Risk Minimization through Portfolio Replication

Stefano Ciliberti\textsuperscript{1,2} and Marc Mézard\textsuperscript{1}

\textsuperscript{1}CNRS; Univ. Paris Sud, UMR8626, LPTMS, ORSAY CEDEX, F-91405 (France)
\textsuperscript{2}Science \& Finance, Capital Fund Management, 6 bd Haussmann, 75009, Paris (France)

Abstract

We use a replica approach to deal with portfolio optimization problems. A given risk measure is minimized using empirical estimates of asset values correlations. We study the phase transition which happens when the time series is too short with respect to the size of the portfolio. We also study the noise sensitivity of portfolio allocation when this transition is approached. We consider explicitly the cases where the absolute deviation and the conditional value-at-risk are chosen as a risk measure. We show how the replica method can study a wide range of risk measures, and deal with various types of time series correlations, including realistic ones with volatility clustering.
I. INTRODUCTION

The portfolio optimization problem dates back to the pioneering work of Markowitz and is one of the main issues of risk management. Given that the input data of any risk measure ultimately come from empirical observations of the market, the problem is directly related to the presence of noise in financial time series. In a more abstract (model-based) approach, one uses Monte Carlo simulations to get “in-sample” evaluations of the objective risk function. In both cases the issue is how to take advantage of the time series of the returns on the assets in order to properly estimate the risk associated with our portfolio. This eventually results in the choice of the risk measure, and a long debate in the recent years has drawn the attention on two important and distinct clues: the mathematical property of coherence, and the noise sensitivity of the optimal portfolio. The rational behind the first of these issues lies in the need of a formal (axiomatic) translation of the basic common principles of risk management, like the fact that portfolio diversification should always lead to risk reduction. Moreover, requiring a risk measure to be coherent implies the existence of a unique optimal portfolio and a well-defined variational principle, of obvious relevance in practical cases. The second issue is also a very delicate one. In a realistic experimental set-up, the number \( N \) of assets included in a portfolio can be of order \( 10^2 \) to \( 10^3 \), while the length of a trustable time series hardly goes beyond a few years, i.e. \( T \sim 10^3 \). A good estimate of any extensive observable would require the condition \( N/T \ll 1 \) to hold, but this is rarely the case. Instead, the ratio of assets to data points, \( N/T \), will be considered as a finite number.

In this note we address analytically the risk minimization problem by studying the dependence of the optimal portfolio on the ratio \( N/T \) and on other potential external parameters. We first assume that the real distribution of returns is multinormal in order to keep the problem tractable from the analytical point of view. Generalizations to more realistic returns distributions are also presented. Our approach consists in writing down the empirical estimate of the risk measure and then reformulating the problem from the point of view of the statistical physics. We work out the analytical solution by means of the replica method and thus get some insights on the optimal portfolios. The analytical solution confirms previous results on the existence of a phase transition. The ratio \( N/T \) plays the role of a control parameter. When it increases, there exists a sharply defined threshold
value where the estimation error of the optimal portfolio diverges. A first account of our method, limited to the Expected Shortfall risk measure, has appeared in ref. 5. Here we give a more general presentation, studying other risk measures and more realistic distributions of returns.

The paper is organized as follows. In section II we introduce the notations we will use throughout the paper and we formulate the problem in its general mathematical form. In section III we consider the case of the absolute deviation (AD) 6. We present the replica calculation of the optimal portfolio and compute explicitly a noise sensitivity measure introduced in ref. 10. In section IV we deal with portfolio optimization under Expected Shortfall 2, 7, which was shown to have a non-trivial phase diagram 4 and then studied analytically 5. The striking point is that, for some values of the external parameters of the problem, the minimization problem is not well defined and thus cannot admit a finite solution. We investigate here the same feature while considering realistic distribution of returns, so as to take into account volatility clustering. The replica approach then turns into a semi-analytic and extremely versatile technique. We discuss this point and then summarize our results in section V.

II. THE GENERAL SETTING

We denote our portfolio by \( \mathbf{w} = \{w_1, \ldots, w_N\} \), where \( w_i \) is the position on asset \( i \). We do not impose any restriction to short selling: \( w_i \) is a real number. The global constraint induced by the total budget reads \( \sum_i w_i = N \), where, due to a later mathematical convenience, we have chosen a slightly different normalization with respect to the previous literature. Calling \( x_i \) the return of the asset \( i \) and assuming the existence of a well-defined probability density function (pdf) \( p(x_1, \ldots, x_N) \), one is interested in computing the pdf of the loss \( \ell \) associated to a given portfolio, i.e.

\[
p_{\mathbf{w}}(\ell) = \int \prod_i dx_i \ p(x_1, \ldots, x_N) \ \delta \left( \ell + \sum_{i=1}^{N} w_i x_i \right).
\]  

(1)

The complete knowledge of this pdf would lead to the precise, though still probabilistic, evaluation of the loss, thus allowing for a straightforward optimization over the space of legal portfolios. This is actually a pretty difficult task and one usually restricts to some characteristic of this pdf (e.g. its first moments, its tail behavior), so as to capture the
consequences of extremely bad events in the global loss. The actual $p(x_1, \ldots x_N)$ is not known in general, and integrals like the one in (1) are usually estimated by time series, coming from market observations or synthetically produced by numerical simulations. Whatever the chosen risk measure then, one typically faces cost functions (to be optimized over all possible portfolios) like

$$
\text{risk}(w; N, T, \lambda) = \frac{1}{T} \sum_{\tau=1}^{T} F_\lambda \left[ \sum_{i=1}^{N} w_i x_i^{(\tau)} \right],
$$

(2)

where $\{x_i^{(1)}, x_i^{(2)} , \ldots x_i^{(T)}\}$ is the whole time series of the return $i$ and where we denoted by $\lambda$ other possible external parameters of the risk measure. The best known example of risk measure is of course the variance, as first suggested by Markowitz. In that case the risk function is obtained by taking $F_\lambda(z) = z^2$ in (2). The evaluation of the variance implies an empirical evaluation of the covariance matrix $\sigma_{ij}$ of the underlying stochastic process, and the extremely noisy character of any estimation of $\sigma_{ij}$ has been underlined a few years ago [8, 9]. However, recent studies [10, 11] have shown that the effect of the noise on the actual portfolio risk is not as dramatic as one might have expected. More in detail, a direct measure of this effect was introduced and explicitly computed in the simplest case of $\sigma_{ij} = \delta_{ij}$. In the next section, we compute the same quantity as far as the absolute deviation of the loss is concerned.

In the statistical physics approach, one studies the limit $N, T \to \infty$, while $N/T \equiv 1/t$ is finite. One introduces the partition function at inverse temperature $\gamma$:

$$
Z_\gamma^{(N)}[t, \lambda; \{x_i^{(\tau)}\}] = \int \prod_{i=1}^{N} dw_i e^{-\gamma \text{risk}[w; N, Nt, \lambda]} \delta \left( \sum_{i=1}^{N} w_i - N \right),
$$

(3)

from which any observable will be computed. For instance, the optimal cost (i.e. the minimum of the risk function in (2)) is computed from

$$
e(t, \lambda) = \lim_{N \to \infty} \frac{1}{N} \min_w \text{risk}[w; N, Nt, \lambda] = \lim_{N \to \infty} \frac{1}{N} \lim_{\gamma \to \infty} -\frac{1}{\gamma} \log Z_\gamma^{(N)}[t, \lambda; \{x_i^{(\tau)}\}].
$$

(4)

It turns out that this expression depends on the actual sample (the time series $\{x_i^{(\tau)}\}$) used to estimate the risk measure. We are mainly interested in the average over all possible time series of this quantity, which we assume to be narrowly distributed around its mean value. Taking the average of eq. (4) means that we have to average the logarithm of the partition function according to the pdf $p(\{x_i^{(\tau)}\})$. The so-called replica method allows to simplify this task as follows. We compute $E[Z^n]$ for integer $n$ and assume we can analytically continue
this result to real $n$: then $E[\log Z] = \lim_{n \to 0}(E[Z^n] - 1)/n$. This is the strategy that we are going to use in the next sections and that will allow to compute the optimal portfolio.

III. REPLICA ANALYSIS: ABSOLUTE DEVIATION

The absolute deviation measure $AD[w; N, T]$ is obtained by choosing $F_\lambda(z) = |z|$ in (2). No other external parameters $\lambda$ are present here. We assume a factorized distribution

$$p\left[\{x_i^{(\tau)}\}\right] \sim \prod_{i, \tau} \exp \left( -\frac{N(x_i^{(\tau)})^2}{2\sigma^2} \right),$$

where the volatilities $\{\sigma_\tau\}$ are distributed according to a pdf which we do not specify for the moment. Following the replica method, we introduce $n$ identical replicas of our portfolio and compute the average of $Z^n$:

$$E[Z^n_\gamma(t)] \sim \int \prod_{a, b = 1}^n dQ^{ab} dQ^{ab} e^{N \sum_{a, b = 1}^n (Q^{ab} - 1)\hat{Q}^{ab} - \frac{N}{2} \text{Tr} \log \hat{Q} - \frac{1}{2} \text{Tr} \log Q + \sum_\tau \log A_\gamma(\{Q^{ab}\}; \sigma_\tau)},$$

where we have introduced the overlap matrix

$$Q^{ab} = \frac{1}{N} \sum_{i = 1}^N w_i^a w_i^b, \quad a, b = 1, \ldots n,$$

as well as its conjugate $\hat{Q}^{ab}$, the Lagrange multipliers introduced to enforce (7). In the limit $N, T \to \infty$, $N/T = 1/t$ finite, the integral in (6) can be solved by a saddle point method. Due to the symmetry of the integrand by permutation of replica indices, there exists a replica-symmetric saddle point [3]: $Q^{aa} = q_1$, $Q^{ab} = q_0$ for $a \neq b$, and the same for $\hat{Q}^{ab}$. We expect the saddle point to be correct in view of the fact that the problem is linear. Under this hypothesis, which will be only justified a posteriori by a direct comparison to numerical data, the replicated partition function in (6) gets simplified into

$$E[Z^n_\gamma(t)] \sim \int dq_0 \int d\Delta q \exp \left[ Nn S_\gamma(q_0, \Delta q)(1 + O(n)) \right],$$

$$S_\gamma(q_0, \Delta q) = \frac{(1 - t)q_0 - 1}{2\Delta q} + \frac{1 - t}{2} \log \Delta q + \frac{1}{T} \sum_\tau \frac{1}{n} \log A_\gamma(q_0, \Delta q; \sigma_\tau),$$

$$A_\gamma(q_0, \Delta q; \sigma_\tau) = \int \frac{ds}{\sqrt{2\pi q_0}} e^{-s^2/2q_0} \left[ 1 + n \int du e^{-\frac{u^2}{2\Delta^2 q_0^2} + \frac{1}{2} \Delta q_0^2 - \gamma |u|} + O(n^2) \right].$$
where $\Delta q = q_1 - q_0$ and $n$ is the number of replicas (which will eventually go to zero). We now assume that in the low temperature limit the overlap fluctuations are of order $1/\gamma$ and introduce $\Delta = \gamma \Delta q$. One can show that if $\Delta$ stays finite at low temperatures

$$
\lim_{n \to 0} \lim_{\gamma \to \infty} \frac{1}{n\gamma} \log A_\gamma(q_0, \Delta/\gamma; \sigma_\tau) = \Delta^2 \sigma_\tau^3 \int_1^\infty ds \, e^{-s^2/(2\Delta^2)} (1 - s)^2 .
$$

(9)

For the sake of clarity, we focus on the simple case $\sigma_\tau = 1\forall \tau$. In the $\gamma \to \infty$ limit, the saddle point equations for (8) are

$$
\frac{1}{t} = \operatorname{erf} \left( \frac{1}{\sqrt{2q_0'}} \right),
$$

(10)

$$
\Delta = \left( 2t \left[ \frac{1 - 1/t}{2} q_0' + \frac{q_0'}{2\pi} e^{-1/2q_0'} - \frac{(1 + q_0')}{2} \left( 1 - \operatorname{erf} \left( \frac{1}{\sqrt{2q_0'}} \right) \right) \right] \right)^{-1/2},
$$

(11)

where $q_0 = q_0' \Delta^2$. The minimum cost function, i.e. the average of eq. (4), is found to be $e(t) = 1/\Delta$. Notice that (10) only admits a solution for $t \geq 1$. There is no solution to the minimization problem if the ratio of assets to data points, $N/T$, is smaller than 1. On the other hand, once this condition is fulfilled, the equation (11) gives a finite $\Delta$ at any $t > 1$. The asymptotic behaviour of $e(t)$ can be worked out analytically: we introduce $\delta \equiv 1 - 1/t$ and consider the limit $\delta \ll 1$. This leads to

$$
e(t) \simeq \sqrt{\frac{\delta}{-2 \log \delta}} \left( 1 - \frac{\log \left( \frac{-\frac{1}{2} \log \delta}{4 \log \delta} \right)}{4 \log \delta} \right).
$$

(12)

The full solution and a comparison with numerics are shown in Fig. 1 (left).

We now address the issue of noise sensitivity, for which a measure was introduced in [10]. The idea is the following: Assume you know the true pdf of the loss (11) and you get some optimal $w^{(0)}$ by minimizing the absolute deviation of $\ell$. We want to compare the optimal risk associated to $w^{(0)}$ with the one obtained by optimizing (2), i.e. the empirical estimation of the same risk measure. A fair comparison is then $q_K - 1$, with

$$
q_K^2(N, T) = \frac{\text{AD}\left[ w^*; N, T \right]}{\text{AD}\left[ w^{(0)}; N \right]},
$$

(13)

where the $w_i^*$ refer to the portfolio obtained by minimizing (2). This is the quantity which we have computed by the replica approach. In our calculation we have assumed to deal with a factorized Gaussian distribution of returns (extensions to more realistic cases will be presented in the next section) and it is straightforward to prove that in this case $q_K = \cdots$
FIG. 1: **Left:** The analytic solution $e(t)$ is compared with the results of numerical simulations, where the constrained optimization is computed directly via linear programming methods [12].

**Right:** Numerical results for $\sqrt{\sum_{i=1}^{N}(w_i^*)^2}$ compared to the analytic behaviour $\sqrt{q_0\Delta}$. The curve denoted by $q_K$ (var) represents the behaviour of $q_K$ in the variance minimization problem.

$\sqrt{\sum_{i=1}^{N}(w_i^*)^2}$. This corresponds in our language to $\sqrt{q_0} = \sqrt{q_0\Delta}$, which diverges like $(1 - 1/t)^{-1/2}$ as $1/t \to 1^-$. Corrections to this leading behavior (which is instead the full shape of $q_K$ in the variance minimization problem) are needed in order to reproduce the data (right panel of Fig. 1). The comparison with the Markowitz optimal portfolio (variance minimization) indicates that the AD measure is actually less stable to perturbations: A geometric interpretation of this result can be found in ref. [4]. Beside this fact, the interesting result is then the existence of a well defined threshold value $t = 1$ at which the estimation error becomes infinite. This is due to the divergence of the variance of the optimal portfolio in the regime $t < 1$, where any minimization attempt is thus totally meaningless.

### IV. EXPECTED SHORTFALL

#### A. The minimization problem

For a fixed value of $\beta < 1$ ($\beta \gtrsim 0.9$ in the interesting cases) the expected-shortfall (ES) of a portfolio $w$ is obtained by choosing $F(z) \propto z\theta(z - \text{VaR})$ in [2], where VaR stands here for the Value-at-Risk [13]. In practice, it is computed from the minimization of a properly
chosen objective function \([14]\): 

\[
\text{ES}[w; N, T, \beta] = \min_v \left\{ v + \frac{1}{(1 - \beta)T} \sum_{\tau=1}^{T} \left[ -v - \sum_{i=1}^{N} w_i x_i^{(\tau)} \right]^+ \right\},
\]

where \([a]^+ \equiv (a + |a|)/2\). Optimizing the ES risk measure over all the possible portfolios satisfying the budget constraint is equivalent to the following linear programming problem:

- **Cost function:** \(E = (1 - \beta)T v + \sum_{\tau=1}^{T} u_\tau\);
- **Variables:** \(Y \equiv \{w_1, \ldots w_N, u_1, \ldots u_T, v\}\);
- **Constraints:** \(u_\tau \geq 0\), \(u_\tau + v + \sum_{i=1}^{N} x_i w_i \geq 0\), \(\sum_{i=1}^{N} w_i = N\).

In a previous work \([5]\) we solved the problem in the case where the historical series of returns is drawn from the oversimplified probability distribution \([15]\), with \(\sigma_\tau = 1\ \forall \tau\). Here we do a first step towards dealing with more realistic data and assume that the series of returns can be obtained by a sequence of normal distributions whose variances depend on time:

\[
p[\{\sigma_\tau\}] \sim \prod_{\tau, \tau'} \exp \left( -\sigma_\tau \sigma_{\tau'} G_{\tau, \tau'}^{-1} \right) \prod_{\tau} q(\sigma_\tau),
\]

for some long range correlator \(G_{\tau, \tau'}\) which takes into account volatility correlations, and \(q(\sigma_\tau)\) equal e.g. to a lognormal distribution.

### B. The replica solution

A straightforward generalization of the replica calculation presented in ref. \([8]\) (and sketched in the previous section for a similar problem) allows to compute the average optimal cost for a given volatility sequence \(\{\sigma_1, \ldots \sigma_T\}\), in the limit when \(N, T \to \infty\) and \(N/T = 1/t\) stays finite. This is given by

\[
e(t, \beta) = \min_{v, q_0, \Delta} \left[ \frac{1}{2\Delta} + \Delta \bar{\varepsilon}(t, \beta; v, q_0|\{\sigma_\tau\}) \right],
\]

\[
\bar{\varepsilon}(t, \beta; v, q_0|\{\sigma_\tau\}) \equiv t(1 - \beta)v - \frac{q_0}{2} + \frac{t}{2\sqrt{\pi T}} \sum_{\tau=1}^{T} \int_{-\infty}^{+\infty} ds \ e^{-s^2} g(v/\sigma_\tau + s\sqrt{2q_0}; \sigma_\tau),
\]

where \(\Delta \equiv \lim_{\gamma \to \infty} \gamma \Delta q\) and the function \(g(x; \sigma)\) is equal to \(x^2\) if \(-\sigma \leq x < 0\), to \(-2\sigma x - \sigma^2\) is \(x < -\sigma\), and 0 otherwise. The minimization over \(v, q_0\) implies that

\[
\partial \bar{\varepsilon}/\partial v = \partial \bar{\varepsilon}/\partial q_0 = 0.
\]
As discussed in [5], the problem admits a finite solution if (17) is minimized by a finite value of $\Delta$. The feasible region is then defined by the condition $\tilde{\varepsilon}(t, \beta; v, q|\{\sigma_i\}) \geq 0$, where $v$ and $q_0$ satisfy (18). This theoretical setup suggests the following semi-analytic protocol for determining the phase diagram of realistic portfolio optimization problems.

1. Fix a value of $\beta \in [0, 1]$, and take $N$ equal to the portfolio size you are interested in.

2. For $T = T_{\text{min}}$ to $T_{\text{max}}$, such that $N/T \in [0.1, 0.9]$, do the following:

   (a) Generate a sequence $\{\sigma_1, \sigma_2, \ldots, \sigma_T\}$ according to (15) and compute the $\tilde{\varepsilon}$ function in (17).

   (b) Minimize $\tilde{\varepsilon}$ with respect to $v$ and $q_0$ according to (18).

   (c) Repeat steps (a) and (b) for $n$ samples, and compute the mean value $\langle \tilde{\varepsilon} \rangle$.

3. Plot $\langle \tilde{\varepsilon} \rangle$ vs. $N/T$ and find the value $(N/T)^*$ where this function changes its sign.

By repeating this procedure for several values of $\beta$ we get the phase separation line $(N/T)^*$ vs. $\beta$.

**C. Results**

A simple way of generating realistic volatility series consists in looking at the return time series as a cascade process [15]. In a multifractal model recently introduced [16] the volatility covariance decreases logarithmically: this is achieved by letting $\sigma_\tau = \exp \xi_\tau$, where $\xi_\tau$ are Gaussian variables and

$$
\langle \xi_\tau \rangle = -\lambda^2 \log T_{\text{cut}}, \quad \langle \xi_\tau \xi_{\tau'} \rangle - \langle \xi_\tau^2 \rangle = \lambda^2 \log \frac{T_{\text{cut}}}{1 + |\tau - \tau'|},
$$

$\lambda$ quantifying volatility fluctuations (the so-called ‘vol of the vol’), and $T_{\text{cut}}$ being a large cutoff. A few samples generated according to this procedure are shown in Fig. 2.

The phase diagram obtained for different values of $\lambda^2$ is shown in Fig. 3. A comparison with the phase diagram computed in absence of volatility fluctuations shows that, while the precise shape of the separating curve depend on the fine details of the volatility pdf, the main message has not changed: There exists a regime, $N/T > (N/T)^*$, where the small number of data with respect to the portfolio size makes the optimization problem ill-defined.
In the “max-loss” limit $\beta \to 1$, where the single worst loss contributes to the risk measure, the threshold value $(N/T)^* = 0.5$ does not seem to depend on the volatility fluctuations. As $\beta$ gets smaller than 1, though, the presence of these fluctuations is such that the feasible region becomes smaller than the ideal multinormal case.

V. CONCLUSIONS

In this paper we have discussed the replica approach to portfolio optimization. The rather general formulation of the problem allows to deal with several risk measures. We have shown here the examples of absolute deviation, expected shortfall and max-loss (which is simply taken as the limit case of ES). In all cases we find that the optimization problem, when the risk measure is estimated by using time series, does not admit a feasible solution if the ratio of assets to data points is larger than a threshold value. As discussed in ref. 4, this is a common feature of various risk measures: the estimation error on the optimal portfolio, originating from in-sample evaluations, diverges as a critical value is approached. In the expected shortfall case, we have also discussed a semi-analytic approach which is suitable
FIG. 3: The phase diagram corresponding to different values of the parameter $\lambda^2$. The full line corresponds to the absence of fluctuations in the volatility distributions (i.e. $\sigma_\tau = 1 \forall \tau$).

for describing realistic time series. Our results suggest that, as far as volatility clustering is taken into account, the phase transition is still there, the only effect being the reduction of the feasible region. As a general remark, we have shown that the replica method may prove extremely useful in dealing with optimization problems in risk management.

Acknowledgments. We thank I. Kondor and J.-P. Bouchaud for interesting discussions. S. C. is supported by EC through the network MTR 2002-00319, STIPCO.

[1] H. Markowitz, *Portfolio selection: Efficient diversification of investments*, J. Wiley & Sons, New York (1959).
[2] P. Artzner, F. Delbaen, J. M. Eber, and D. Heath, Mathematical Finance 9, 203–228 (1999).
[3] M. Mézard, G. Parisi, and M. A. Virasoro, “Spin Glass theory and Beyond”, World Scientific Lecture Notes in Physics Vol. 9, Singapore (1987).
[4] I. Kondor, Lectures given at the Summer School on “Risk Measurement and Control”, Rome, June 9-17, 2005; I. Kondor, S. Pafka, and G. Nagy, Noise sensitivity of portfolio selection under various risk measures, submitted to Journal of Banking and Finance.
[5] S. Ciliberti, I. Kondor, and M. Mézard, *arXiv:physics/0606015*. 

11
[6] H. Konno, H. Yamazaki, Management Science 37, 519 (1991); H. Konno, T. Koshizuka, IIE Transactions 37, (10) 893 (2005).

[7] C. Acerbi, and D. Tasche, Journal of Banking and Finance 26, 1487–1503 (2002).

[8] L. Laloux, P. Cizeau, J.-P. Bouchaud, and M. Potters, Phys. Rev. Lett. 83, 1467 (1999).

[9] V. Plerou, P. Gopikrishnan, B. Rosenow, L. Nunes Amaral, and H. Stanley, Phys. Rev. Lett. 83, 1471 (1999).

[10] S. Pafka, and I. Kondor, Europ. Phys. J. B27, 277 (2002).

[11] S. Pafka, and I. Kondor, Physica A319, 487 (2003).

[12] W. H. Press, S. H. Teukolsky, W. T. Wetterling, and B. P. Flannery, “Numerical Recipes in C”, Cambridge University Press (Cambridge, UK, 1992).

[13] see http://www.gloriamundi.org for a complete source of references.

[14] R. Rockafellar, and S. Uryasev, The Journal of Risk 2, 21–41 (2000).

[15] B. B. Mandelbrot, J. Fluid Mech. 62, 331 (1974).

[16] J.-F. Muzy, J. Delour, and E. Bacry, Eur. Phys. J. B17, 537 (2000).