



## Optimal allocation of trend following strategies



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 Investments, 38 Avenue Franklin Roosevelt, 77210 Fontainebleau-Avon, France

### HIGHLIGHTS

- Allocation of trend following strategies is studied within a Gaussian market model.
- Analytical formulas for the mean and variance of the portfolio return are derived.
- The optimal portfolio is constructed by introducing lead-lag corrections.
- Inter-asset correlations are not deteriorative but beneficial for trend followers.
- These concepts are illustrated on three highly correlated futures markets.

### ARTICLE INFO

#### Article history:

Received 30 October 2014

Received in revised form 23 February 2015

Available online 2 April 2015

#### Keywords:

Optimal allocation  
Trend following  
Asset correlations  
Diversification  
Gaussian returns  
Futures

### ABSTRACT

We consider a portfolio allocation problem for trend following (TF) strategies on multiple correlated assets. Under simplifying assumptions of a Gaussian market and linear TF strategies, we derive analytical formulas for the mean and variance of the portfolio return. We construct then the optimal portfolio that maximizes risk-adjusted return by accounting for inter-asset correlations. The dynamic allocation problem for  $n$  assets is shown to be equivalent to the classical static allocation problem for  $n^2$  virtual assets that include lead-lag corrections in positions of TF strategies. The respective roles of asset auto-correlations and inter-asset correlations are investigated in depth for the two-asset case and a sector model. In contrast to the principle of diversification suggesting to treat uncorrelated assets, we show that inter-asset correlations allow one to estimate apparent trends more reliably and to adjust the TF positions more efficiently. If properly accounted for, inter-asset correlations are not deteriorative but beneficial for portfolio management that can open new profit opportunities for trend followers. These concepts are illustrated using daily returns of three highly correlated futures markets: the E-mini S&P 500, Euro Stoxx 50 index, and the US 10-year T-note futures.

© 2015 Elsevier B.V. All rights reserved.

## 1. Introduction

For decades, market participants have attempted to detect potential trends in asset price fluctuations on exchange markets. In systematic trading, trend following (TF) strategies that generate buy or sell signals to adjust their market exposure according to past price variations, were developed to profit from trends at various time horizons [1–5]. While its actual profitability is highly controversial [6–9], trend following remains a widely used strategy among professional asset managers. Since many traders search for the same profit opportunities, the expected (net) gains are small, especially at short

\* Corresponding author.

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

times, and are subject to stochastic fluctuations. In order to enhance profit and reduce risk, fund managers build diversified portfolios, aiming to decorrelate constituent TF strategies as much as possible. Our goal is to show that this conventional approach leads to sub-optimal portfolios. In particular, we illustrate that inter-asset correlations, if accounted for properly, facilitate trend detection and thus significantly improve the risk-adjusted portfolio return.

In a previous work, we considered a linear TF strategy applied to an asset with auto-correlated returns [10]. In order to study trend following from a risk-reward perspective, we introduced an explicit persistence in asset returns by adding a stochastic trend term to a Gaussian market model. This simple framework allowed us to derive analytical formulas for the mean and variance of the strategy profit-and-losses (P&L). We were able to derive a threshold in auto-correlation below which a trend follower has no hope to realize profit given market transaction costs. Fund managers often use such criteria to select a set of assets/markets that are of interest for trend following trading. Many examples of TF strategies applied to stock markets, foreign exchange markets, and commodities were reported [2,11–13]. In the financial industry, diversified funds apply TF strategies to a large number of assets with hope to benefit from the so-called diversification effect [14,15].

In this paper we extend the model from Ref. [10] to the multivariate case. We introduce an asset covariance structure into a stochastic trend process and study how correlations between stochastic trends affect the portfolio risk-reward profile. We solve the portfolio optimization problem by taking into account the trend following nature of trading strategies. Our goal is to show that failure to account for trend correlations (i.e., only using asset returns covariance) leads to sub-optimal risk-adjusted portfolio return.

Modern portfolio theory [16–20], initiated by Markowitz [21], offers numerous solutions to the asset allocation problem [22]. The original problem was to find portfolio weights, i.e. amount of capital allocated to each asset, that maximize a portfolio mean-variance objective given expected market returns and covariance structure. In our approach, we consider a similar problem for TF strategies. The expected return of TF strategy depends directly on the asset expected excess variance (or auto-correlation) that characterizes market trends [10,23]. We solve the problem of static allocation of dynamic strategies by specifying the correlation structure of trend and noise components of asset price fluctuations.

We keep simple modeling assumptions from Ref. [10] in an effort to derive an exact form of the optimal allocation. We model price variations as a stochastic trend added to a white noise and we introduce separately the correlation in trends and the correlation in noises. For instance, two assets can exhibit similar long-term trends and be negatively correlated on the short term. Under these assumptions, we find that the static allocation problem in which an optimal weight is assigned to each asset, leads to sub-optimal risk-adjusted portfolio return. Even if the correlations in trends and noises are equal, the application of a classical Markowitz approach to TF strategies is sub-optimal. We then formulate a dynamic allocation framework that leads to an improved risk-adjusted return of the portfolio. Our approach to dynamic allocation consists in correcting each strategy signal by a linear combination of other strategy signals. This cross-correcting term can be seen as a lead-lag correction [24,25]. We show that the allocation problem for  $n$  dynamic strategies can be reduced to solving a static Markowitz problem for a set of  $n^2$  virtual assets with explicitly derived expected returns and covariance structure.

The paper is organized as follows. In Section 2, we introduce the standard mathematical tools to solve the portfolio allocation problem. In Section 3, we study in detail the two-asset portfolio problem while Section 4 extends to the case of multiple assets with identical correlations (e.g., a sector of the market). We quantify the improvement in terms of the expected Sharpe ratio (or risk-adjusted return) of the portfolio and the Sharpe gain compared to a static allocation scheme. In Section 5, we discuss some challenges in developing inference tools for estimating the model parameters from empirical data and illustrate the effect of the lead-lag correction onto the profitability of TF portfolios for futures markets. Conclusion section summarizes the main results, while technical derivations are reported in Appendices.

## 2. Market model and trading strategy

We first introduce a mathematical market model for  $n$  assets and describe linear trend following strategies. We then present the dynamic portfolio allocation based on a linear combination of strategy signals. In this frame, we derive mean and variance of portfolio returns, formulate the optimal allocation problem, and show its reduction to a standard static allocation problem for  $n^2$  virtual assets.

### 2.1. Market model

We assume that the return<sup>1</sup>  $r_t^j$  of the  $j$ th asset at time  $t$  has two contributions: an instantaneous fluctuation (noise)  $\varepsilon_t^j$ , and a stochastic trend which in general is given as a linear combination of random fluctuations  $\xi_{t'}^j$ ,

$$r_t^j = \varepsilon_t^j + \sum_{t'=1}^{t-1} \mathbf{A}_{t,t'}^j \xi_{t'}^j, \quad (1)$$

<sup>1</sup> Throughout this paper, daily price variations are called “returns” for the sake of simplicity. Rigorously speaking, we consider additive logarithmic returns resized by realized volatility which is a common practice on futures markets [26,27]. Although asset returns are known to exhibit various non-Gaussian features (so-called “stylized facts” [28–33]), resizing by realized volatility allows one to reduce, to some extent, the impact of changes in volatility and its correlations [34,35], and to get closer to the Gaussian hypothesis of returns [36].

where the matrix  $\mathbf{A}^j$  describes the stochastic trend of the  $j$ th asset, while  $\varepsilon_1^j, \dots, \varepsilon_t^j$  and  $\xi_1^j, \dots, \xi_t^j$  are two sets of independent Gaussian variables with mean zero and the following covariance structure:

$$\langle \varepsilon_t^j \varepsilon_{t'}^k \rangle = \delta_{t,t'} \mathbf{C}_\varepsilon^{j,k}, \quad \langle \xi_t^j \xi_{t'}^k \rangle = \delta_{t,t'} \mathbf{C}_\xi^{j,k}, \quad \langle \varepsilon_t^j \xi_{t'}^k \rangle = 0, \quad (2)$$

where  $\delta_{t,t'} = 1$  for  $t = t'$  and 0 otherwise. Here  $\mathbf{C}_\varepsilon$  and  $\mathbf{C}_\xi$  are the covariance matrices that describe inter-asset correlations of noises  $\varepsilon_t^j$  and of stochastic trend components  $\xi_t^j$ , respectively. The covariance matrix of Gaussian asset returns is then

$$\mathbf{C}_{t,t'}^{j,k} \equiv \langle r_t^j r_{t'}^k \rangle = \delta_{t,t'} \mathbf{C}_\varepsilon^{j,k} + \mathbf{C}_\xi^{j,k} (\mathbf{A}^j \mathbf{A}^{k,\dagger})_{t,t'}, \quad (3)$$

where  $\dagger$  denotes the matrix transposition. For each asset, the stochastic trend induces auto-correlations due to a linear combination of exogenous random variables  $\xi_t^j$  which are independent from short-time noises  $\varepsilon_t^j$ . Moreover, the structure of these *auto-correlations* (which is described by the matrix  $\mathbf{A}^j$ ) is considered to be independent from *inter-asset correlations* (which are described by matrices  $\mathbf{C}_\varepsilon$  and  $\mathbf{C}_\xi$ ). In particular, the covariance matrices  $\mathbf{C}_\varepsilon$  and  $\mathbf{C}_\xi$  do not depend on time. As discussed in Ref. [10], the presence of auto-correlations makes TF strategies profitable even for assets with zero mean returns. In other words, we consider asset auto-correlations as the origin of profitability of TF strategies. Although the whole analysis can be performed for nonzero mean returns, it is convenient to impose  $\langle \varepsilon_t^j \rangle = \langle \xi_t^j \rangle = \langle r_t^j \rangle = 0$  in order to accentuate the gain of the TF strategy over a simple buy-and-hold strategy (which is profitless in this case).

## 2.2. Profit-and-loss of a TF portfolio

The incremental profit-and-loss of a TF portfolio (i.e., the total return of the portfolio at time  $t$ ) is

$$\delta \mathcal{P}_t = \sum_{j=1}^n r_t^j S_{t-1}^j, \quad (4)$$

where  $S_{t-1}^j$  is the position<sup>2</sup> of the TF strategy on the  $j$ th asset at time  $t - 1$ . In a conventional setting, the position  $S_{t-1}^j$  is determined from earlier returns  $r_1^j, \dots, r_{t-1}^j$  of the  $j$ th asset. In this paper, we will show that this conventional choice is sub-optimal due to inter-asset corrections. To overcome this limitation, we introduce the position  $S_{t-1}^j$  as a weighted linear combination of the signals from all assets:

$$S_{t-1}^j = \sum_{k=1}^n \omega_{j,k} s^k(r_1^k, \dots, r_{t-1}^k), \quad (5)$$

where  $s^k(r_1^k, \dots, r_{t-1}^k)$  is the signal from the  $k$ th asset, with weights  $\omega_{j,k}$  to be determined. Note that the weights are considered to be time-independent, in coherence with the earlier assumption of time-independent inter-asset correlations. The incremental P&L of the portfolio becomes

$$\delta \mathcal{P}_t = \sum_{j,k=1}^n \omega_{j,k} r_t^j s^k(r_1^k, \dots, r_{t-1}^k), \quad (6)$$

where  $\omega_{j,k}$  can be interpreted as the weight of the  $k$ th signal onto the position of  $j$ th asset. The particular case of diagonal weights (when  $\omega_{j,k} = 0$  for  $j \neq k$ ) corresponds to a portfolio of  $n$  TF strategies with weights  $\omega_{j,j}$ . Therefore, the standard portfolio allocation problem is included in our framework, in which the diagonal weight  $\omega_{j,j}$  represents the amount of capital allocated to the  $j$ th asset. In general, non-diagonal terms allow one to benefit from inter-asset correlations to enhance the profitability of the TF portfolio.

Following Ref. [10], we consider a TF strategy whose signal is determined by a *linear* combination of earlier returns (e.g., an exponential moving average, see below):

$$s^k(r_1^k, \dots, r_{t-1}^k) = \sum_{t'=1}^{t-1} \mathbf{S}_{t,t'}^k r_{t'}^k, \quad (7)$$

so that

$$\delta \mathcal{P}_t = \sum_{j,k=1}^n \omega_{j,k} \sum_{t'=1}^{t-1} \mathbf{S}_{t,t'}^k r_t^j r_{t'}^k, \quad (8)$$

with given matrices  $\mathbf{S}_{t,t'}$ .

<sup>2</sup> The term “position” refers to the exposure or investment in a given asset. It is generally used in futures trading where position can be either positive (long) or negative (short) [37].

Using the Gaussian character of the model, we compute in [Appendix A](#) the mean and variance of the incremental profit-and-loss of a portfolio with  $n$  assets:

$$\begin{aligned}\langle \delta \mathcal{P}_t \rangle &= \sum_{j,k=1}^n \omega_{j,k} M_t^{j,k}, \\ \text{var}\{\delta \mathcal{P}_t\} &= \sum_{j_1,k_1,j_2,k_2=1}^n \omega_{j_1,k_1} \omega_{j_2,k_2} V_t^{j_1,k_1;j_2,k_2},\end{aligned}\tag{9}$$

where

$$\begin{aligned}M_t^{j,k} &= \mathbf{C}_{\xi}^{j,k} (\mathbf{S}^k \mathbf{A}^k \mathbf{A}^{j,\dagger})_{t,t}, \\ V_t^{j_1,k_1;j_2,k_2} &= \mathbf{C}_{\varepsilon}^{j_1,j_2} \mathbf{C}_{\varepsilon}^{k_1,k_2} (\mathbf{S}^{k_1} \mathbf{S}^{k_2,\dagger})_{t,t} + \mathbf{C}_{\varepsilon}^{j_1,j_2} \mathbf{C}_{\xi}^{k_1,k_2} (\mathbf{S}^{k_1} \mathbf{A}^{k_1} \mathbf{A}^{k_2,\dagger} \mathbf{S}^{k_2,\dagger})_{t,t} + \mathbf{C}_{\varepsilon}^{k_1,k_2} \mathbf{C}_{\xi}^{j_1,j_2} (\mathbf{S}^{k_1} \mathbf{S}^{k_2,\dagger})_{t,t} (\mathbf{A}^{j_1} \mathbf{A}^{j_2,\dagger})_{t,t} \\ &\quad + \mathbf{C}_{\xi}^{j_1,j_2} (\mathbf{A}^{j_1} \mathbf{A}^{j_2,\dagger})_{t,t} \mathbf{C}_{\xi}^{k_1,k_2} (\mathbf{S}^{k_1} \mathbf{A}^{k_1} \mathbf{A}^{k_2,\dagger} \mathbf{S}^{k_2,\dagger})_{t,t} + \mathbf{C}_{\xi}^{j_1,k_2} \mathbf{C}_{\xi}^{k_1,j_2} (\mathbf{S}^{k_1} \mathbf{A}^{k_1,\dagger} \mathbf{A}^{j_2})_{t,t} (\mathbf{S}^{k_2} \mathbf{A}^{j_1,\dagger} \mathbf{A}^{k_2})_{t,t}.\end{aligned}\tag{10}$$

The structural separation between auto-correlations and inter-asset corrections from Eq. [\(3\)](#) is also reflected in these formulas.

### 2.3. Optimization problem

Once the mean and variance of the incremental P&L are derived in the form [\(9\)](#), the dynamic allocation problem for a portfolio of trend following strategies is reduced to the standard optimization problem for a portfolio composed of  $n^2$  “virtual” assets (indexed by a double index  $j, k$ ) whose means are  $M_t^{j,k}$  and the covariance is  $V_t^{j_1,k_1;j_2,k_2}$ . One can therefore search for the weights  $\omega_{j,k}$  that optimize the chosen criterion (e.g., to minimize variance under a fixed expected return for the Markowitz theory). In this work, we search for the optimal weights  $\omega_{j,k}$  that maximize the squared Sharpe ratio (or squared risk-adjusted return of the portfolio):

$$\delta^2 \equiv \frac{\langle \delta \mathcal{P}_t \rangle^2}{\text{var}\{\delta \mathcal{P}_t\}} = \frac{(M_t^\dagger \omega)^2}{(\omega^\dagger V_t \omega)},\tag{11}$$

where the weights  $\omega_{j,k}$  are denoted here by a single vector  $\omega$  of size  $n^2$ . The optimization equations are obtained by setting

$$\frac{\partial \delta^2}{\partial \omega_{j,k}} = \frac{2(M_t^\dagger \omega)}{(\omega^\dagger V_t \omega)^2} [M_t^{j,k} (\omega^\dagger V_t \omega) - (V_t \omega)^{j,k} (M_t^\dagger \omega)] = 0 \quad (j, k = 1, \dots, n),\tag{12}$$

and we used the symmetry of the matrix  $V_t : V_t^{j_1,k_1;j_2,k_2} = V_t^{j_2,k_2;j_1,k_1}$ . More explicitly, these equations read

$$\sum_{j_1,k_1,j_2,k_2=1}^n [M_t^{j,k} V_t^{j_1,k_1;j_2,k_2} - V_t^{j,k;j_1,k_1} M_t^{j_2,k_2}] \omega_{j_1,k_1} \omega_{j_2,k_2} = 0\tag{13}$$

for all indices  $j, k = 1, \dots, n$ . In general, this is a set of  $n^2$  quadratic equations onto  $n^2$  unknown weights  $\omega_{j,k}$ . However, the original expressions for the mean and variance of  $\delta \mathcal{P}_t$  are invariant under the substitution of  $\omega_{j,k}$  by  $\omega_{k,j}$ . This is related to the linearity of the considered trend following strategy. In what follows, we consider the symmetric weights so that there remain  $n(n+1)/2$  unknown weights, with the same number of equations. For instance, one needs to solve 3, 6 and 10 equations for a portfolio with two, three and four assets, respectively. Note also that any solution of the above optimization problem is defined up to a multiplicative factor. In fact, the squared Sharpe ratio in Eq. [\(11\)](#) is invariant under multiplication of weights  $\omega_{j,k}$  by any nonzero constant. As a consequence,  $\omega_{j,k}$  should be interpreted as relative weights.

In general, the solution of Eqs. [\(13\)](#) depends on two covariance matrices  $\mathbf{C}_{\varepsilon}$  and  $\mathbf{C}_{\xi}$  (inter-asset correlations), matrices  $\mathbf{A}^j$  (asset auto-correlations), and matrices  $\mathbf{S}^j$  (signals of TF strategies). Once all these matrices are specified, the optimization problem can be solved numerically. However, it is impossible in practice to infer such a large number of parameters from market data, as well as to understand their influences on the optimal weights. For this reason, we further specify the problem in order to reduce the original, very large set of parameters. First, we choose in [Section 2.4](#) a particular form of auto-correlations (matrices  $\mathbf{A}^j$ ) and TF signals (matrices  $\mathbf{S}^j$ ). After that, we consider several particular forms of the covariances matrices  $\mathbf{C}_{\varepsilon}$  and  $\mathbf{C}_{\xi}$  for two-asset case and a sector model. In this way, we identify a relatively small number of the most relevant parameters and investigate their influence onto the optimal TF portfolio.

### 2.4. Exponential moving averages

In Ref. [\[10\]](#), we employed exponential moving averages (EMAs) to describe both stochastic trends and signals of TF strategies. This is equivalent to choosing stochastic trends as induced by a discrete Ornstein–Uhlenbeck process for

which

$$\mathbf{A}_{t,t'}^j = \begin{cases} \beta^j(1 - \lambda^j)^{t-t'-1}, & t > t', \\ 0, & t \leq t', \end{cases} \quad (14)$$

where  $\beta^j$  and  $\lambda^j$  are the strength and the rate of the  $j$ th stochastic trend. Similarly, the signal of a TF strategy is also chosen to be an EMA [38,39]:

$$\mathbf{S}_{t,t'}^j = \begin{cases} \gamma^j(1 - \eta^j)^{t-t'-1}, & t > t', \\ 0, & t \leq t', \end{cases} \quad (15)$$

where  $\gamma^j$  and  $\eta^j$  are the strength and the rate of the  $j$ th TF strategy. Setting the elements of these matrices to 0 for  $t \leq t'$  implements the causality: the trend and the signal at time  $t$  rely only upon the earlier returns with  $t' < t$ .

In what follows, we focus on the particular situation when the rates  $\lambda^j$  of all assets are identical ( $\lambda^j = \lambda$ ), and the rates  $\eta^j$  of all strategies are identical ( $\eta^j = \eta$ ). In the stationary limit  $t \rightarrow \infty$ , we derive in [Appendix A](#):

$$\begin{aligned} M_{\infty}^{j,k} &= \frac{q\sqrt{1-p^2}}{(1-pq)(1-q^2)} \mathbf{C}_{\xi,\beta}^{j,k}, \\ V_{\infty}^{j_1,k_1;j_2,k_2} &= \mathbf{C}_{\varepsilon}^{j_1,j_2} \mathbf{C}_{\varepsilon}^{k_1,k_2} + \frac{2\mathbf{C}_{\varepsilon}^{j_1,j_2} \mathbf{C}_{\xi,\beta}^{k_1,k_2}}{(1-pq)(1-q^2)} + \mathbf{C}_{\xi,\beta}^{j_1,j_2} \mathbf{C}_{\xi,\beta}^{k_1,k_2} \frac{1+q^2-2p^2q^2}{(1-pq)^2(1-q^2)^2}, \end{aligned} \quad (16)$$

where  $q = 1 - \lambda$ ,  $p = 1 - \eta$ , and we set  $\gamma^k = \gamma = \sqrt{1-p^2}$  as an appropriate normalization (see Ref. [10]).

In general, the optimal weights maximizing the squared Sharpe ratio can be found by solving numerically either the set of Eqs. (13), or the unconstrained maximization problem for  $\delta^2$  in Eq. (11). In order to understand the mechanisms behind the optimal TF strategy, we first focus on the particular case of two assets, for which many results can be derived analytically and then easily illustrated (Section 3). After that, we consider in Section 4 a sector model of  $n$  similar assets.

### 3. Two assets

For two assets, there are three independent weights:  $\omega_{11}$ ,  $\omega_{22}$ , and  $\omega_{12}$ . The covariance matrices take a simple form:

$$\mathbf{C}_{\varepsilon} = \begin{pmatrix} \sigma^1 \sigma^1 & \sigma^1 \sigma^2 \rho_{\varepsilon} \\ \sigma^1 \sigma^2 \rho_{\varepsilon} & \sigma^2 \sigma^2 \end{pmatrix}, \quad \mathbf{C}_{\xi} = \begin{pmatrix} 1 & \rho_{\xi} \\ \rho_{\xi} & 1 \end{pmatrix}, \quad (17)$$

where  $\rho_{\varepsilon}$  and  $\rho_{\xi}$  are two correlation coefficients (between inter-asset price noises  $\varepsilon_t^1$ ,  $\varepsilon_t^2$ , and stochastic trend components  $\xi_t^1$ ,  $\xi_t^2$ , respectively), while  $\sigma^1$  and  $\sigma^2$  are the volatilities of noises  $\varepsilon_t^j$ . Note that the volatility of stochastic trend components  $\xi_t^j$  can be included into auto-correlation strengths  $\beta^j$  that allows one to write a simplified form of the covariance matrix  $\mathbf{C}_{\xi}$ .

Substituting these relations in Eq. (16), we get explicit formulas for the mean and variance of the incremental profit-and-loss  $\delta \mathcal{P}_{\infty}$  in the stationary regime:

$$\begin{aligned} \langle \delta \mathcal{P}_{\infty} \rangle &= \frac{q\sqrt{1-p^2} [\beta_0^2]^2}{1-pq} [\kappa^2 \omega_{11} + 2\rho_{\xi}\kappa\omega_{12} + \omega_{22}], \\ \text{var}\{\delta \mathcal{P}_{\infty}\} &= [\sigma^2]^4 (\Omega_1 + 2\Omega_2/Q + R\Omega_3/Q^2) \end{aligned} \quad (18)$$

(with  $\beta_0^{1,2} = \beta^{1,2}/\sqrt{1-q^2}$ ) so that the squared Sharpe ratio becomes

$$\delta^2 = \frac{q^2(1-p^2)(\kappa^2 \omega_{11} + 2\rho_{\xi}\kappa\omega_{12} + \omega_{22})^2}{Q^2\Omega_1 + 2Q\Omega_2 + R\Omega_3}, \quad (19)$$

where

$$\begin{aligned} \Omega_1 &\equiv \nu^4 \omega_{11}^2 + 4\rho_{\varepsilon}\nu\omega_{12}[\nu^2\omega_{11} + \omega_{22}] + 2\nu^2\rho_{\varepsilon}^2\omega_{11}\omega_{22} + 2\nu^2(1+\rho_{\varepsilon}^2)\omega_{12}^2 + \omega_{22}^2, \\ \Omega_2 &\equiv \nu^2\kappa^2\omega_{11}^2 + 2\omega_{11}\omega_{12}\nu\kappa(\nu\rho_{\xi} + \kappa\rho_{\varepsilon}) + 2\omega_{11}\omega_{22}\nu\kappa\rho_{\varepsilon}\rho_{\xi} \\ &\quad + \omega_{12}^2(\nu^2 + 2\nu\kappa\rho_{\varepsilon}\rho_{\xi} + \kappa^2) + 2\omega_{12}\omega_{22}(\nu\rho_{\varepsilon} + \kappa\rho_{\xi}) + \omega_{22}^2, \\ \Omega_3 &\equiv \kappa^4\omega_{11}^2 + 4\rho_{\xi}\kappa\omega_{12}[\kappa^2\omega_{11} + \omega_{22}] + 2\kappa^2\rho_{\xi}^2\omega_{11}\omega_{22} + 2\kappa^2(1+\rho_{\xi}^2)\omega_{12}^2 + \omega_{22}^2, \end{aligned} \quad (20)$$

and

$$Q \equiv \frac{(1-pq)[\sigma^2]^2}{[\beta_0^2]^2}, \quad R \equiv 1 + q^2 - 2p^2q^2, \quad \kappa \equiv \frac{\beta^1}{\beta^2}, \quad \nu \equiv \frac{\sigma^1}{\sigma^2}. \quad (21)$$

With no loss of generality, we assume that  $\beta_1 \leq \beta_2$ , i.e.  $\kappa \leq 1$ .

As discussed in Section 2, the optimization procedure to maximize  $\delta^2$  leads to three quadratic equations on weights  $\omega_{11}$ ,  $\omega_{12}$ , and  $\omega_{22}$ . Defining two independent ratios,

$$z = \frac{\omega_{11}}{\omega_{22}}, \quad x = \frac{\omega_{12}}{\omega_{22}}, \quad (22)$$

one gets three equations containing terms  $z^2$ ,  $x^2$ ,  $zx$ ,  $z$ ,  $x$ , and constants. The equation  $\frac{\partial \delta^2}{\partial \omega_{11}} = 0$  does not contain the term  $z^2$ , while the equation  $\frac{\partial \delta^2}{\partial \omega_{12}} = 0$  does not contain the term  $x^2$ . Taking appropriate linear combinations, one can express and then eliminate the term  $zx$ . Finally, one would deal with a single fourth degree equation. Although an explicit analytical solution of this equation is possible, it is too cumbersome for any practical use. In turn, the original problem of maximizing the squared Sharpe ratio can be solved numerically as a standard minimization problem.

The squared Sharpe ratio in Eq. (19) depends on the following parameters of the model: two rates  $\lambda (=1-q)$  and  $\eta (=1-p)$  of the EMAs for stochastic trends and for TF strategies; two asset volatilities  $\sigma^1$  and  $\sigma^2$ ; two auto-correlation strengths  $\beta^1$  and  $\beta^2$ ; and two correlation coefficients  $\rho_\varepsilon$  and  $\rho_\xi$ . In order to illustrate and discuss various features of the optimal solution, we consider several particular cases of practical interest for which explicit analytical solutions are relatively simple.

### 3.1. Uncorrelated assets ( $\rho_\varepsilon = \rho_\xi = 0$ )

We first consider the case of two uncorrelated assets:  $\rho_\varepsilon = \rho_\xi = 0$ . In this case, the condition  $\frac{\partial \delta^2}{\partial \omega_{12}} = 0$  leads to the equation  $(\nu^2 Q^2 + Q(\nu^2 + \kappa^2) + \kappa^2 R)xz = 0$ . Since the coefficient in front of  $xz$  is strictly positive, one has either  $x = 0$ , or  $z = 0$ . The second option ( $z = 0$ ) does not satisfy other equations while the former case yields

$$x_{\text{opt}} = \frac{\omega_{12}}{\omega_{22}} = 0, \quad z_{\text{opt}} = \frac{\omega_{11}}{\omega_{22}} = \frac{\kappa^2(Q^2 + 2Q + R)}{\nu^4 Q^2 + 2\nu^2 \kappa^2 Q + \kappa^4 R}, \quad (23)$$

with the squared optimal Sharpe ratio

$$\delta_{\text{opt}}^2 = q^2(1-p^2) \left[ \frac{1}{Q^2 + 2Q + R} + \frac{\kappa^4}{\nu^4 Q^2 + 2\nu^2 \kappa^2 Q + \kappa^4 R} \right]. \quad (24)$$

Since two assets are uncorrelated, no additional information can be gained by including lead-lag term so that  $\omega_{12} = 0$ , in agreement with Eq. (23). In turn,  $z_{\text{opt}}$  determines the weights of two assets in the optimal portfolio, up to a multiplicative factor. Adding a constrain  $\omega_{11} + \omega_{22} = 1$ , one can identify  $\omega_{11}$  and  $\omega_{22}$  are relative weights.

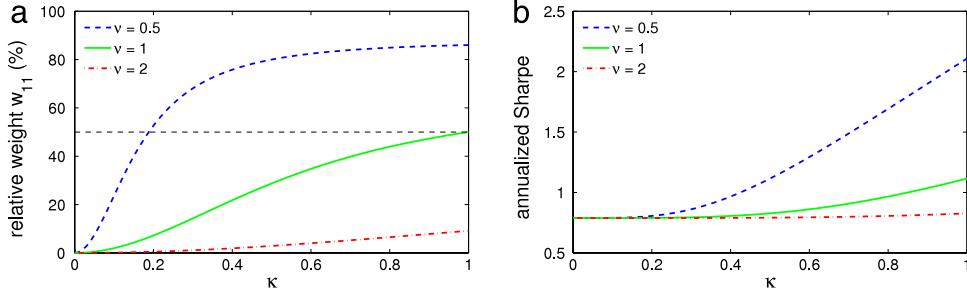
When two assets have identical characteristics (i.e.,  $\sigma^1 = \sigma^2$  and  $\beta^1 = \beta^2$  from which  $\nu = \kappa = 1$ ), Eq. (23) yields  $z_{\text{opt}} = 1$ , i.e., both assets enter with equal weights as expected. In this case, the squared Sharpe ratio of the optimal portfolio is twice larger than the squared Sharpe ratio of each asset:  $\delta_{\text{opt}}^2(\kappa = 1) = 2 \frac{q^2(1-p^2)}{Q^2 + 2Q + R}$ , in agreement with the principle of diversification.

In the limit  $\kappa = 0$ , the first asset has no auto-correlation ( $\beta^1 = 0$ ) so that the underlying TF strategy is profitless and therefore excluded from the portfolio:  $\omega_{11} = z_{\text{opt}} = 0$ . One retrieves the squared Sharpe ratio of the second asset:  $\delta_{\text{opt}}^2(\kappa = 0) = \frac{q^2(1-p^2)}{Q^2 + 2Q + R}$ . Fig. 1 shows the relative weight  $\omega_{11} = 100\% \frac{z_{\text{opt}}}{1+z_{\text{opt}}}$  of the first asset in the optimal portfolio and the annualized<sup>3</sup> optimal Sharpe ratio as functions of  $\kappa$  varying from 0 to 1. For two assets with the same volatility (i.e.,  $\nu = 1$ ), the relative weight  $\omega_{11}$  of the less profitable first asset varies from 0 to 50%, while the squared Sharpe ratio doubles, as expected. When the first (less profitable) asset is in addition twice more volatile ( $\nu = 2$ ), its relative weight does not exceed 10% even at  $\kappa = 1$ , while the Sharpe ratio has almost not improved (dash-dotted lines). In turn, if the less profitable asset is twice less volatile ( $\nu = 0.5$ ), its relative weight rapidly grows and exceeds 50% at  $\kappa \simeq 0.18$ . The inclusion of the less volatile asset greatly improves the Sharpe ratio. Using Eqs. (23), (24), one can investigate the relative impacts of volatilities and auto-correlations onto the optimal portfolio for uncorrelated assets.

### 3.2. Indistinguishable correlated assets ( $\kappa = \nu = 1$ )

In order to reveal the role of inter-asset corrections, we consider two assets with the same structure of auto-correlations (i.e.,  $\kappa = \nu = 1$ ) that makes them indistinguishable from each other. Since each asset offers the same expected TF returns,

<sup>3</sup> Since we investigate the incremental profit-and-loss  $\delta P_t$ , the Sharpe ratio  $\delta$  in Eq. (11) characterizes the risk-adjusted return of the portfolio over one time step of TF strategies. In many trading platforms, TF strategies are realized on a daily basis (i.e., they are updated once per day). At the same time, it is conventional to rescale the daily Sharpe ratio  $\delta$  to the annualized Sharpe ratio  $\sqrt{255} \delta$ , where  $\sqrt{255}$  is the standard pre-factor matching a calendar year of 255 business days. For a quick grasp of this rescaling, one can use as a reference point an asset delivering a 10% annual return with 10% annualized volatility, corresponding to a Sharpe ratio of 1. In the systematic hedge industry, only a few funds deliver the Sharpe ratio above 1 over the long run [27,40,41].



**Fig. 1.** The relative weight  $\omega_{11} = 100\% \frac{z_{\text{opt}}}{1+z_{\text{opt}}}$  of the first asset in the optimal portfolio (a) and the annualized optimal Sharpe ratio  $\sqrt{255} \delta_{\text{opt}}$  (b), versus  $\kappa = \beta_0^1/\beta_0^2$ , for uncorrelated assets (i.e.,  $\rho_\varepsilon = \rho_\xi = 0$ ), with  $\beta_0^2 = 0.1$ ,  $\beta_0^1 = \kappa \beta_0^2$ ,  $\lambda = \eta = 0.01$ ,  $\sigma^2 = 1$ , and  $\sigma^1 = \nu \sigma^2$ , with  $\nu = 0.5, 1, 2$ .

the diagonal weights are expected to be identical:  $\omega_{11} = \omega_{22}$ . We study therefore the optimal value of the lead-lag correction  $\omega_{12}$  (which is the same for both assets), and the gain in the Sharpe ratio that this cross-correcting term brings to the optimal portfolio.

In Appendix B, we derive the optimal solution of this minimization problem:  $z_{\text{opt}} = 1$ ,

$$x_{\text{opt}} = -\frac{Q^2(2\rho_\varepsilon - \rho_\xi - \rho_\varepsilon^2 \rho_\xi) + 2Q\rho_\varepsilon(1 - \rho_\xi^2) + R\rho_\xi(1 - \rho_\xi^2)}{Q^2(1 - 2\rho_\varepsilon \rho_\xi + \rho_\varepsilon^2) + 2Q(1 - \rho_\xi^2) + R(1 - \rho_\xi^2)}, \quad (25)$$

and the squared optimal Sharpe ratio is

$$\begin{aligned} \delta_{\text{opt}}^2 &= 2q^2(1 - p^2) \left[ (1 - \rho_\xi^2) \left( Q^2(1 - \rho_\varepsilon^2) + 2Q(1 - \rho_\varepsilon \rho_\xi) + R(1 - \rho_\xi^2) \right) \right. \\ &\quad \left. + 2Q^2(\rho_\varepsilon - \rho_\xi)^2 \right] \left\{ \left( Q^2(1 - \rho_\varepsilon^2) + 4Q(1 - \rho_\varepsilon \rho_\xi) + R(1 - \rho_\xi^2) \right) \right. \\ &\quad \times \left. \left( Q^2(1 - \rho_\varepsilon^2) + R(1 - \rho_\xi^2) \right) + 4Q^2(1 - \rho_\varepsilon^2)(1 - \rho_\xi^2) + 4Q^2R(\rho_\varepsilon - \rho_\xi)^2 \right\}^{-1}. \end{aligned} \quad (26)$$

In the typical situation,  $R \ll Q$  (see Eq. (21)), and if  $\rho_\varepsilon$  is not too close to 1, one can neglect terms with  $R$  in order to get a simpler approximate relation:

$$\delta_{\text{opt}}^2 \simeq 2q^2(1 - p^2) \frac{1 - \rho_\xi^2}{1 - \rho_\varepsilon^2} \frac{Q[1 - \rho_\varepsilon^2 + 2(\rho_\varepsilon - \rho_\xi)^2] + 2(1 - \rho_\varepsilon \rho_\xi)}{Q^2(1 - \rho_\varepsilon^2) + 4Q(1 - \rho_\varepsilon \rho_\xi) + 4(1 - \rho_\xi^2)}. \quad (27)$$

In order to assess the gain of including the lead-lag term, one can compare  $\delta_{\text{opt}}$  with the Sharpe ratio  $\delta_0$  from Eq. (19) with  $\omega_{11} = \omega_{22}$  and  $\omega_{12} = 0$  (i.e., without lead-lag term):

$$\delta_0^2 = \frac{q^2(1 - p^2)}{Q^2(1 + \rho_\varepsilon^2) + 2Q(1 + \rho_\varepsilon \rho_\xi) + R(1 + \rho_\xi^2)}. \quad (28)$$

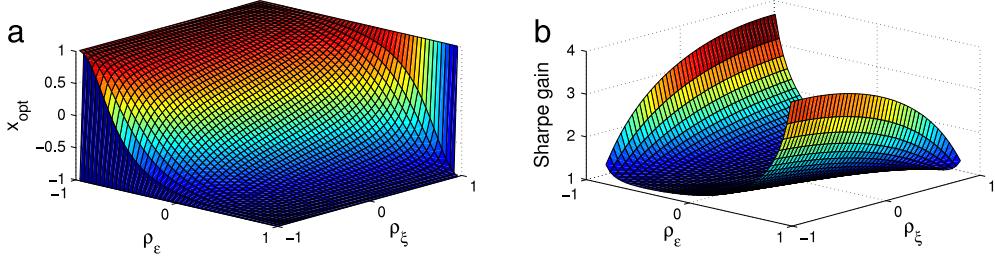
**Fig. 2** shows how  $x_{\text{opt}}$  and the Sharpe gain  $\delta_{\text{opt}}/\delta_0$  of the optimal portfolio depend on two correlation coefficients  $\rho_\varepsilon$  and  $\rho_\xi$ . As expected, the highest gain can be achieved when  $\rho_\varepsilon \rho_\xi$  is close to  $-1$ , e.g., when stochastic trends are strongly correlated ( $\rho_\xi \simeq 1$ ) while noises are strongly anti-correlated ( $\rho_\varepsilon \simeq -1$ ). This extreme situation illustrates the need to distinguish correlations in trends and noises. Although the extreme cases with  $\rho_\varepsilon \rho_\xi \simeq -1$  are not realistic for financial markets, some profit opportunities can be overlooked in the conventional portfolio management in which there is no explicit separation of the two effects in inter-asset correlations. At the same time, a reliable estimation of two correlation coefficients from financial data remains challenging.

In order to better grasp the behavior of  $x_{\text{opt}}$  and  $\delta_{\text{opt}}$ , we consider several particular cases:

- (i) when  $\rho_\xi = \pm 1$  (fully correlated stochastic trends), one gets  $x_{\text{opt}} = \pm 1$ , independently of  $\rho_\varepsilon$ . As seen on Fig. 2(a), this behavior is unstable: it is sufficient to take  $|\rho_\xi|$  slightly smaller than 1 to retrieve the dependence of  $x_{\text{opt}}$  on  $\rho_\varepsilon$ .
- (ii) when  $\rho_\varepsilon = \pm 1$  (fully correlated noises), one gets  $x_{\text{opt}} \simeq \pm 1$ , with a very weak dependence on  $\rho_\xi$ . In contrast to the above case, this behavior persists for all  $\rho_\varepsilon$  near  $\pm 1$  (Fig. 2(a)).
- (iii) when  $\rho_\varepsilon = 0$  (uncorrelated noises), one finds

$$x_{\text{opt}} = \rho_\xi \frac{Q^2 - R(1 - \rho_\xi^2)}{Q^2 + (2Q + R)(1 - \rho_\xi^2)} \approx \frac{\rho_\xi}{1 + (1 - \rho_\xi^2) \frac{R}{Q}}, \quad (29)$$

where the small term  $R$  was neglected in the second relation. Here one can see the impact of correlated trends on apparently uncorrelated markets. Note that the lead-lag correction has the same sign as the trend correlation coefficient  $\rho_\xi$ .



**Fig. 2.** The optimal lead-lag weight ratio  $x_{\text{opt}} = \omega_{12}/\omega_{22}$  (a) and the Sharpe gain  $s_{\text{opt}}/s_0$  (b) as functions of two correlation coefficients  $\rho_\epsilon$  and  $\rho_\xi$  for two indistinguishable assets with  $\beta_0^1 = \beta_0^2 = 0.1$ ,  $\sigma^1 = \sigma^2 = 1$ , and  $\lambda = \eta = 0.01$ . For the second plot, both correlation coefficients vary from  $-0.9$  to  $0.9$  in order to exclude unrealistically large Sharpe gains at  $\rho_\epsilon$  and  $\rho_\xi$  around  $\pm 1$ .

(iv) when  $\rho_\xi = 0$  (uncorrelated stochastic trends), one gets

$$x_{\text{opt}} = -\rho_\epsilon \frac{2Q(Q+1)}{Q^2(1+\rho_\epsilon^2) + 2Q + R} \approx \frac{-\rho_\epsilon}{1 - (1 - \rho_\epsilon^2) \frac{Q}{2(Q+1)}}, \quad (30)$$

where the small term with  $R$  was neglected in the second relation. One can see the impact of correlated noises on a market with independent trend components. In contrast to the above case, the lead-lag correction has the opposite sign of the noise correlation coefficient  $\rho_\epsilon$ .

(v) when  $\rho_\epsilon = \rho_\xi = \rho$ , one gets  $x_{\text{opt}} = -\rho$ , i.e., the position of the first asset should be reduced by a relative amount  $\rho$  of the second asset in order to maximally decorrelate them. In addition, the squared optimal Sharpe ratio does not depend on correlations:

$$s_{\text{opt}}^2 = \frac{2q^2(1-p^2)}{Q^2 + 2Q + R}. \quad (31)$$

In other words, such correlations cannot improve the optimal Sharpe ratio but one needs to correct the TF strategy to remove the effect of correlations. This case is particularly interesting as it helps to show that the static allocation is sub-optimal without introducing the lead-lag correction.

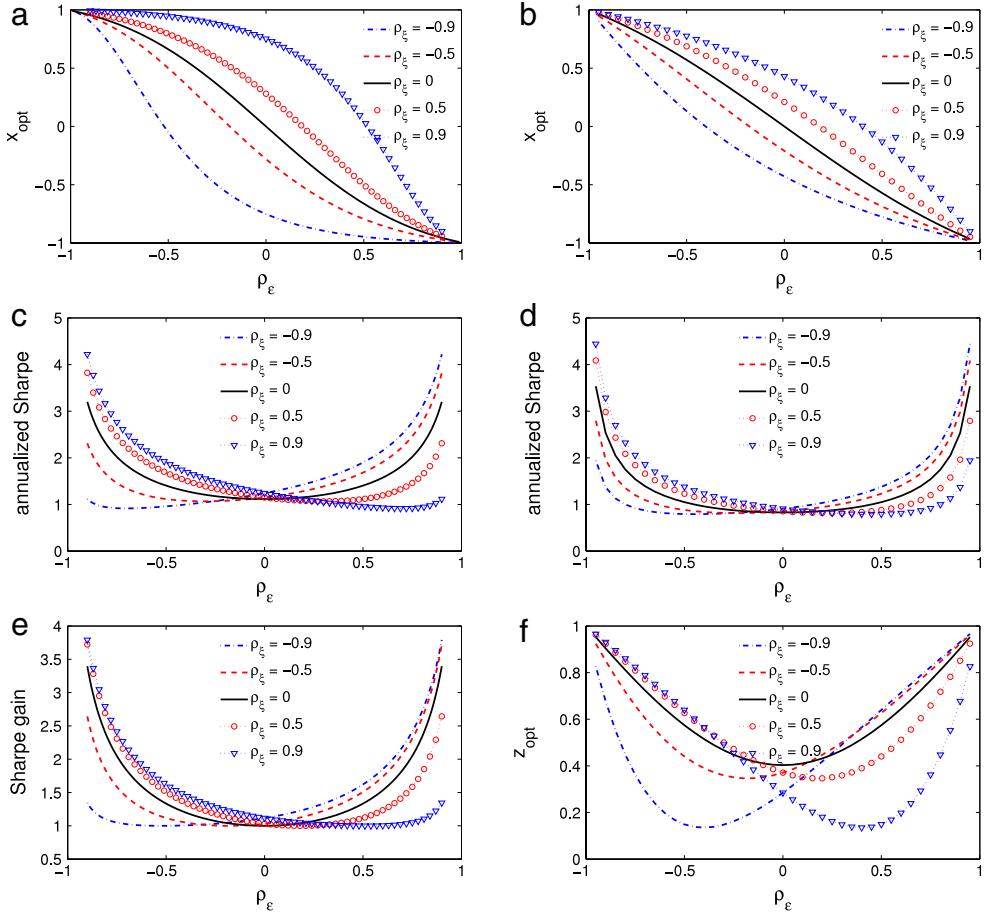
The plots in the left column of Fig. 3 further illustrate some properties of the optimal portfolio of TF strategies on two indistinguishable assets for different correlation coefficients  $\rho_\epsilon$  and  $\rho_\xi$ . Both assets are traded with identical weights,  $\omega_{11} = \omega_{22}$ , i.e.  $z_{\text{opt}} = 1$ . The optimal lead-lag correction  $x_{\text{opt}} = \omega_{12}/\omega_{22}$  from Eq. (25) monotonously decreases with  $\rho_\epsilon$  from 1 at  $\rho_\epsilon = -1$  (fully anticorrelated noises) to  $-1$  at  $\rho_\epsilon = 1$  (fully correlated noises), as shown on Fig. 3(a). The rate of decrease depends on  $\rho_\xi$ . As expected, no correction is needed when  $\rho_\epsilon = \rho_\xi = 0$ . Generally, the horizontal line at  $x_{\text{opt}} = 0$  determines the set of correlation coefficients for which no lead-lag correction term is needed.

Fig. 3(c) shows how the annualized optimal Sharpe ratio  $\sqrt{255} s_{\text{opt}}$  changes with correlation coefficients  $\rho_\epsilon$  and  $\rho_\xi$ . When  $\rho_\epsilon = 0$ , the annualized Sharpe ratio is close 1 (for the chosen level  $\beta_0 = 0.1$  of auto-correlations), and it depends weakly on stochastic trend correlations ( $\rho_\xi$ ). In turn, this ratio is strongly enhanced when  $|\rho_\epsilon| > 0.5$ , and the enhancement occurs when  $\rho_\epsilon$  and  $\rho_\xi$  are of opposite signs. In contrast, the annualized optimal Sharpe ratio may be decreased when both correlations are of the same sign. Finally, the enhancement is even stronger when  $|\rho_\epsilon|$  is close to 1, independently of mutual signs of  $\rho_\epsilon$  and  $\rho_\xi$ . However, this region is unrealistic for financial markets.

Fig. 3(e) illustrates the Sharpe gain  $s_{\text{opt}}/s_0$  due to inclusion of the lead-lag correction term. This gain is substantial for highly correlated assets (when  $|\rho_\epsilon|$  and/or  $|\rho_\xi|$  are large). As expected, the gain is always greater than (or equal to) 1, as  $s_{\text{opt}}$  is the optimal solution.

### 3.3. Example of two distinct assets: $\kappa = 0.5$

For comparison, we present on the right column of Fig. 3 similar quantities for two assets with different stochastic trends ( $\beta_0^1 = 0.05$ ,  $\beta_0^2 = 0.1$ , i.e.,  $\kappa = 0.5$ ). In this example, the first asset with lower auto-correlations is less profitable for TF strategies. Although the inclusion of the first asset does not improve the mean profit-and-loss, it still allows to increase the Sharpe ratio by reducing the variance due to diversification. Fig. 3(f) shows the optimal weights ratio  $z_{\text{opt}} = \omega_{11}/\omega_{22}$ . When there is no correlations ( $\rho_\epsilon = \rho_\xi = 0$ ), Eq. (23) yields  $z_{\text{opt}} \approx 0.40$ , i.e., two assets enter with relative weights 28.6% and 71.4%, respectively (for chosen parameters). For correlated assets, the optimal weights ratio can be either smaller or larger than 0.4. For instance, when both  $\rho_\epsilon$  and  $\rho_\xi$  are of the same sign, the relative weight of the first asset can be reduced to almost zero ( $z_{\text{opt}}$  is close to 0). In contrast, when  $|\rho_\epsilon|$  is close to 1, both assets have to be included with almost the same weights ( $z_{\text{opt}}$  approaches 1).



**Fig. 3.** Comparison between optimal portfolios for two indistinguishable assets ( $\beta_0^1 = \beta_0^2 = 0.1, \kappa = 1$ , left) and two assets with different stochastic trends ( $\beta_0^1 = 0.05, \beta_0^2 = 0.1, \kappa = 0.5$ , right): the optimal lead-lag correction  $x_{\text{opt}} = \omega_{12}/\omega_{22}$  (a, b), the annualized optimal Sharpe ratio  $\sqrt{255} s_{\text{opt}}$  (c, d), the gain in the Sharpe ratio  $s_{\text{opt}}/s_0$  due to the lead-lag correction (e), and the optimal asset weights ratio  $z_{\text{opt}} = \omega_{11}/\omega_{22}$  (f) (note that  $z_{\text{opt}} = 1$  for indistinguishable assets on the left). These quantities are presented as functions of  $\rho_\epsilon$  for different  $\rho_\xi$ , and we set  $\sigma^1 = \sigma^2 = 1$  (i.e.,  $v = 1$ ) and  $\lambda = \eta = 0.01$ .

Fig. 3(b), (d) show the optimal lead-lag correction  $x_{\text{opt}}$  and the annualized optimal Sharpe ratio. Both quantities exhibit similar features as in the case of indistinguishable assets (left column). The Sharpe gain is also similar to the earlier case (not shown).

#### 4. A sector model

The above analysis can be extended to multiple assets. Although the theoretical solution is formally available, a large number of weights (growing as  $n(n + 1)/2$ ) makes challenging its investigation in general. At the same time, one can still perform numerical minimization to determine the optimal solution by standard optimization tools. In this section, we consider the particular case of a market sector when inter-asset correlation is the same for all assets [14]. We also assume that all returns are normalized by realized volatilities, i.e.,  $\sigma^j = 1$ . In other words, we consider the covariance matrices for noises and stochastic trends to be

$$\mathbf{C}_\epsilon = \begin{pmatrix} 1 & \rho_\epsilon & \rho_\epsilon & \cdots & \rho_\epsilon \\ \rho_\epsilon & 1 & \rho_\epsilon & \cdots & \rho_\epsilon \\ \rho_\epsilon & \rho_\epsilon & 1 & \cdots & \rho_\epsilon \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \rho_\epsilon & \rho_\epsilon & \rho_\epsilon & \cdots & 1 \end{pmatrix}, \quad \mathbf{C}_\xi = \begin{pmatrix} 1 & \rho_\xi & \rho_\xi & \cdots & \rho_\xi \\ \rho_\xi & 1 & \rho_\xi & \cdots & \rho_\xi \\ \rho_\xi & \rho_\xi & 1 & \cdots & \rho_\xi \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \rho_\xi & \rho_\xi & \rho_\xi & \cdots & 1 \end{pmatrix}. \quad (32)$$

Each of these matrices has two eigenvalues,  $\nu_1 = 1 + (n - 1)\rho$  and  $\nu_2 = 1 - \rho$ . Since covariance matrices must be positive definite, all eigenvalues should be non-negative that implies  $\rho_\epsilon \geq -1/(n - 1)$  and  $\rho_\xi \geq -1/(n - 1)$ . In what follows, we only consider  $\rho_\epsilon \geq 0$  and  $\rho_\xi \geq 0$ .

When the assets are indistinguishable (i.e.,  $\beta^j = \beta$ ), they are expected to have the same weights in the optimal portfolio of TF strategies,  $\omega_{jj} = \omega_{11}$ , as well as all lead-lag corrections are the same:  $\omega_{jk} = \omega_{12}$  for all  $j \neq k$ . For this particular case of indistinguishable assets, the number of unknown weights is reduced from  $n(n + 1)/2$  to 2 that allows one to derive an analytical solution and to investigate its behavior as a function of the number of assets.

In Appendix C, we derived the optimal solution:

$$x_{\text{opt}} = \frac{\omega_{12}}{\omega_{11}} = -\frac{V_2 - (n - 1)\rho_{\xi}V_1}{V_3 - (n - 1)\rho_{\xi}V_2}, \quad (33)$$

$$\delta_{\text{opt}}^2 = nq^2(1 - p^2) \frac{(n - 1)^2\rho_{\xi}^2V_1 - 2\rho_{\xi}(n - 1)V_2 + V_3}{V_1V_3 - V_2^2} \quad (34)$$

where

$$\begin{aligned} V_1 &= Q^2(1 + (n - 1)\rho_{\xi}^2) + 2Q(1 + (n - 1)\rho_{\xi}\rho_{\varepsilon}) + R(1 + (n - 1)\rho_{\xi}^2), \\ V_2 &= (n - 1)[Q^2(2\rho_{\xi} + (n - 2)\rho_{\varepsilon}^2) + 2Q(\rho_{\xi} + \rho_{\varepsilon} + (n - 2)\rho_{\xi}\rho_{\varepsilon}) + R(2\rho_{\xi} + (n - 2)\rho_{\xi}^2)], \\ V_3 &= n[Q^2(1 + (n - 1)\rho_{\xi})^2 + 2Q(1 + (n - 1)\rho_{\xi})(1 + (n - 1)\rho_{\varepsilon}) + R(1 + (n - 1)\rho_{\xi})^2] - V_1 - 2V_2. \end{aligned} \quad (35)$$

Although these formulas are explicit, they are rather cumbersome for theoretical analysis. For  $n = 2$ , one retrieves the results of Section 3.2 for two indistinguishable assets. In what follows, we consider several particular cases in order to illustrate the main features of the optimal solution and the role of the portfolio size (number of assets).

#### 4.1. Conventional trading

It is instructive to start with a “conventional” trading without lead-lag correction term ( $\omega_{12} = 0$ ) for which Eq. (C.4) yields the squared Sharpe ratio

$$\delta_n^2 = \frac{nq^2(1 - p^2)}{Q^2(1 + (n - 1)\rho_{\xi}^2) + 2Q(1 + (n - 1)\rho_{\xi}\rho_{\varepsilon}) + R(1 + (n - 1)\rho_{\xi}^2)}. \quad (36)$$

For  $n = 1$ , one retrieves the squared Sharpe ratio of a single asset:

$$\delta_1^2 = \frac{q^2(1 - p^2)}{Q^2 + 2Q + R}. \quad (37)$$

If there was no inter-asset correlation ( $\rho_{\xi} = \rho_{\varepsilon} = 0$ ), the squared Sharpe ratio  $\delta_n^2$  for  $n$  assets would be  $n$  times larger than  $\delta_1^2$  for a single asset:  $\delta_n^2 = n\delta_1^2$ , as expected due to diversification. In this uncorrelated case, one also finds the optimal solution to be

$$x_{\text{opt}} = 0, \quad \delta_{\text{opt}}^2 = \delta_n^2(\rho_{\xi} = \rho_{\varepsilon} = 0) = n\delta_1^2. \quad (38)$$

The presence of correlations reduces the effect of diversification and diminishes  $\delta_n$ . Moreover, this reduction is stronger for large  $n$ . In what follows, we show that inclusion of the lead-lag term allows one to recover or even further enhance the Sharpe ratio. In other words, although strong correlations may reduce diversification, their proper accounting makes the underlying assets even more profitable.

#### 4.2. Equal trend and noise correlations ( $\rho_{\xi} = \rho_{\varepsilon}$ )

For  $\rho_{\xi} = \rho_{\varepsilon} = \rho$ , we get

$$\begin{aligned} V_1 &= (1 + (n - 1)\rho^2)[Q^2 + 2Q + R], \\ V_2 &= (n - 1)(2\rho + (n - 2)\rho^2)[Q^2 + 2Q + R], \\ V_3 &= (n - 1)(1 + 2(n - 2)\rho + \rho^2(n^2 - 3n + 3))[Q^2 + 2Q + R]. \end{aligned}$$

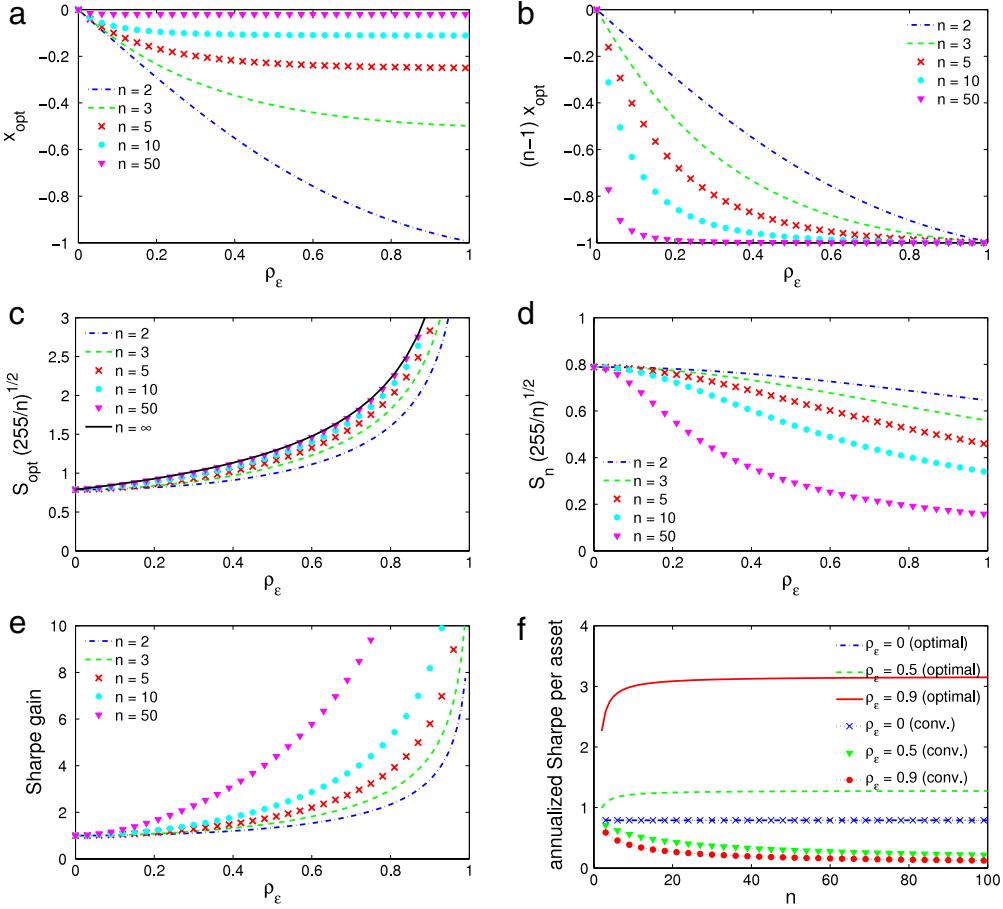
Substituting these expressions into Eq. (34), we deduce

$$x_{\text{opt}} = -\frac{\rho}{1 + (n - 2)\rho}, \quad \delta_{\text{opt}}^2 = n \frac{q^2(1 - p^2)}{Q^2 + 2Q + R} = n\delta_1^2. \quad (39)$$

As for the case of two assets, the optimal Sharpe ratio does not depend on correlations, while the lead-lag term does depend on  $\rho$ . As previously, one can compare the squared optimal Sharpe ratio to  $\delta_n^2$  (i.e., the case without lead-lag correction):

$$\frac{\delta_{\text{opt}}^2}{\delta_n^2} = 1 + (n - 1)\rho^2. \quad (40)$$

One can see that accounting for correlations by inclusion of the lead-lag term may significantly increase the squared Sharpe ratio. This effect obviously disappears at  $n = 1$ . Note that even weak correlations can be enhanced by including many assets



**Fig. 4.** The optimal lead-lag correction  $x_{\text{opt}}$  (a),  $(n-1)x_{\text{opt}}$  (b), the annualized optimal Sharpe ratio per asset,  $\sqrt{255} \delta_{\text{opt}}/\sqrt{n}$  (c), the annualized conventional Sharpe ratio per asset,  $\sqrt{255} \delta_n/\sqrt{n}$  (d), and the Sharpe gain  $\delta_{\text{opt}}/\delta_n$  (e), as functions of  $\rho_\varepsilon$ , for  $n$  indistinguishable assets, with  $\rho_\xi = 0$ ,  $\beta_0^j = 0.1$ ,  $\lambda = \eta = 0.01$ . The last plot (f) shows the annualized optimal Sharpe ratio per asset  $\sqrt{255} \delta_{\text{opt}}/\sqrt{n}$  (lines) and the annualized conventional Sharpe ratio per asset  $\sqrt{255} \delta_n/\sqrt{n}$  (symbols) as functions of  $n$ .

in a portfolio. This secondary effect (in addition to diversification that increases  $\delta_n$ ) favors large portfolios, in agreement with a common practice of fund managers.

#### 4.3. Uncorrelated stochastic trends ( $\rho_\xi = 0$ )

When  $\rho_\xi = 0$ , Eqs. (33), (34) yield

$$x_{\text{opt}} = -\frac{V_2}{V_3}, \quad \delta_{\text{opt}}^2 = nq^2(1-p^2) \frac{V_3}{V_1 V_3 - V_2^2}. \quad (41)$$

Fig. 4 illustrates the dependence of the lead-lag correction  $x_{\text{opt}}$  and the annualized optimal Sharpe ratio  $\sqrt{255} \delta_{\text{opt}}$  on  $\rho_\varepsilon$ . As expected, negative optimal lead-lag corrections are needed to reduce positive noise correlations. Larger  $n$  require smaller correction amplitude  $|x_{\text{opt}}|$  (Fig. 4(a)). At the same time, each asset has  $n-1$  identical lead-lag corrections (from other  $n-1$  assets) so that the total correction appears as  $(n-1)x_{\text{opt}}$  (Fig. 4(b)). For large  $n$ , the total correction rapidly reaches the level -1, even for relatively small  $\rho_\varepsilon$ . In other words, when a large number of assets is traded, even small inter-asset correlations, if ignored, can significantly reduce the Sharpe ratio. One needs therefore to include lead-lag corrections. In order to understand the rapid approach to the limiting value -1, one can expand  $(n-1)x_{\text{opt}}$  in terms of a small parameter  $1/(n-1)$  in the limit of large  $n$  as

$$(n-1)x_{\text{opt}} \simeq -1 + \frac{Q^2(1-\rho_\varepsilon)^2 + 2Q(1-\rho_\varepsilon) + R}{Q^2\rho_\varepsilon^2(n-1)^2} + O\left(\frac{1}{(n-1)^3}\right). \quad (42)$$

Note that the first correction term here is of the order of  $1/(n-1)^2$ .

**Fig. 4(c)** shows the annualized optimal Sharpe ratio normalized by the number of assets,  $\sqrt{255} \delta_{\text{opt}}/\sqrt{n}$ , as a function of  $\rho_\varepsilon$ . When there is no correlation ( $\rho_\varepsilon = 0$ ), this quantity does not depend on  $n$ , as expected from Eq. (38), and all curves come to the same point  $\sqrt{255} \delta_1 \approx 0.7885$  (for the chosen set of parameters). When  $\rho_\varepsilon$  increases, the annualized optimal Sharpe ratio per asset also monotonously increases. This is in sharp contrast to the annualized Sharpe ratio per asset for conventional trading without lead-lag correction,  $\sqrt{255} \delta_n/\sqrt{n}$ , shown on **Fig. 4(d)**. As expected, inter-asset correlations reduce diversification and thus diminish  $\delta_n/\sqrt{n}$  if lead-lag terms are ignored. Comparison of **Fig. 4(c)** and (d) suggests that inter-asset correlations can significantly increase the Sharpe ratio by inclusion of the lead-lag terms. In contrast to conventional views, these correlations, if correctly accounted for, are not deteriorative but beneficial. **Fig. 4(e)** shows the Sharpe gain (i.e., the ratio between the optimal and conventional Sharpe ratios,  $\delta_{\text{opt}}/\delta_n$ ) due to accounting for lead-lag corrections. This effect is particularly important for large  $\rho_\varepsilon$  and large  $n$ . It is also worth noting the difference with the earlier case  $\rho_\varepsilon = \rho_\xi$ , for which the annualized optimal Sharpe ratio per asset was independent of both  $n$  and correlations. In other words, equal inter-asset correlations between noises and stochastic trends do not provide opportunities for increasing the Sharpe ratio with correlations. In turn, correlations only between noises allow to TF strategies to better estimate and then eliminate their effects, enhancing contributions from stochastic trends.

Interestingly, the curves on **Fig. 4(c)** for different  $n$  do not coincide, as one might expect from the uncorrelated case. The larger the number of assets  $n$ , the faster increase of  $\sqrt{255} \delta_{\text{opt}}/\sqrt{n}$ . In other words, the Sharpe ratio of the optimal portfolio grows slightly faster than  $\sqrt{n}$ . At the same time, these curves progressively approach to the limiting curve as  $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} \frac{\delta_{\text{opt}}^2}{n} = \frac{q^2(1-p^2)}{Q^2(1-\rho_\varepsilon)^2 + 2Q(1-\rho_\varepsilon) + R}. \quad (43)$$

This explicit function (shown by black solid line) accurately reproduces the annualized optimal Sharpe ratio per asset for moderate  $n = 50$ . A rapid approach to the limit is illustrated on **Fig. 4(f)** which shows the annualized optimal Sharpe ratio per asset,  $\sqrt{255} \delta_{\text{opt}}/\sqrt{n}$ , as a function of  $n$  (shown by lines). This quantity rapidly saturates to a constant level, in contrast to  $\sqrt{255} \delta_n/\sqrt{n}$  for conventional trading which progressively diminishes with  $n$  (shown by symbols).

#### 4.4. Uncorrelated noises ( $\rho_\varepsilon = 0$ )

The other limiting case  $\rho_\varepsilon = 0$  is even more intriguing. **Fig. 5** shows the optimal lead-lag correction  $x_{\text{opt}}$ , the annualized optimal Sharpe ratio per asset,  $\sqrt{255} \delta_{\text{opt}}/\sqrt{n}$ , the annualized conventional Sharpe ratio per asset  $\sqrt{255} \delta_n/\sqrt{n}$ , and the Sharpe gain  $\delta_{\text{opt}}/\delta_n$ , as functions of  $\rho_\xi$ .

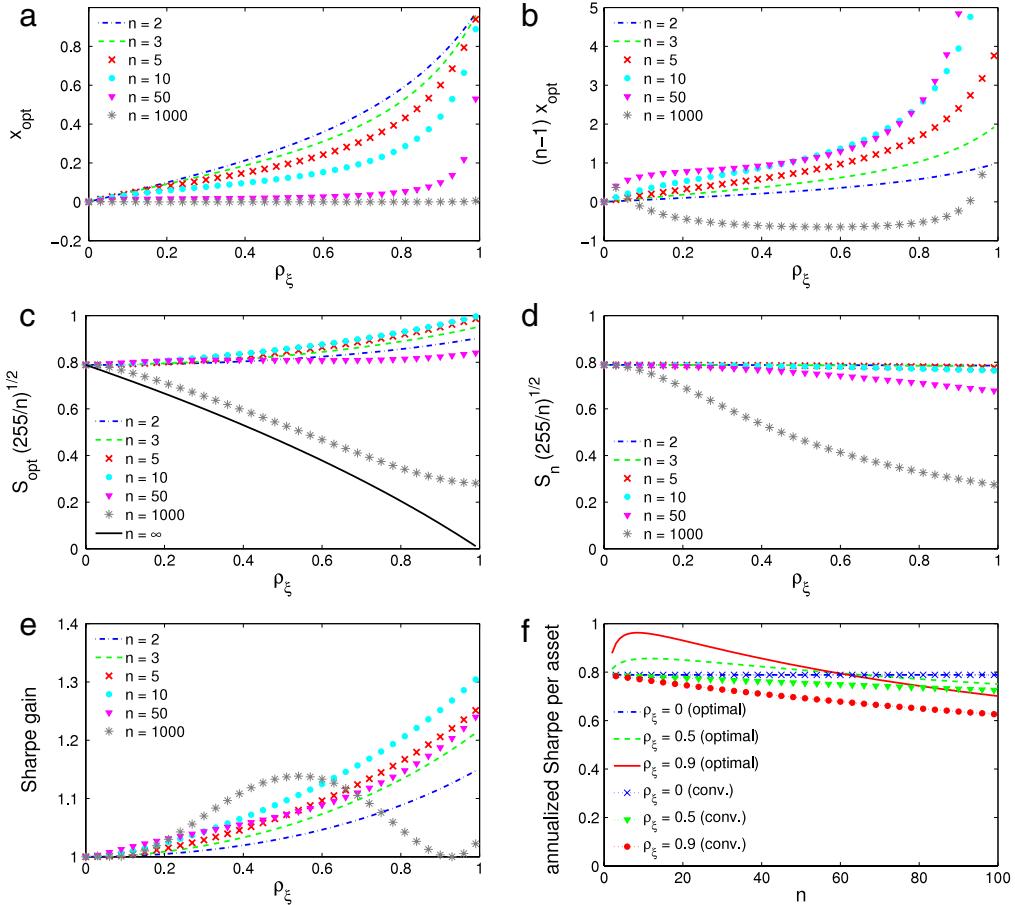
The optimal lead-lag correction  $x_{\text{opt}}$  is positive (**Fig. 5(a)**), as for the two-asset case. This observation may sound counter-intuitive because one might expect that a negative lead-lag term is needed to correct for positively correlated stochastic trends. Moreover, once  $x_{\text{opt}}$  is multiplied by  $n - 1$  (**Fig. 5(b)**), one can clearly see that the lead-lag term is indeed negative for  $n = 1000$  (but still remaining positive for small and moderate  $n$ ). Most surprisingly, the lead-lag correction for  $n = 1000$  is positive for small and large  $\rho_\xi$ , while negative for intermediate values. In order to clarify this situation, we consider the asymptotic behavior of the optimal lead-lag correction for large  $n$ :

$$(n-1)x_{\text{opt}} \simeq -1 + \frac{Q^2 + 2Q(1-\rho_\xi) + R(1-\rho_\xi)^2}{\rho_\xi(1-\rho_\xi)R(n-1)} + O\left(\frac{1}{(n-1)^2}\right). \quad (44)$$

As intuitively expected, the total lead-lag correction approaches to  $-1$ , as for the earlier case  $\rho_\xi = 0$  from Section 4.3. Although the first-order correction term vanishes in the limit  $n \rightarrow \infty$ , the coefficient in front of this term is large because  $R \ll Q$ . As a consequence, one needs to consider thousands of assets in order to approach the limit  $-1$ , in sharp contrast to the earlier case from Section 4.3. We conclude that, for the present case  $\rho_\varepsilon = 0$ , the asymptotic formulas in the limit  $n \rightarrow \infty$  are not useful, and may even be misleading when applied to moderate number of assets. In particular, we found positive lead-lag correction for the two-asset case in Section 3.2.

Another surprising feature appears in the non-monotonic behavior of the annualized optimal Sharpe ratio per asset,  $\sqrt{255} \delta_{\text{opt}}/\sqrt{n}$  (**Fig. 5(c)**). One can see that this quantity monotonously grows with  $\rho_\xi$  for small number of assets (up to 10), exhibits non-monotonic behavior for moderate  $n = 50$ , and decreases for very large number of assets ( $n = 1000$ ). Moreover, the curve for  $n = 10$  lies above the curve for  $n = 50$ , in sharp contrast to **Fig. 4(c)**. This is also seen on **Fig. 4(f)** which shows the same quantity as a function of  $n$ . Note, however, that the total annualized Sharpe ratio,  $\sqrt{255} \delta_{\text{opt}}$ , is still higher for larger  $n$ , due to the factor  $\sqrt{n}$ . In other words, if adding more assets with correlated noises ( $\rho_\varepsilon > 0$ ) increased the Sharpe ratio even more than by factor  $\sqrt{n}$ , adding more assets with correlated stochastic trends ( $\rho_\xi > 0$ ) may increase the Sharpe ratio by less than the factor  $\sqrt{n}$ . This is particularly clear from the comparison of the Sharpe gain on **Figs. 4(e)** and **5(e)**: for correlated stochastic trends ( $\rho_\xi > 0$ ), the Sharpe gain is modest (up to 20%) even at very high correlation coefficient  $\rho_\xi$ ; in contrast, the Sharpe gain was extremely large (a factor 10 or higher) for correlated noises ( $\rho_\varepsilon > 0$ ). The decrease of the annualized optimal Sharpe ratio per asset for  $n = 1000$  can be understood by considering the limit  $n \rightarrow \infty$ , for which

$$\lim_{n \rightarrow \infty} \frac{\delta_{\text{opt}}^2}{n} = q^2(1-p^2) \frac{(1-\rho_\xi)^2}{Q^2 + 2Q(1-\rho_\xi) + R(1-\rho_\xi)^2}. \quad (45)$$



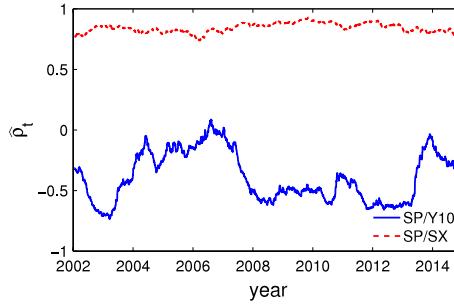
**Fig. 5.** The optimal lead-lag correction  $x_{\text{opt}}$  (a),  $(n-1)x_{\text{opt}}$  (b), the annualized optimal Sharpe ratio per asset,  $\sqrt{255} S_{\text{opt}}/\sqrt{n}$  (c), the annualized conventional Sharpe ratio per asset,  $\sqrt{255} S_n/\sqrt{n}$  (d), and the Sharpe gain  $S_{\text{opt}}/S_n$  (e), as functions of  $\rho_{\xi}$ , for  $n$  indistinguishable assets, with  $\rho_{\varepsilon} = 0$ ,  $\beta_0^j = 0.1$ ,  $\lambda = \eta = 0.01$ . The last plot (f) shows the annualized optimal Sharpe ratio per asset  $\sqrt{255} S_{\text{opt}}/\sqrt{n}$  (lines) and the annualized conventional Sharpe ratio per asset  $\sqrt{255} S_n/\sqrt{n}$  (symbols) as functions of  $n$ .

This limiting function (shown by black solid line on Fig. 5(c)) monotonously decreases with  $\rho_{\xi}$ . As for  $x_{\text{opt}}$ , the approach to the limit is very slow so that the limiting curve does not capture the behavior for moderately large  $n$ . We conclude that the case of correlated stochastic trends exhibits some counter-intuitive features that require particular attention from fund managers.

## 5. Illustrations and limitations

### 5.1. Inference of model parameters

A practical implementation of the present approach to optimal allocation of TF strategies requires advanced inference tools for estimating model parameters. It is well known that a robust estimation of the covariance matrix of many assets from their time series is a difficult problem [42–47]. The present approach requires estimating two covariance matrices describing inter-asset correlations: the matrix  $\mathbf{C}_{\varepsilon}$  for instantaneous fluctuations (noises)  $\varepsilon_t^j$ , and the matrix  $\mathbf{C}_{\xi}$  for trend components  $\xi_t^j$ . One also needs to characterize auto-correlations of each asset, i.e., the rate  $\lambda^j$  and the strength  $\beta^j$  of the stochastic trend component. Inferring all these parameters from a finite sample of price time series is a truly challenging problem for multiple assets. In practice, statistical estimations are often combined with trader's own market information, intuition, or belief. Note also that the hypothesis of time-independent inter-asset correlations is oversimplified and often fail for long time series that one may want to use for more reliable estimations (see Fig. 6). While the qualitative conclusions of the present theoretical work are expected to hold in real markets, the quantitative rules for optimal allocation of TF strategies require further investigations. Futures spread (e.g., between stocks and bonds) and specific sectors on futures markets (e.g., energy) in which assets are expected to be highly corrected, are promising financial instruments for studying optimal portfolios of TF strategies.



**Fig. 6.** EMA estimator  $\hat{\rho}_t$  of the correlation coefficient between the S&P 500 (SP) and the Euro Stoxx 50 (SX) futures (dashed line, top) and between the S&P 500 (SP) and the 10-year T-note futures (10Y) (solid line, bottom), over the timespan of 13 years (from January 2002 to December 2014).

## 5.2. Futures markets illustration

For illustrative purposes, we focus on the two-asset case described in Section 3. We consider two pairs of futures markets: an intra-sector pair with positive correlations (generally high, from 0.5 up to 1), and an inter-sector pair with time-varying correlations (generally between  $-0.5$  and  $+0.5$ ). As representative examples, we choose (i) the pair SP/SX of the E-mini S&P 500 (SP) and Euro Stoxx 50 Index (SX) futures, and (ii) the pair SP/10Y of the S&P 500 and the US 10-year T-note futures (10Y), respectively. The data set consists of 3314 normalized daily returns<sup>4</sup> for each of these three markets, from January 2002 to December 2014 (recorded at the same time to avoid synchronization issues).

Fig. 6 shows an EMA estimator  $\hat{\rho}_t$  of the time-dependent inter-asset correlation coefficient, computed with  $\lambda = 0.008$  (i.e., a timescale of a year). As mentioned earlier, the correlation coefficient remains positive and nearly constant for the intra-sector pair SP/SX while it exhibits several structural changes for the inter-sector pair SP/10Y, remaining negative almost all the time. While the estimate  $\hat{\rho}_t$  can be used as a proxy of the noise correlation coefficient  $\rho_\varepsilon$ , estimating the trend correlation  $\rho_\xi$  remains challenging. Although such cross-correlation measures are very sensitive to the time window (e.g.,  $\lambda$ ), they still provide a strong evidence of complex correlation structure between markets.

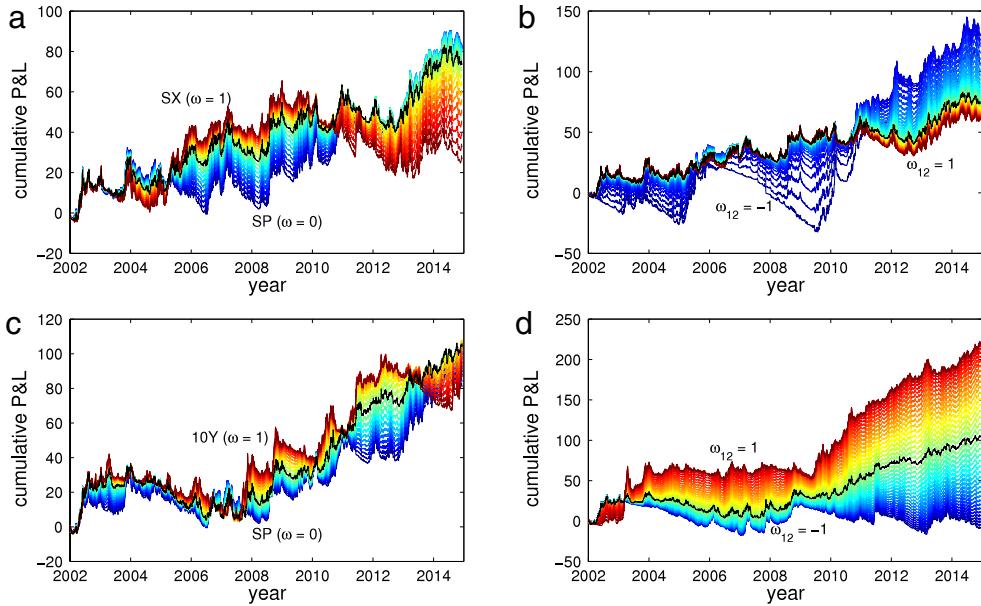
Without appropriate inference tools for fitting the model parameters from the datasets, we do not attempt to construct the optimal TF portfolio from the theoretical formulas of Section 3. We restrict ourselves to illustrating the effect of the lead-lag correction onto the profitability of two-asset portfolios.

We first consider a conventional allocation scheme when two TF strategies are treated as “virtual assets”  $X$  and  $Y$  with relative weights,  $\omega X + (1 - \omega)Y$ ,  $\omega$  ranging from 0 to 1. Such linear combinations of individual TF strategies are used in practice to build TF portfolios. Fig. 7(a), (c) show the cumulative P&L (i.e., “backtest”) of these classical portfolios for two pairs of markets: the intra-sector pair SP/SX and the inter-sector pair SP/10Y, respectively. In other words,  $X$  is the TF strategy applied to SP, while  $Y$  is the TF strategy applied to either SX (first case) or 10Y (second case). In all cases, the TF strategy relies on the EMA of earlier normalized returns with  $\lambda = 0.01$ , as defined in Eqs. (7), (15). The extreme values  $\omega = 0$  and  $\omega = 1$  correspond to the cumulative P&Ls of the underlying assets. Remarkably, even though the S&P 500 and Euro Stoxx 50 futures are highly positively correlated all the time (see Fig. 6), their cumulative P&Ls behave differently. This observation suggests that estimating the correlation of instantaneous price fluctuations (which is captured by the estimator  $\hat{\rho}_t$ ) may be insufficient for trend followers. The concept of correlated trends can provide useful complementary information.

Since the classical allocation does not account for inter-asset trend correlations, we expect that the best cumulative P&L of the classical TF portfolio remains sub-optimal. To illustrate this point, we consider a generalized TF portfolio introduced in Section 3.2 with equal diagonal weights  $\omega_{11} = \omega_{22} = 1$ . This choice can be justified under the assumption of equal auto-correlation excess  $\beta$  for three markets (we recall that all returns were resized by their realized volatility). We emphasize that this choice is used only for illustrations while the optimal portfolio for the considered market pairs remains unknown. Varying the lead-lag correction  $\omega_{12}$  from  $-1$  to  $1$ , we construct a family of generalized TF portfolios whose cumulative P&Ls are shown on Fig. 7(b), (d). The comparison to Fig. 7(a), (c) reveals that the “best” generalized portfolio greatly outperforms the “best” classical portfolio, the gain being particularly important for the inter-sector pair SP/10Y.

How can these backtests be interpreted within our model? If the correlation coefficients  $\rho_\varepsilon$  and  $\rho_\xi$  were equal, the optimal lead-lag correction would be  $\omega_{12} = x_{\text{opt}} = -\rho_\varepsilon$ . For instance, strong negative correlations in the inter-sector pair SP/10Y would require positive lead-lag corrections. One can see on Fig. 7(d) that the generalized portfolio with the highest cumulative P&L (since 2003) corresponds to  $\omega_{12} \simeq 1$  or, equivalently, to the TF strategy applied to the combined asset SP + 10Y. The value  $\omega_{12} \simeq 1$  is higher than the value  $\omega_{12} \simeq 0.5$  that one might expect from  $\hat{\rho}_t \simeq -0.5$  for the pair SP/10Y (Fig. 6). This quantitative discrepancy can have multiple origins such as distinct values of  $\rho_\varepsilon$  and  $\rho_\xi$  (see Fig. 2(a)),

<sup>4</sup> As mentioned earlier, we call by “returns” additive logarithmic price variations resized by realized volatility. Futures contracts are rolled (according to volume criteria), while transaction costs are not taken into account (though it is straightforward to include them as discussed in Ref. [10]). The data set has been provided by John Locke Investments.



**Fig. 7.** (a, c) Cumulative P&L (“backtest”) of a conventional portfolio, in which the cumulative P&Ls of two TF strategies  $X$  and  $Y$  enter through a weighted linear combination  $\omega X + (1 - \omega)Y$ , with  $X$  corresponding to SP, while  $Y$  corresponding to either SX (a) or 10Y (c). 21 color lines represent 21 portfolios with  $\omega$  varying from 0 (blue) to 1 (red) at step 0.05, the black line in the “middle” indicating the portfolio with equal weights ( $\omega = 0.5$ ). (b), (d) Cumulative P&L (“backtest”) of a generalized portfolio composed of two TF strategies with  $\lambda = 0.01$ , in which the diagonal weights are fixed to be  $\omega_{11} = \omega_{22} = 1$ , while the lead-lag correction  $\omega_{12}$  ranges from  $-1$  to  $1$ . The underlying two assets are SP/SX (b) and SP/10Y (d). 41 color lines represent 41 portfolios with  $\omega_{12}$  varying from  $-1$  (blue) to 1 (red) at step 0.05, the black line in the “middle” indicating the portfolio without lead-lag correction ( $\omega_{12} = 0$ ). The P&Ls have been risk adjusted over the period. Note different ranges of the vertical axes of these four plots. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

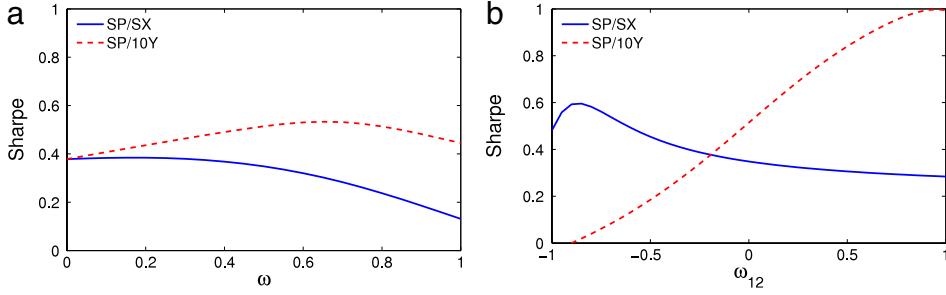
time dependence of correlation coefficients, different auto-correlation excess of two assets, or other real market features that were ignored in our simplified model. While backtests absolute performance is very sensitive to parameters choice, the relative performance between conventional and generalized portfolios highlights the impact of the lead-lag term.

Fig. 8 shows the empirical annualized Sharpe ratio (i.e., the realized risk-adjusted returns) of both conventional and generalized portfolios presented in Fig. 7. While the best conventional portfolios yield the Sharpe ratios 0.38 and 0.53 for the intra-sector (SP/SX) and inter-sector (SP/10Y) pairs, the best generalized portfolios (with equal diagonal weights) perform much better, ensuring the Sharpe ratios 0.60 and 1.00, respectively. In other words, accounting for the lead-lag correction improves the Sharpe ratio by factor 1.58 for the SP/SX pair and by factor 1.89 for the SP/10Y pair. We conclude that inter-asset correlations present an important source of potential improvements for trend followers. Further analysis is necessary to systematically apply these concepts to various assets and to exploit inter-asset correlations in real markets.

## 6. Conclusion

The principle of diversification in portfolio management [14,15] calls for investing in as many uncorrelated assets as possible in order to reduce a portfolio risk. The same principle is often applied to trend following portfolios. In a typical setting, the exposure of each TF strategy is determined by earlier returns of the traded asset, independently of other assets. Then, the weight of each strategy is adjusted to account for asset covariance structure. In this setting, investing in correlated assets is more risky and thus yields smaller risk-adjusted returns (or Sharpe ratio). However, properly modeling the source of correlations can be beneficial as a mean to estimate apparent trends more reliably, to adjust the TF portfolio more efficiently, and thus to enhance the Sharpe ratio. The paper aimed at analyzing this concept in a quantitative way. For this purpose, we introduced a simple Gaussian model, in which volatility-normalized returns of each asset had two contributions: short-range noises mimicking instantaneous *uninformative* price fluctuations (e.g., daily), and stochastic trends modeling TF profitability on a longer time scale (e.g., monthly). The Gaussian assumption allowed us to derive analytically the mean and variance of the portfolio profit-and-loss and thus to formulate explicitly the problem of Sharpe ratio maximization. This problem was solved analytically for the two-asset case and for a sector model of  $n$  similar assets, while an exact numerical solution is possible in general.

As mentioned earlier, we argue that each strategy should incorporate information from other strategies. We considered each asset investment as a linear combination of strategy signals of all assets. The weight  $\omega_{j,k}$  represents a correction term to the  $j$ th asset investment due to its correlation to the  $k$ th asset. This cross correcting, or lead-lag, term is indeed proportional



**Fig. 8.** Annualized Sharpe ratio for conventional (a) and generalized (b) portfolios from Fig. 7 as functions of  $\omega$  and  $\omega_{12}$ , respectively.

to the  $k$ th strategy signal, itself being a linear combination of  $k$ th asset past returns. We compared optimal portfolios formed by these augmented strategies with optimal portfolios of individual strategies. Instead of determining the weights of  $n$  TF strategies (one for each asset), our generalized portfolio allocation problem operates with  $n^2$  weights  $\omega_{j,k}$ .

One may look at it as a static allocation problem on  $n^2$  virtual strategies, which include the  $n$  conventional individual trading strategies, plus  $n(n-1)$  lead-lag strategies, contributing to an asset position using past price variations of another one. We showed that when there is no correlation between assets, both approach are equivalent and the optimal weight matrix  $\omega_{j,k}$  is diagonal. However, optimal portfolio weights contain non-diagonal terms in presence of inter-asset correlations. As demonstrated in the two-asset case, the static allocation scheme is sub-optimal whereas the introduction of the lead-lag terms with  $\omega_{j,k}$  can make inter-asset correlations highly beneficial.

We also investigated the respective roles of noise–noise ( $\rho_\varepsilon$ ) and trend–trend ( $\rho_\xi$ ) inter-asset correlations. The separate accounting for these two mechanisms is a new feature of our model. When  $\rho_\varepsilon = \rho_\xi$  (i.e., the same structure of inter-asset correlations for noises and trends), the optimal Sharpe ratio for  $n$  indistinguishable assets is equal to  $\sqrt{n}$  times the Sharpe ratio of one asset, as expected due to diversification. The inclusion of the lead-lag corrections  $\omega_{j,k}$  allows one to reach this optimal Sharpe ratio even for strongly correlated assets but one cannot overperform here the benchmark case of uncorrelated assets. In turn, when  $\rho_\varepsilon \neq \rho_\xi$ , the optimal Sharpe ratio can be much larger than that for the benchmark case due to statistically more reliable estimations. Since economical and financial mechanisms behind short-range price fluctuations and longer trends may be quite different, one can speculate about hidden opportunities due to possible mismatches between  $\rho_\varepsilon$  and  $\rho_\xi$ . Perhaps, numerous algorithmic tricks and empirical hints in trend following strategies, as well as managers' intuition, aim to catch these opportunities in practice. Our study presents a first step towards better understanding of these mechanisms.

The effect of the lead-lag correction onto the profitability of TF portfolios was illustrated on two pairs of futures markets: the intra-sector pair of positively correlated S&P 500 and Euro Stoxx 50 futures, and the inter-sector pair of negatively correlated S&P 500 and 10-year T-note futures. Without building an optimal portfolio, a mere inclusion of the lead-lag term led to a significant increase of the annualized Sharpe ratio of the generalized TF portfolio. The optimization of TF strategies (e.g., the optimal choice of  $\lambda$ ) and the use of optimal allocation weights  $\omega_{j,k}$  can further improve the overall performance. Although a practical use of the optimization framework requires further work (notably about parameters inference), the developed theoretical model provides a first characterization of inter-asset trend correlations and shows a promising way for trend followers to enhance the profitability of their portfolios.

## Appendix A. Mean and variance of incremental P&L of a portfolio

From Eq. (8), the mean incremental profit-and-loss of a portfolio is simply

$$\langle \delta \mathcal{P}_t \rangle = \sum_{j,k=1}^n \omega_{j,k} \sum_{t'=1}^{t-1} \mathbf{S}_{t,t'}^k \langle r_t^j r_{t'}^k \rangle = \sum_{j,k=1}^n \omega_{j,k} \mathbf{C}_\xi^{j,k} (\mathbf{S}^k \mathbf{A}^k \mathbf{A}^{j,\dagger})_{t,t}, \quad (\text{A.1})$$

where the explicit structure (3) of the covariance  $\langle r_t^j r_{t'}^k \rangle$  was used.

Next, we compute the variance of  $\delta \mathcal{P}_t$  as

$$\text{var}\{\delta \mathcal{P}_t\} = \sum_{j_1,k_1=1}^n \omega_{j_1,k_1} \sum_{t'_1=1}^{t-1} (\mathbf{S}^{k_1})_{t,t'_1} \sum_{j_2,k_2=1}^n \omega_{j_2,k_2} \sum_{t'_2=1}^{t-1} (\mathbf{S}^{k_2})_{t,t'_2} \left[ \langle r_t^{j_1} r_{t'}^{j_2} \rangle \langle r_{t'_1}^{k_1} r_{t'_2}^{k_2} \rangle + \langle r_t^{j_1} r_{t'}^{k_2} \rangle \langle r_{t'_2}^{j_2} r_{t'_1}^{k_1} \rangle \right], \quad (\text{A.2})$$

where we used Wick's theorem for Gaussian returns  $r_t^k$ . Substituting again the covariances from Eq. (3), one gets

$$\begin{aligned} \text{var}\{\delta\mathcal{P}_t\} = & \sum_{j_1,k_1=1}^n \omega_{j_1,k_1} \sum_{t'_1=1}^{t-1} (\mathbf{S}^{k_1})_{t,t'_1} \sum_{j_2,k_2=1}^n \omega_{j_2,k_2} \sum_{t'_2=1}^{t-1} (\mathbf{S}^{k_2})_{t,t'_2} \\ & \times \left[ (\mathbf{C}_{\varepsilon}^{j_1,j_2} + \mathbf{C}_{\xi}^{j_1,j_2} (\mathbf{A}^{j_1} \mathbf{A}^{j_2,\dagger})_{t,t}) (\delta_{t'_1,t'_2} \mathbf{C}_{\varepsilon}^{k_1,k_2} + \mathbf{C}_{\xi}^{k_1,k_2} (\mathbf{A}^{k_1} \mathbf{A}^{k_2,\dagger})_{t'_1,t'_2}) \right. \\ & \left. + (\mathbf{C}_{\xi}^{j_1,k_2} (\mathbf{A}^{j_1} \mathbf{A}^{k_2,\dagger})_{t,t'_2}) (\mathbf{C}_{\xi}^{j_2,k_1} (\mathbf{A}^{j_2} \mathbf{A}^{k_1,\dagger})_{t,t'_1}) \right], \end{aligned}$$

from which

$$\begin{aligned} \text{var}\{\delta\mathcal{P}_t\} = & \sum_{j_1,k_1,j_2,k_2=1}^n \omega_{j_1,k_1} \omega_{j_2,k_2} \left[ (\mathbf{C}_{\varepsilon}^{j_1,j_2} + \mathbf{C}_{\xi}^{j_1,j_2} (\mathbf{A}^{j_1} \mathbf{A}^{j_2,\dagger})_{t,t}) \right. \\ & \times \left. (\mathbf{C}_{\varepsilon}^{k_1,k_2} (\mathbf{S}^{k_1} \mathbf{S}^{k_2,\dagger})_{t,t} + \mathbf{C}_{\xi}^{k_1,k_2} (\mathbf{S}^{k_1} \mathbf{A}^{k_1} \mathbf{A}^{k_2,\dagger} \mathbf{S}^{k_2,\dagger})_{t,t}) + \mathbf{C}_{\xi}^{j_1,k_2} \mathbf{C}_{\xi}^{k_1,j_2} (\mathbf{S}^{k_1} \mathbf{A}^{k_1,\dagger} \mathbf{A}^{j_2})_{t,t} (\mathbf{S}^{k_2} \mathbf{A}^{j_1,\dagger} \mathbf{A}^{k_2})_{t,t} \right]. \quad (\text{A.3}) \end{aligned}$$

This is the variance of the incremental P&L in the general case. Note that Eqs. (A.1) and (A.3) can be re-written in the form (9), (10).

In what follows, we introduce a simplifying assumption that all the assets have the same rates:  $\lambda^j = \lambda$  (i.e.,  $q^j = q$ ). Similarly, we assume that all the trend following strategies have the same rate:  $\eta^k = \eta$  (i.e.,  $p^k = p$ ). In this case, the elements  $M_t^{j,k}$  and  $V_t^{j_1,k_1;j_2,k_2}$  from Eq. (10) become

$$\begin{aligned} M_t^{j,k} &= \gamma^k (\mathbf{E}_p \mathbf{E}_q \mathbf{E}_q^\dagger)_{t,t} \mathbf{C}_{\xi,\beta}^{j,k}, \\ V_t^{j_1,k_1;j_2,k_2} &= \gamma^{k_1} \gamma^{k_2} \left[ \mathbf{C}_{\varepsilon}^{j_1,j_2} \mathbf{C}_{\varepsilon}^{k_1,k_2} (\mathbf{E}_p \mathbf{E}_p^\dagger)_{t,t} + \mathbf{C}_{\varepsilon}^{k_1,k_2} \mathbf{C}_{\xi,\beta}^{j_1,j_2} (\mathbf{E}_p \mathbf{E}_p^\dagger)_{t,t} (\mathbf{E}_q \mathbf{E}_q^\dagger)_{t,t} \right. \\ &\quad \left. + \mathbf{C}_{\varepsilon}^{j_1,j_2} \mathbf{C}_{\xi,\beta}^{k_1,k_2} (\mathbf{E}_p \mathbf{E}_q \mathbf{E}_q^\dagger \mathbf{E}_p^\dagger)_{t,t} + \mathbf{C}_{\xi,\beta}^{j_1,j_2} \mathbf{C}_{\xi,\beta}^{k_1,k_2} (\mathbf{E}_q \mathbf{E}_q^\dagger)_{t,t} (\mathbf{E}_p \mathbf{E}_q \mathbf{E}_q^\dagger \mathbf{E}_p^\dagger)_{t,t} + \mathbf{C}_{\xi,\beta}^{j_1,k_2} \mathbf{C}_{\xi,\beta}^{k_1,j_2} [(\mathbf{E}_p \mathbf{E}_q \mathbf{E}_q^\dagger)_{t,t}]^2 \right], \quad (\text{A.4}) \end{aligned}$$

where  $[\mathbf{E}_q]_{t,t'} = q^{t-t'-1}$  for  $t > t'$ , and 0 otherwise, and  $\mathbf{C}_{\xi,\beta}^{j,k} \equiv \beta^j \beta^k \mathbf{C}_{\xi}^{j,k}$ . Supplementary Materials to Ref. [10] provide the explicit formulas for various products of matrices  $\mathbf{E}_p$  and  $\mathbf{E}_q$ . In the stationary limit  $t \rightarrow \infty$ , one gets

$$\begin{aligned} \lim_{t \rightarrow \infty} (\mathbf{E}_p \mathbf{E}_p^\dagger)_{t,t} &= \frac{1}{1-p^2}, \\ \lim_{t \rightarrow \infty} (\mathbf{E}_p \mathbf{E}_q \mathbf{E}_q^\dagger)_{t,t} &= \frac{q}{(1-q^2)(1-pq)}, \\ \lim_{t \rightarrow \infty} (\mathbf{E}_p \mathbf{E}_q \mathbf{E}_q^\dagger \mathbf{E}_p^\dagger)_{t,t} &= \frac{1+pq}{(1-pq)(1-q^2)(1-p^2)}, \quad (\text{A.5}) \end{aligned}$$

from which we obtain

$$\begin{aligned} M_\infty^{j,k} &= \frac{q\sqrt{1-p^2}}{1-pq} \mathbf{C}_{\xi,\beta_0}^{j,k}, \\ V_\infty^{j_1,k_1;j_2,k_2} &= \mathbf{C}_{\varepsilon}^{j_1,j_2} \mathbf{C}_{\varepsilon}^{k_1,k_2} + \mathbf{C}_{\varepsilon}^{k_1,k_2} \mathbf{C}_{\xi,\beta_0}^{j_1,j_2} + \mathbf{C}_{\varepsilon}^{j_1,j_2} \mathbf{C}_{\xi,\beta_0}^{k_1,k_2} \frac{1+pq}{1-pq} + \mathbf{C}_{\xi,\beta_0}^{j_1,j_2} \mathbf{C}_{\xi,\beta_0}^{k_1,k_2} \frac{1+pq}{1-pq} + \mathbf{C}_{\xi,\beta_0}^{j_1,k_2} \mathbf{C}_{\xi,\beta_0}^{k_1,j_2} \frac{q^2(1-p^2)}{(1-pq)^2}, \quad (\text{A.6}) \end{aligned}$$

where  $\mathbf{C}_{\xi,\beta_0}^{j,k} \equiv \mathbf{C}_{\xi,\beta}^{j,k}/(1-q^2)$ , and we set  $\gamma^k = \sqrt{1-(1-\eta^k)^2} = \sqrt{1-p^2}$ . This normalization was proposed in Ref. [10] to set the unit variance of the stationary incremental P&L of a single asset without auto-correlations (i.e., when  $\beta = 0$ ). For the multivariate case, this normalization yields the expected form  $V_\infty^{j_1,k_1;j_2,k_2} = \mathbf{C}_{\varepsilon}^{j_1,j_2} \mathbf{C}_{\varepsilon}^{k_1,k_2}$  when all  $\beta^j = 0$ . For symmetric weights, the above expression for the covariance matrix can be further simplified to get Eq. (16).

## Appendix B. Two indistinguishable assets

When  $\beta_0^1 = \beta_0^2 = \beta_0$  (i.e.,  $\kappa = 1$ ) and  $\sigma^1 = \sigma^2 = 1$  (i.e.,  $\nu = 1$ ), two assets have the same structure of auto-correlations that makes them indistinguishable from each other. In this case, Eqs. (18), (20) are reduced to

$$\delta^2 = q^2(1-p^2) \frac{(\omega_{11} + 2\rho_\xi \omega_{12} + \omega_{22})^2}{Q^2 \Omega_1 + 2Q \Omega_2 + R \Omega_3}, \quad (\text{B.1})$$

with

$$\begin{aligned}\mathcal{Q}_1 &= \omega_{11}^2 + 2\omega_{12}^2 + \omega_{22}^2 + 2\rho_\varepsilon^2(\omega_{12}^2 + \omega_{11}\omega_{22}) + 4\rho_\varepsilon\omega_{12}(\omega_{11} + \omega_{22}), \\ \mathcal{Q}_2 &= \omega_{11}^2 + 2\omega_{12}^2 + \omega_{22}^2 + 2\rho_\varepsilon\rho_\xi(\omega_{12}^2 + \omega_{11}\omega_{22}) + 2(\rho_\varepsilon + \rho_\xi)\omega_{12}(\omega_{11} + \omega_{22}), \\ \mathcal{Q}_3 &= \omega_{11}^2 + 2\omega_{12}^2 + \omega_{22}^2 + 2\rho_\xi^2(\omega_{12}^2 + \omega_{11}\omega_{22}) + 4\rho_\xi\omega_{12}(\omega_{11} + \omega_{22}).\end{aligned}\quad (\text{B.2})$$

In this case, three quadratic equations determining the weights ratios,  $z = \omega_{11}/\omega_{22}$  and  $x = \omega_{12}/\omega_{22}$ , are

$$\begin{aligned}2Ax^2 + 2Bxz + 2Cx - Dz + D &= 0, \\ Dz^2 + 2Ax^2 + 2Cxz - Dz + 2Bx &= 0, \\ Bz^2 + Axz + 2Cz + Ax + B &= 0,\end{aligned}\quad (\text{B.3})$$

where

$$\begin{aligned}A &= Q^2(1 - 2\rho_\varepsilon\rho_\xi + \rho_\varepsilon^2) + 2Q(1 - \rho_\xi^2) + R(1 - \rho_\xi^2), \\ B &= Q(Q + 1)(\rho_\varepsilon - \rho_\xi), \\ C &= Q^2\rho_\varepsilon(1 - \rho_\varepsilon\rho_\xi) + Q(\rho_\varepsilon + \rho_\xi - 2\rho_\varepsilon\rho_\xi^2) + R\rho_\xi(1 - \rho_\xi^2), \\ D &= Q^2(1 - \rho_\varepsilon^2) + 2Q(1 - \rho_\varepsilon\rho_\xi) + R(1 - \rho_\xi^2).\end{aligned}\quad (\text{B.4})$$

The difference between the first two relations in Eqs. (B.3) yields the equation  $(z - 1)[2(B - C)x - D(1 + z)] = 0$ , from which one determines both  $z$  and  $x$ . One can show that the quadratic equation corresponding to the choice  $z = 2(B - C)x/D - 1$  does not have real solutions. As a consequence, we get the following solution of the minimization problem:  $z_{\text{opt}} = 1$  (i.e.,  $\omega_{11} = \omega_{22}$ ), while  $x_{\text{top}}$  is given by Eq. (25).

## Appendix C. Derivation for a sector model

We consider the case of  $n$  indistinguishable assets (with  $\beta^j = \beta$  and  $\sigma^j = \sigma = 1$ ). In the optimal portfolio of TF strategies, all assets are expected to have the same weight,  $\omega_{jj} = \omega_{11}$ , as well as all lead-lag corrections are the same:  $\omega_{jk} = \omega_{12}$  for all  $j \neq k$ . Substituting these weights into Eqs. (9), (16), one gets

$$\begin{aligned}\langle \delta\mathcal{P}_\infty \rangle &= \frac{q\sqrt{1-p^2}}{1-pq} \sum_{j,k}^n \mathbf{C}_{\xi,\beta_0}^{j,k} \omega_{j,k} = \frac{q\sqrt{1-p^2}\beta_0^2}{1-pq} [\omega_{11}n + \omega_{12}n(n-1)\rho_\xi], \\ \text{var}\{\delta\mathcal{P}_\infty\} &= \sum_{j_1,k_1,j_2,k_2}^n V^{j_1,k_1;j_2,k_2} \omega_{j_1,k_1} \omega_{j_2,k_2} = \omega_{11}^2 \tilde{V}_1 + 2\omega_{11}\omega_{12} \tilde{V}_2 + \omega_{12}^2 \tilde{V}_3,\end{aligned}\quad (\text{C.1})$$

where

$$\begin{aligned}\tilde{V}_1 &\equiv \sum_{j,k} V^{j,j;k,k} = \sum_j \left[ \mathbf{C}_\varepsilon^{j,k} \mathbf{C}_\varepsilon^{j,k} + \frac{2\beta_0^2}{1-pq} \mathbf{C}_\varepsilon^{j,k} \mathbf{C}_\xi^{j,k} + \frac{\beta_0^4 R}{(1-pq)^2} \mathbf{C}_\xi^{j,k} \mathbf{C}_\xi^{j,k} \right] \\ &= n \left[ 1 + \frac{2\beta_0^2}{1-pq} + \frac{\beta_0^4 R}{(1-pq)^2} \right] + n(n-1) \left[ \rho_\varepsilon^2 + \frac{2\beta_0^2}{1-pq} \rho_\varepsilon \rho_\xi + \frac{\beta_0^4 R}{(1-pq)^2} \rho_\xi^2 \right], \\ \tilde{V}_2 &\equiv \sum_{j,j_2 \neq k_2} V^{j,j;j_2,k_2} = \sum_{j,j_2 \neq k_2} \left[ \mathbf{C}_\varepsilon^{j,j_2} \mathbf{C}_\varepsilon^{j,k_2} + \frac{2\beta_0^2}{1-pq} \mathbf{C}_\varepsilon^{j,j_2} \mathbf{C}_\xi^{j,k_2} + \frac{\beta_0^4 R}{(1-pq)^2} \mathbf{C}_\xi^{j,j_2} \mathbf{C}_\xi^{j,k_2} \right] \\ &= n(n-1) \left[ 2\rho_\varepsilon + (n-2)\rho_\varepsilon^2 + \frac{2\beta_0^2}{1-pq} (\rho_\varepsilon + \rho_\xi + (n-2)\rho_\varepsilon\rho_\xi) + \frac{\beta_0^4 R}{(1-pq)^2} (2\rho_\xi + (n-2)\rho_\xi^2) \right], \\ \tilde{V}_3 &\equiv \sum_{j_1 \neq k_1; j_2 \neq k_2} V^{j_1,k_1;j_2,k_2} = \tilde{V}_0 - \tilde{V}_1 - 2\tilde{V}_2, \\ \tilde{V}_0 &\equiv \sum_{j_1,k_1,j_2,k_2} V^{j_1,k_1;j_2,k_2} = \sum_{j_1,k_1,j_2,k_2} \left[ \mathbf{C}_\varepsilon^{j_1,j_2} \mathbf{C}_\varepsilon^{k_1,k_2} + \frac{2\beta_0^2}{1-pq} \mathbf{C}_\varepsilon^{j_1,j_2} \mathbf{C}_\xi^{k_1,k_2} + \frac{\beta_0^4 R}{(1-pq)^2} \mathbf{C}_\xi^{j_1,j_2} \mathbf{C}_\xi^{k_1,k_2} \right] \\ &= n^2 \left[ (1 + (n-1)\rho_\varepsilon)^2 + \frac{2\beta_0^2}{1-pq} (1 + (n-1)\rho_\varepsilon)(1 + (n-1)\rho_\xi) + \frac{\beta_0^4 R}{(1-pq)^2} (1 + (n-1)\rho_\xi)^2 \right],\end{aligned}\quad (\text{C.2})$$

and we used the identity

$$\mathbf{C}_\varepsilon \mathbf{C}_\xi = \begin{pmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ b & b & b & \cdots & a \end{pmatrix}, \quad \begin{cases} a = 1 + (n-1)\rho_\varepsilon\rho_\xi, \\ b = \rho_\varepsilon + \rho_\xi + (n-2)\rho_\varepsilon\rho_\xi. \end{cases} \quad (\text{C.3})$$

We get then

$$\delta^2 = \frac{q^2(1-p^2)n(\omega_{11} + \omega_{12}(n-1)\rho_\xi)^2}{\omega_{11}^2 V_1 + 2\omega_{11}\omega_{12}V_2 + \omega_{12}^2 V_3}, \quad (\text{C.4})$$

where the new coefficients  $V_j \equiv \tilde{V}_j \frac{(1-pq)^2}{n\rho_0^4}$  are obtained from Eqs. (C.2) and given in Eqs. (35). Setting to zero the derivative of  $\delta^2$  with respect to  $\omega_{11}$ , one gets the optimal weights ratio  $x_{\text{opt}}$  in Eq. (33), from which the corresponding squared optimal Sharpe ratio in Eq. (34) is deduced.

Since the explicit formulas for  $x_{\text{opt}}$  and  $\delta_{\text{opt}}$  are too cumbersome, it is instructive to consider their asymptotic behavior as  $n$  goes to infinity. Since each asset has  $n-1$  lead-lag terms, it is convenient to introduce the small parameter as  $1/(n-1)$  (instead of  $1/n$ ). In particular, one gets

$$(n-1)x_{\text{opt}} \simeq -1 + \frac{\rho_\xi [Q^2(1-\rho_\varepsilon)^2 + 2Q(1-\rho_\varepsilon)(1-\rho_\xi) + R(1-\rho_\xi)^2]}{(1-\rho_\xi)[Q^2\rho_\varepsilon^2 + 2Q\rho_\varepsilon\rho_\xi + R\rho_\xi^2](n-1)} + \dots \quad (\text{C.5})$$

Setting  $\rho_\varepsilon = 0$ , one retrieves Eq. (44). In turn, when  $\rho_\xi = 0$ , the first-order term in Eq. (C.5) vanishes, and one needs to compute the next-order term to get Eq. (42).

## 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] W. Fung, D.A. Hsieh, Rev. Financ. Stud. 14 (2001) 313.
- [4] A. Clare, J. Seaton, P.N. Smith, S. Thomas, The trend is our friend: Risk parity, momentum and trend following in global asset allocation, Cass Business School Working Paper, 2012; SSRN-id2126478.
- [5] C.S. Asness, T.J. Moskowitz, L.H. Pedersen, J. Finance 68 (2013) 929.
- [6] M. Potters, J.-P. Bouchaud, Wilmott Magazine, January 2006; online: ArXiv:physics-0508104, 2005.
- [7] L.K.C. Chan, N. Jegadeesh, J. Lakonishok, J. Finance 51 (1996) 1681.
- [8] N. Jegadeesh, S. Titman, J. Finance 56 (2001) 699.
- [9] T.J. Moskowitz, Y.H. Ooi, L.H. Pedersen, J. Financ. Econ. 104 (2012) 228.
- [10] D.S. Grebenkov, J. Serror, Physica A 394 (2014) 288.
- [11] H. Allen, M.P. Taylor, Econom. J. 100 (1990) 49.
- [12] C. Wilcox, E. Crittenden, Does trend-following work on stocks? Tech. Anal. 14 (2005).
- [13] A.C. Szakmary, Q. Shen, S.C. Sharma, J. Bank. Financ. 34 (2010) 409.
- [14] W.F. Sharpe, G.J. Alexander, J.V. Bailey, Investments, sixth ed., Prentice Hall, Englewood Cliffs, NJ, 1999.
- [15] A. Ilmanen, J. Kizer, J. Portf. Manag. 38 (2012) 15.
- [16] N. Gärleanu, L.H. Pedersen, J. Finance 68 (2013) 2309.
- [17] R.C. Merton, J. Econom. Theory 3 (1971) 373.
- [18] E.J. Elton, M.J. Gruber, J. Bank. Financ. 21 (1997) 1743.
- [19] B. Pfaff, Financial Risk Modelling and Portfolio Optimization with R, John Wiley & Sons, 2013.
- [20] D. Li, W.-L. Ng, Math. Finance 10 (1998) 387.
- [21] H.M. Markowitz, Portfolio selection, J. Finance 7 (1952) 77.
- [22] J.Y. Campbell, L.M. Viceira, Strategic Asset Allocation: Portfolio Choice for Long-Term Investors, Oxford University Press, 2002.
- [23] J.-P. Bouchaud, R. Cont, Eur. Phys. J. B 6 (1998) 543.
- [24] H.R. Stoll, R.E. Whaley, J. Finan. Quant. Anal. 25 (1990) 441.
- [25] K. Chan, Rev. Financ. Stud. 5 (1992) 123.
- [26] R. Martin, D. Zou, Momentum trading: 'skews me, Risk Mag. (2012).
- [27] S. Thomas, A. Clare, P.N. Smith, J. Seaton, The trend is our friend: risk parity, momentum and trend following in global asset allocation. Working Paper, 2012; online: SSRN-id2126478.
- [28] J.-P. Bouchaud, M. Potters, Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management, Cambridge University Press, 2003.
- [29] R. Mantegna, H.E. Stanley, An Introduction to Econophysics, Cambridge University Press, Cambridge, 1999.
- [30] R. Mantegna, H.E. Stanley, Nature 376 (1995) 46.
- [31] J.-P. Bouchaud, M. Potters, Physica A 299 (2001) 60.
- [32] D. Sornette, Phys. Rep. 378 (2003) 1.
- [33] J.-P. Bouchaud, Y. Gefen, M. Potters, M. Wyart, Quant. Finance 4 (2004) 176.
- [34] J.-P. Bouchaud, A. Matacz, M. Potters, Phys. Rev. Lett. 87 (2001) 1.
- [35] S. Valérye, D.S. Grebenkov, S. Aboura, Q. Liu, Quant. Finance 13 (2013) 1697.
- [36] T.G. Andersen, T. Bollerslev, F.X. Diebold, P. Labys, Multinat. Financ. J. 4 (2000) 159.
- [37] J.C. Hull, Options, Futures and Other Derivatives, seventh ed., Pearson Prentice Hall, Upper Saddle River, NJ, 2009.
- [38] P.R. Winters, Manag. Sci. 6 (1960) 324.
- [39] R.G. Brown, Smoothing Forecasting and Prediction of Discrete Time Series, Prentice-Hall, Englewood Cliffs, NJ, 1963.
- [40] G. Burghardt, B. Walls, Managed Futures for Institutional Investors: Analysis and Portfolio Construction, Wiley & Sons, NJ, Hoboken, 2010.
- [41] A.W. Lo, Financ. Anal. J. 58 (2002) 36.
- [42] P. Embrechts, A. McNeil, D. Straumann, Correlation and dependence in risk management: properties and pitfalls, in: Risk Management: Value at Risk and Beyond, Cambridge University Press, 2002, pp. 176–223.
- [43] D.J. Fenn, M.A. Porter, S. Williams, M. McDonald, N.F. Johnson, N.S. Jones, Phys. Rev. E 84 (2011) 026109.
- [44] L. Laloux, P. Cizeau, J.-P. Bouchaud, M. Potters, Phys. Rev. Lett. 83 (1998) 1467.
- [45] Z. Burda, A. Gorlich, A. Jarosz, J. Jurkiewicz, Physica A 343 (2004) 295.
- [46] J.-P. Bouchaud, L. Laloux, M.A. Miceli, M. Potters, Eur. Phys. J. B 55 (2007) 201.
- [47] R. Allez, J.-P. Bouchaud, Phys. Rev. E 86 (2012) 046202.