0 = 0.1$ ,  $\lambda = 0.01$ ): (a) linear scale, (b) semi-logarithmic scale. The Gaussian distribution of price variation,  $r_{t_0+1} + \dots + r_{t_0+t}$ , is plotted for comparison (dotted black line). Vertical lines indicate the mean value  $\langle \mathcal{P}_{t,t_0} \rangle$  for both cases.

can be ignored (by setting  $\nu_{\pm} = 0$ ) so that

$$F(z) \approx A_- |\mu_-| \exp(-z/\mu_-) \quad (z \rightarrow -\infty),$$

$$1 - F(z) \approx A_+ \mu_+ \exp(-z/\mu_+) \quad (z \rightarrow \infty).$$

Extreme negative values of the cumulative P&L correspond to the limit  $q \rightarrow 0$  for which the equation  $F(z_q) = q$  can be approximately solved as

$$z_q \approx -|\mu_-| \ln \left( \frac{A_- |\mu_-|}{q} \right) \quad (q \rightarrow 0). \tag{35}$$

As expected, the behavior of the small  $q$ -quantile  $z_q$  is essentially determined by the smallest eigenvalue  $\mu_-$  of the matrix **MC**. In turn, extreme positive values of P&L correspond to the limit  $q \rightarrow 1$  for which

$$z_q \approx \mu_+ \ln \left( \frac{A_+ \mu_+}{1 - q} \right) \quad (q \rightarrow 1). \tag{36}$$

The large  $q$ -quantile is therefore mainly determined by the largest eigenvalue  $\mu_+$  of the matrix **MC**.

The behavior of the smallest and the largest eigenvalues  $\mu_{\pm}$  of the matrix  $\mathbf{M}_{1-\eta}^{(t,t_0)} \mathbf{C}$  for the cumulative P&L is illustrated in Fig. 8. The largest eigenvalue  $\mu_+$  grows with time  $t$  and slowly approaches a constant value at long  $t$ . In turn, the smallest eigenvalue  $\mu_-$  decreases and approaches a constant value much faster. For independent returns, we compute in Appendix C the asymptotic values  $\mu_{\pm}^{\infty}$ :

$$\mu_+^{\infty} = \frac{2\sqrt{\eta(2-\eta)}}{\eta} \approx \frac{\sqrt{8}}{\sqrt{\eta}}, \tag{37}$$

$$\mu_-^{\infty} = -\frac{\sqrt{\eta(2-\eta)}}{2\eta(1-\eta)(2-\eta)} \approx -\frac{1}{\sqrt{8\eta}}. \tag{38}$$

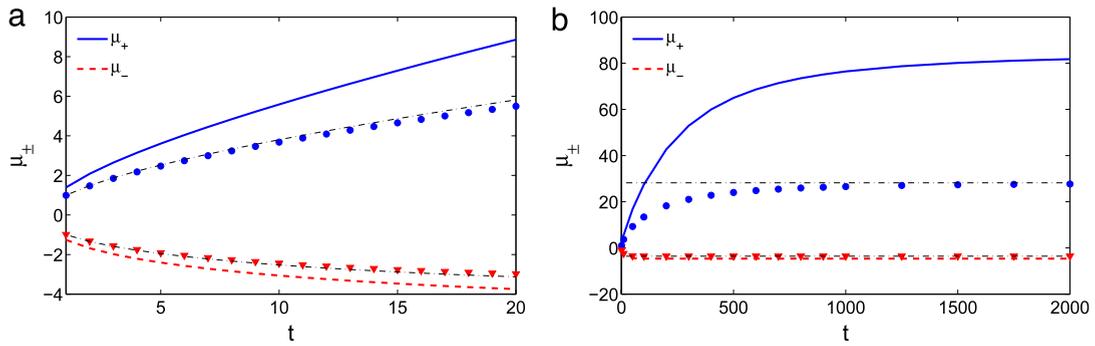
These values are shown in Fig. 8(b) by horizontal dash-dotted lines. One can see that  $\mu_+^{\infty}$  is 8 times larger than  $|\mu_-^{\infty}|$ . Most importantly, Eq. (38) turns out to be an accurate approximation for the smallest eigenvalue  $\mu_-$  even for auto-correlated returns. In other words, extreme negative P&Ls weakly depend on the market features (here,  $\beta_0$  and  $\lambda$ ) and are mainly determined by the trend following strategy (timescale  $\eta$ ). In turn, the largest eigenvalue  $\mu_+$  for auto-correlated returns may attain much larger values than  $\mu_+^{\infty}$  from Eq. (37). In other words, the presence of trends due to auto-correlations of returns increases  $\mu_+$  and thus enhances the probability of extreme positive P&Ls. In contrast to  $\mu_-$ , the largest eigenvalue  $\mu_+$  is sensitive to the market features.

We also analyzed the behavior of the largest and the smallest eigenvalues  $\mu_{\pm}$  in the opposite case of short times  $t$ . For independent returns, we compute the eigenvalues of the matrix  $\mathbf{M}_{1-\eta}^{(t,t_0)} \mathbf{C}$  for  $t = 2, 3, 4$  and  $t_0 \rightarrow \infty$  (see Appendix D). These explicit results suggest the conjectural asymptotic relation

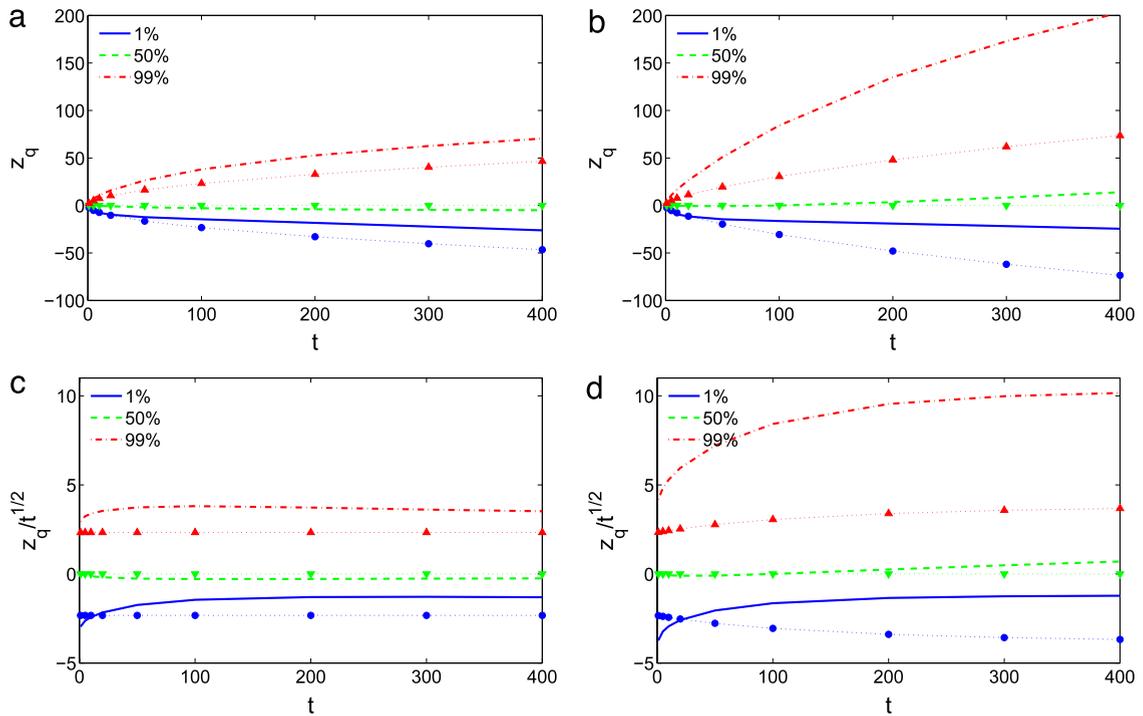
$$\mu_{\pm} \simeq \pm\sqrt{t} + (t-1)\sqrt{\eta/2} + O(\eta), \tag{39}$$

which is applicable for small  $\eta$  and moderate values of  $t$  (see Fig. 8(a)). At short times, the small  $q$ -quantile can be approximated as

$$z_q \approx -\sqrt{t} \left( 1 - \frac{t-1}{\sqrt{2t}} \sqrt{\eta} \right) \ln \left( \frac{A_- \sqrt{t}}{q} \right) \quad (q \rightarrow 0). \tag{40}$$



**Fig. 8.** The largest and the smallest eigenvalues of the matrix  $\mathbf{M}_{1-\eta}^{(t,t_0)} \mathbf{C}$  (with  $t_0 = 200, \eta = 0.01$ ), for independent returns (symbols) and for auto-correlated returns (lines,  $\beta_0 = 0.1, \lambda = 0.01$ ). (a) Short-time behavior. Black dash-dotted lines indicate the asymptotic behavior (39). (b) Long-time behavior. Horizontal black dash-dotted lines indicate the asymptotic values  $\mu_{\pm}^{\infty}$  from Eqs. (37), (38).



**Fig. 9.** Quantiles  $z_{0.01}, z_{0.5}$  and  $z_{0.99}$  (1%, 50%, and 99%) of the cumulative P&L,  $\mathcal{P}_{t,t_0}$  as functions of time  $t$ , with  $\eta = 0.01$  and  $t_0 = 200$ : (a) for independent returns ( $\beta_0 = 0$ ), and (b) auto-correlated returns ( $\beta_0 = 0.1, \lambda = 0.01$ ). The same quantiles of price variation over time  $t, r_{t_0+1} + \dots + r_{t_0+t}$ , are shown by symbols. Since Gaussian quantiles grow as  $\sqrt{t}$ , the ratio  $z_q/\sqrt{t}$  is also plotted for independent returns (c) and for auto-correlated returns (d), with the same parameters.

This behavior can be compared to the quantile of price variation over the time  $t, r_{t_0+1} + \dots + r_{t_0+t}$ . For independent returns, this is a Gaussian variable with mean zero and variance  $t$ , independently of the initiation period duration  $t_0$ . The cumulative probability distribution is  $\frac{1}{2}(1 + \text{erf}(z/\sqrt{2t}))$ , from which the  $q$ -quantile is given by the inverse error function:

$$z_q^0 = 2\sqrt{t}\text{erf}^{-1}(2q - 1). \tag{41}$$

One can see that this quantile grows as  $\sqrt{t}$ , while  $z_q^0/\sqrt{t}$  is constant. In turn, the quantile for the P&L, even after normalization by  $\sqrt{t}$ , exhibits a power law increase according to Eq. (40).

### 3.6.1. Auto-correlated returns

For auto-correlated returns ( $\beta_0 \neq 0$ ), the quantiles were computed numerically by solving the equation  $F(z_q) = q$ , in which the cumulative probability distribution  $F(z)$  was found by integrating the probability density  $p(z)$ .

Fig. 9 illustrates the behavior of quantiles for the cumulative P&L,  $\mathcal{P}_{t,t_0}$ , and for price variation over the same time  $t, r_{t_0+1} + \dots + r_{t_0+t}$ . For independent returns ( $\beta_0 = 0$ ), the Gaussian quantile grows as  $\sqrt{t}$  according to Eq. (41). In turn,

the quantiles for  $\mathcal{P}_{t,t_0}$  exhibit quite different behavior showing a strong asymmetry between positive and negative values. This asymmetry is further enhanced by auto-correlations of returns. The most interesting feature is the behavior of the renormalized quantile  $z_q/\sqrt{t}$  for small  $q = 0.01$  illustrated in Fig. 9(d). At short time  $t$ , the negative values of this quantile are smaller for P&L than for price variation. In other words, a trend following strategy may lead to more significant losses than one could naively anticipate from a Gaussian distribution of price variations. At larger times, the situation changes to the opposite: the quantile for the P&L exceeds that for price variation quite significantly. In other words, the trend following strategy ensures *smaller* losses at *longer* times, even in the absence of trends (when  $\beta_0 = 0$ , see Fig. 9(a), (c)).

One can also observe many similarities between Fig. 2(a), (b) (Dow Jones) and Fig. 9(a), (b) (present model). In particular, the 1% quantiles in both cases are close to each other, confirming that quantiles for extreme negative P&Ls weakly depend on market features and may be well approximated even by a simple model. In turn, the 99% quantiles in Figs. 2(b) and 9(b) are not so close, i.e., quantiles for extreme positive P&Ls are sensitive to market features.

#### 4. Conclusion

In this paper, we investigated how price variations of a stock are transformed into profits and losses of a trend following strategy. We started by deriving simple formulas for the mean and variance of the P&L, as well as the mean turnover of the strategy. The explicit expression for the net annualized risk adjusted P&L allowed us to analyze the profitability of trend following strategies in the presence of auto-correlations and transaction costs, and their sensitivity to the choice of parameters. We next proceeded in computing explicitly the probability distribution of P&L and investigating its asymptotic behavior. The theoretical analysis was mainly done for independent returns and confronted to numerical results for auto-correlated returns. Although the model of correlated returns was over-simplified, it allowed us to illustrate the basic features and mechanisms of a trend following strategy. Moreover, the general matrix formulas provided at the beginning of Section 3 are applicable to arbitrary Gaussian models.

It is worth emphasizing that quantitative results of this study are model dependent. For instance, we analyzed the asymptotic behavior of the probability density  $p(z)$  for extreme losses and showed its exponential decay with the rate  $\mu_-^\infty \approx -1/\sqrt{8\eta}$ . An exponential decay is a universal feature for quadratic forms of Gaussian vectors. However, the dependence of the smallest eigenvalue  $\mu_-$  on the parameter  $\eta$  was only derived for the studied trend following strategy. Moreover, the unknown prefactor  $A_-$  in the asymptotic formula (20) and related quantiles may strongly depend on other parameters, rendering quantitative estimates of quantiles model dependent. Finally, the asymptotic exponential decay of  $p(z)$  may settle at extremely large values of  $|z|$ , at which the probability density is negligible and out of practical interest. At the same time, qualitative conclusions of the study are expected to be general. In fact, a trend following strategy strongly modifies probabilistic properties of price time series, yielding skewed asymmetric distributions, with often small losses and less frequent high profits. The probability of extreme losses decays much faster than the probability of extreme profits. Moreover, the occurrence of extreme losses is more influenced by the trend following strategy than by the market itself. In turn, the occurrence of extreme profits depends on both the strategy and the market. We showed that the usual Gaussian paradigm may lead to erroneous conclusions about trend following strategies. For instance, at short times, trend following strategies admit larger losses than one may anticipate from standard Gaussian estimates. This is an important message for systematic traders and risk analysts.

The present analysis can be extended to arbitrary Gaussian models of returns and to multiple correlated stocks. The practical advantage of choosing *linear* equation (13) is that the position from EMAs of individual stocks is simply the sum of the related positions. As a consequence, the P&L of a portfolio is again a quadratic form of Gaussian vectors for which general matrix formulas at the beginning of Section 3 are still applicable. One can therefore study the role of inter-stock correlations which may significantly improve risk control of trend following strategies.

#### Appendix A. Mean turnover of trend following strategy

Accounting for transaction costs is important for a comprehensive analysis of trading strategies. We define the daily turnover of the trend following strategy as

$$\mathcal{T}_t = \theta |s_t - s_{t-1}|^\alpha, \quad (\text{A.1})$$

where  $\theta$  represents transaction cost, and  $\alpha$  is an appropriate exponent (typically  $\alpha = 1$  or  $\alpha = 2$ ). The mean turnover can be evaluated by using the identity

$$\langle f(\mathbf{r}^\dagger \mathbf{a}) \rangle = \int_{-\infty}^{\infty} dz f(z) \frac{1}{\sqrt{2\pi} (\mathbf{a}^\dagger \mathbf{C} \mathbf{a})} \exp\left(-\frac{z^2}{2(\mathbf{a}^\dagger \mathbf{C} \mathbf{a})}\right), \quad (\text{A.2})$$

where  $f(z)$  is a continuous function of the scalar product  $(\mathbf{r}^\dagger \mathbf{a})$ ,  $\mathbf{r}$  is a Gaussian vector with mean zero and covariance matrix  $\mathbf{C}$ , and  $\mathbf{a}$  is a fixed vector. Setting  $f(z) = |z|^\alpha$  and  $a_j = \gamma [(\mathbf{E}_{1-\eta})_{t,j} - (\mathbf{E}_{1-\eta})_{t-1,j}]$ , one gets

$$\langle \mathcal{T}_t \rangle = \theta \frac{\Gamma\left(\frac{1+\alpha}{2}\right)}{\sqrt{\pi}} [2(\mathbf{a}^\dagger \mathbf{C} \mathbf{a})]^{\alpha/2}$$

$$= \theta \frac{\Gamma\left(\frac{1+\alpha}{2}\right)}{\sqrt{\pi}} (2\gamma^2)^{\alpha/2} \left( [\mathbf{E}_p \mathbf{C} \mathbf{E}_p^\dagger]_{t,t} - 2[\mathbf{E}_p \mathbf{C} \mathbf{E}_p^\dagger]_{t,t-1} + [\mathbf{E}_p \mathbf{C} \mathbf{E}_p^\dagger]_{t-1,t-1} \right)^{\alpha/2}, \quad (\text{A.3})$$

where  $\Gamma(z)$  is Gamma function. Using Eq. (5), one obtains explicitly

$$\langle \mathcal{T}_t \rangle = \theta \frac{\Gamma\left(\frac{1+\alpha}{2}\right)}{\sqrt{\pi}} (2\gamma^2)^{\alpha/2} \left( \frac{2 - (1-p)p^{2t-4}}{1+p} + \frac{\beta_0^2(1-q^2)}{(1-pq)(p-q)} \left[ \frac{2(p-q)}{(1+p)(1+q)} - \frac{(1-p)p^{2t-3}}{1+p} + \frac{(1-q)q^{2t-3}}{1+q} - \frac{(p^{t-1} - p^{t-2} - q^{t-1} + q^{t-2})^2}{p-q} \right] \right)^{\alpha/2}. \quad (\text{A.4})$$

In the special case  $p = q$ , one gets

$$\langle \mathcal{T}_t \rangle = \theta \frac{\Gamma\left(\frac{1+\alpha}{2}\right)}{\sqrt{\pi}} (2\gamma^2)^{\alpha/2} \left( \frac{2 - (1-q)q^{2t-4}}{1+q} + \frac{\beta_0^2}{(1+q)^2} \left[ 2 - q^{2t-4} \left( 1 + [(q^{-1} - q)(t-2) - 1]^2 \right) \right] \right)^{\alpha/2}. \quad (\text{A.5})$$

In the stationary limit, one finds

$$\langle \mathcal{T}_\infty \rangle = \theta \frac{\Gamma\left(\frac{1+\alpha}{2}\right)}{\sqrt{\pi}} (2\gamma^2)^{\alpha/2} \left( \frac{2}{1+p} + \frac{2\beta_0^2(1-q^2)}{(1-pq)(1+p)(1+q)} \right)^{\alpha/2} \quad (\text{A.6})$$

that simplifies when  $\lambda \ll 1$  and  $\eta \ll 1$  to

$$\langle \mathcal{T}_\infty \rangle \approx \theta \frac{\Gamma\left(\frac{1+\alpha}{2}\right) 2^\alpha}{\sqrt{\pi}} \left( \eta + \frac{\beta_0^2 \lambda \eta}{\lambda + \eta} \right)^{\alpha/2} \approx \theta \frac{\Gamma\left(\frac{1+\alpha}{2}\right) 2^\alpha}{\sqrt{\pi}} \eta^{\alpha/2}. \quad (\text{A.7})$$

## Appendix B. Variogram of incremental P&L

We sketch the derivation of the variogram of incremental P&Ls in the stationary limit  $t_0 \rightarrow \infty$ . The variogram is defined as

$$V_{t,t_0}^{\text{P\&L}} = \frac{\text{var}\{\delta \mathcal{P}_{t_0+1} + \dots + \delta \mathcal{P}_{t_0+t}\}}{\text{var}\{\delta \mathcal{P}_{t_0+1}\} + \dots + \text{var}\{\delta \mathcal{P}_{t_0+t}\}}. \quad (\text{B.1})$$

The variances  $\text{var}\{\delta \mathcal{P}_k\}$  in the denominator are given by Eq. (27). One can explicitly compute their sum for  $k$  ranging from  $t_0 + 1$  to  $t_0 + t$  and then take the limit  $t_0 \rightarrow \infty$ . As expected, this limit is simply equal  $t v_\infty$ , where the stationary variance  $v_\infty$  is given by Eq. (28). The major difficulties rely in the computation of the numerator of Eq. (B.1) which contains correlations between incremental P&Ls.

We start by writing the definition of the variance

$$\begin{aligned} \text{var}\{\delta \mathcal{P}_{t_0+1} + \dots + \delta \mathcal{P}_{t_0+t}\} &= \sum_{j,k=t_0+1}^{t_0+t} [\langle r_j s_j r_k s_k \rangle - \langle r_j s_j \rangle \langle r_k s_k \rangle] \\ &= \sum_{j,k=t_0+1}^{t_0+t} [\langle r_j r_k \rangle \langle s_j s_k \rangle + \langle r_j s_k \rangle \langle r_k s_j \rangle], \end{aligned} \quad (\text{B.2})$$

where the second relation implied by Wick's theorem. Lengthy but straightforward computation yields

$$\begin{aligned} \langle s_j s_k \rangle &= p^{|j-k|} - p^{j+k-2} + \beta_0^2 \left( \frac{p(1-q^2)}{(1-pq)(p-q)} (p^{|j-k|} - p^{j+k-2}) - \frac{q(1-p^2)}{(1-pq)(p-q)} [p^{j+k-2} + q^{|j-k|} \right. \\ &\quad \left. - p^{j-1} q^{k-1} - p^{k-1} q^{j-1}] - \frac{1-p^2}{(p-q)^2} (p^{j-1} - q^{j-1})(p^{k-1} - q^{k-1}) \right) \end{aligned}$$

and

$$\frac{\langle r_j s_k \rangle}{\gamma \beta_0^2} = \begin{cases} \frac{p^{k-j-1}}{\beta_0^2} + p^{k-j-1} \frac{1-(pq)^j}{1-pq} + q \frac{p^{k-j-1} - q^{k-j-1}}{p-q} - q^{j-1} \frac{p^{k-1} - q^{k-1}}{p-q} & (k > j), \\ q^{j-k+1} \frac{1-(pq)^{k-1}}{1-pq} - q^{j-1} \frac{p^{k-1} - q^{k-1}}{p-q} & (k \leq j). \end{cases}$$

The numerator of Eq. (B.1) is then obtained by computing the double sum in Eq. (B.2). Since we are interested in the stationary limit  $t_0 \rightarrow \infty$ , it is sufficient to keep only the terms with  $j = k$ , in which the dependence on  $t_0$  is canceled. In the stationary limit, one gets the following variance

$$\begin{aligned} & \lim_{t_0 \rightarrow \infty} \text{var}\{\delta \mathcal{P}_{t_0+1} + \dots + \delta \mathcal{P}_{t_0+t}\} \\ &= \left[ 1 + 2\beta_0^2 \frac{1+q^2-2p^2q^2}{(1-pq)^2} + \beta_0^4 c_2 \right] t - 2\beta_0^2 q \frac{p+q-2p^2q}{(1-pq)^3} \left[ 1 + \beta_0^2 \frac{p(1-q^2)}{(1-pq)(p-q)} \right] (1-(pq)^t) \\ &+ \frac{4\beta_0^4 q^3 (1-p^2)}{(1-pq)(p-q)(1-q^2)^2} (1-q^{2t}), \end{aligned} \tag{B.3}$$

where

$$c_2 = \frac{2p^3q^5 + 2p^3q^3 - 6p^2q^2 + 2q^4p^2 - 4q^3p - q^5p + pq + 1 - q^4 + 4q^2}{(1-q^2)(1-pq)^3}.$$

For the special case  $p = q$ , one gets

$$\begin{aligned} \lim_{t_0 \rightarrow \infty} \text{var}\{\delta \mathcal{P}_{t_0+1} + \dots + \delta \mathcal{P}_{t_0+t}\} &= \left[ 1 + 2\beta_0^2 \frac{1+2q^2}{1-q^2} + \beta_0^4 \frac{1+7q^2+2q^4}{(1-q^2)^2} \right] t \\ &- \frac{4\beta_0^2 q^2}{(1-q^2)^2} \left[ 1 + \beta_0^2 \frac{2q^2+3/2}{1-q^2} \right] (1-q^{2t}) + \frac{4\beta_0^4 q^2}{(1-q^2)^2} q^{2t} t. \end{aligned}$$

Dividing Eq. (B.3) by  $t v_\infty$ , one gets the variogram  $V_{t,\infty}^{\text{P\&L}}$ . This expression can be compared to the variogram of returns from Eq. (9). Both variograms behave similarly, exhibiting both rapid exponential decay and slow power law decay. The asymptotic value of the variogram  $V_{t,\infty}^{\text{P\&L}}$  as  $t \rightarrow \infty$  is

$$V_{\infty,\infty}^{\text{P\&L}} = \frac{1 + 2\beta_0^2 \frac{1+q^2-2p^2q^2}{(1-pq)^2} + \beta_0^4 c_2}{1 + \frac{2\beta_0^2}{1-pq} + \frac{\beta_0^4(1+q^2-2p^2q^2)}{(1-pq)^2}}. \tag{B.4}$$

In the special case  $p = q = 1 - \lambda$ , one gets for  $\lambda \ll 1$  and  $\beta_0 \ll 1$

$$V_{\infty,\infty}^{\text{P\&L}} \approx 1 + \frac{2\beta_0^2}{\lambda} + \frac{(\beta_0^2/\lambda)^2}{2(1+\beta_0^2/\lambda)}. \tag{B.5}$$

For comparison, the variogram of returns from Eq. (9) gets the asymptotic value  $1 + 2\beta_0^2/\lambda$ . It is worth noting that  $\beta_0^2/\lambda$  does not need to be a small parameter.

### Appendix C. The largest and the smallest eigenvalues at long times

We present explicit formulas for the eigenvalues of the matrix  $\mathbf{MC}$  determining the cumulative P&L for independent returns at long times. In that case,  $\mathbf{MC} = \mathbf{M}_{1-\eta}^{(t,t_0)}$ . We first consider the simpler case when an initiation period is not ignored (i.e.,  $t_0 = 0$ ). In the limit  $t \rightarrow \infty$ , the matrix  $\mathbf{M}_{1-\eta}^{(t,0)} = \gamma[\mathbf{E}_{1-\eta} + \mathbf{E}_{1-\eta}^\dagger]$  becomes close to a cyclic matrix whose eigenvalues can be computed as

$$\mu_{1-\omega} = \gamma \sum_{j=1}^{\infty} |p|^{j-1} [e^{i\pi j\omega} + e^{-i\pi j\omega}] = 2\gamma \frac{\cos(\pi\omega) - p}{1 - 2p \cos(\pi\omega) + p^2}, \tag{C.1}$$

where  $\omega \in [0, 1]$  is the “index”, and  $p = 1 - \eta$ . The largest eigenvalue of the limiting matrix is then

$$\mu_+^\infty = \mu_{\omega=1} = \frac{2\gamma}{\eta}, \tag{C.2}$$

while the smallest eigenvalue is  $\mu_{\omega=0} = -\frac{\gamma}{1-\eta/2}$ .

When  $t_0 > 0$ , discarding the first  $t_0$  points corresponds to setting the block of size  $t_0 \times t_0$  of the matrix  $\mathbf{M}_{1-\eta}^{(t,0)}$  to zero. This modification introduces  $t_0$  zero eigenvalues into the spectrum and also changes the smallest eigenvalue  $\mu_-$  which becomes significantly smaller than  $\mu_{\omega=0}$ . In what follows, we compute the smallest eigenvalue  $\mu_-$  in the double limit  $t_0 \rightarrow \infty$  and

$t \rightarrow \infty$ . For this purpose, we first guess the corresponding eigenvector  $\mathbf{U}$  and then justify explicitly the correctness of the guess. We take the vector  $\mathbf{U}$  of the form:

$$\mathbf{U} = \underbrace{(p^{t_0-1}, p^{t_0-2}, \dots, p, 1, -a, -a\tilde{p}, -a\tilde{p}^2, \dots)}_{t_0 \text{ elements}}^\dagger,$$

where  $a$  and  $\tilde{p}$  are two parameters to be determined. Applying the matrix  $\mathbf{M}_{1-\eta}^{(t, t_0)}$  to this vector, one gets

$$\mathbf{M}_{1-\eta}^{(t, t_0)} \mathbf{U} = \gamma \begin{pmatrix} -ap^{t_0-1}(1+p\tilde{p}+\dots) \\ -ap^{t_0-2}(1+p\tilde{p}+\dots) \\ \dots \\ -a(1+p\tilde{p}+\dots) \\ \frac{(1+p^2+\dots) - a(\tilde{p}[1+p\tilde{p}+\dots])}{p(1+p^2+\dots) - a(1+\tilde{p}^2[1+p\tilde{p}+\dots])} \\ p^{k+1}(1+\dots) - a([p^k + p^{k-1}\tilde{p} + \dots + \tilde{p}^k] + \tilde{p}^{k+2}[1+\dots]) \\ \dots \end{pmatrix},$$

where the geometrical series  $1+p\tilde{p}+(p\tilde{p})^2+\dots$  and  $1+p^2+p^4+\dots$  contain infinitely many terms in the limits  $t \rightarrow \infty$  and  $t_0 \rightarrow \infty$ , respectively. If  $\mathbf{U}$  is an eigenvector, the right hand side has to be identified with  $\mu_-^\infty \mathbf{U}$ . The first  $t_0$  identities read  $\mu_-^\infty = \frac{-a\gamma}{1-p\tilde{p}}$ . The identity in the  $(t_0+1)$ -th row gives  $-a\mu_-^\infty = \gamma[\frac{1}{1-p^2} - \frac{a\tilde{p}}{1-p\tilde{p}}]$ , from which

$$\frac{a(a+\tilde{p})}{1-p\tilde{p}} = \frac{1}{1-p^2}. \quad (\text{C.3})$$

Finally, an identity in the  $(t_0+k+2)$ -th row reads

$$-a\tilde{p}^{k+1}\mu_-^\infty = \gamma \left[ \frac{p^{k+1}}{1-p^2} - a(p^k + p^{k-1}\tilde{p} + \dots + \tilde{p}^k) - \frac{a\tilde{p}^{k+2}}{1-p\tilde{p}} \right],$$

from which one expresses  $\tilde{p} = p - a(1-p^2)$ . Substituting this relation into Eq. (C.3), one gets  $a = 1/p$ , from which  $\tilde{p} = 2p - 1/p$  and finally

$$\mu_-^\infty = -\frac{\gamma}{2p(1-p^2)}. \quad (\text{C.4})$$

In this way, we constructed explicitly an eigenvalue  $\mu_-^\infty$  and the corresponding eigenvector  $\mathbf{U}$  of the limiting matrix  $\mathbf{M}_{1-\eta}^{(\infty, \infty)}$ . In turn, we did not show that  $\mu_-^\infty$  is the smallest eigenvalue. Although the related demonstration could in principle be performed, this analysis is beyond the scope of the paper. We checked numerically that  $\mu_-^\infty$  from Eq. (C.4) accurately approximates the smallest eigenvalue of the matrix  $\mathbf{M}_{1-\eta}^{(t, t_0)}$  for large enough  $t_0$  and  $t$ .

## Appendix D. Supplementary material

Supplementary material related to this article can be found online at <http://dx.doi.org/10.1016/j.physa.2013.10.007>.

## References

- [1] M.W. Covel, Trend Following (Updated Edition): Learn to Make Millions in Up or Down Markets, Pearson Education, New Jersey, 2009.
- [2] A.F. Clenow, Following the Trend: Diversified Managed Futures Trading, Wiley & Sons, Chichester UK, 2013.
- [3] L.K.C. Chan, N. Jegadeesh, J. Lakonishok, J. Finance 51 (1996) 1681.
- [4] L.K.C. Chan, N. Jegadeesh, J. Lakonishok, Financ. Anal. J. 55 (1999) 80.
- [5] N. Jegadeesh, S. Titman, J. Finance 56 (2001) 699.
- [6] T.J. Moskowitz, Y.H. Ooi, L.H. Pedersen, J. Financ. Econ. 104 (2012) 228.
- [7] C.S. Asness, T.J. Moskowitz, L.H. Pedersen, J. Finance 68 (2013) 929.
- [8] N. Vandewalle, M. Ausloos, Phys. Rev. E 58 (1998) 6832.
- [9] N. Vandewalle, A. Ausloos, P. Boveroux, Physica A 269 (1999) 170.
- [10] A. Carbone, G. Castellì, H.E. Stanley, Phys. Rev. E 69 (2004) 026105.
- [11] S. Arianos, A. Carbone, C. Türk, Phys. Rev. E 84 (2011) 046113.
- [12] M. Potters, J.-P. Bouchaud, Wilmott Mag. (2006) online: ArXiv physics-0508104 (2005).
- [13] R. Martin, D. Zou, Risk Mag. (2012) 52.
- [14] T.G. Andersen, T. Bollerslev, F.X. Diebold, P. Labys, Multinat. Finance J. 4 (2000) 159.
- [15] J.-P. Bouchaud, M. Potters, Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management, Cambridge University Press, 2003.
- [16] R. Mantegna, H.E. Stanley, An Introduction to Econophysics, Cambridge University Press, Cambridge, 1999.
- [17] R. Mantegna, H.E. Stanley, Nature 376 (1995) 46.

- [18] Gabaix, Gopikrishnan, Plerou, Stanley, *Nature* 423 (2003) 267.
- [19] J.-P. Bouchaud, M. Potters, *Physica A* 299 (2001) 60.
- [20] D. Sornette, *Phys. Rep.* 378 (2003) 1.
- [21] J.-P. Bouchaud, Y. Gefen, M. Potters, M. Wyart, *Quant. Finance* 4 (2004) 176.
- [22] A.L. Stella, F. Baldovin, *Pramana J. Phys.* 71 (2008) 341.
- [23] J.-P. Bouchaud, A. Matacz, M. Potters, *Phys. Rev. Lett.* 87 (2001) 1.
- [24] S. Valeyre, D.S. Grebenkov, S. Aboura, Q. Liu, *Quant. Finance* (2013) in press.
- [25] G. Box, G.M. Jenkins, G.C. Reinsel, *Time Series Analysis: Forecasting and Control*, third ed., Prentice-Hall, 1994.
- [26] C.C. Holt, Office of Naval Research Memorandum 52 (1957); reprinted in *Int. J. Forecast.* 20 (2004) 5.
- [27] P.R. Winters, *Manag. Sci.* 6 (1960) 324.
- [28] R.G. Brown, *Smoothing Forecasting and Prediction of Discrete Time Series*, Prentice-Hall, Englewood Cliffs, NJ, 1963.
- [29] D.S. Grebenkov, *Phys. Rev. E* 88 (2013) 032140.
- [30] D.S. Grebenkov, *Phys. Rev. E* 84 (2011) 031124.