Portfolio instability and linear constraints

Nándor Gulyás[a,b], Imre Kondor[a,c]

[a] *Collegium Budapest - Institute for Advanced Study, Szentháromság u. 2., H-1014 Budapest, Hungary*

[b] *Department of Information Systems, Eötvös University, Pázmány P. sétány 1/c, H-1117 Budapest, Hungary*

[c] *Department of Physics of Complex Systems, Eötvös University, Pázmány P. sétány 1/a, H-1117 Budapest, Hungary*

Abstract

It is well known that portfolio optimization is unstable: the out-of-sample variance and the weights of the optimal portfolio show large sample to sample fluctuations. Moreover, it has been shown recently that the estimation error in the minimum risk portfolio diverges at a critical value of the ratio of the portfolio size N and the length T of the available time series. This divergence is a signal of an algorithmic phase transition, and is accompanied with a number of critical phenomena, scaling laws, universality, etc.

Divergent fluctuations can only appear in an infinite volume, however. In reality, portfolio selection is always performed under various constraints (on short selling, industrial sectors, geographical regions, etc.) that render the volume where the optimum is sought finite. These constraints, together with the budget constraint, define finite volume polyhedra in the space of portfolio weights. As a vestige of the infinite volume phase transition, the vector of portfolio weights will still exhibit large fluctuations within the allowed region, so the solution to the optimization problem will contain very little reliable information concerning the true solution that would only be attainable in the infinite T limit. In addition, as a result of these large fluctuations, the solution tends to stick to the faces of the constraint-polyhedron. As the ratio N/T increases, more and more components of the solution vector will be found on the boundaries of the allowed region. It is clear that under these conditions the solution is defined more by the constraints than by the cost function. In particular, when short selling is banned, a large number of the weights will become zero. This spontaneous reduction of diversification is often observed in portfolio optimization, but we are not aware of any link previously recognized between this phenomenon and the critical growth of estimation error.

This paper describes the results of numerical measurements of these effects under different risk measures (variance and maximal loss) and different sets of linear constraints as function of the ratio N/T. Results are given for the relative estimation error, for the average angle of the solution vector and the true solution, and for the percentage of the solutions sticking to the walls of the allowed region.

Key words: estimation noise, portfolio optimization, linear constraints
JEL C13, C15, C61, G11

1 Introduction

It is well known that portfolio optimization, as formulated in the seminal paper by Markowitz [1], suffers from the "curse of dimensions": for large values of the portfolio size N and limited length T of the available time series the estimation error will be huge. Finance theory and econometrics have been struggling with this problem for decades, and a large number of efficient methods have been proposed to reduce the effect of measurement noise, see e.g. [2],[3],[4],[5],[6].

Our purpose here will not be to introduce another filtering method, but to study the effect of noise under various linear constraints.

The problem of estimation error has been put into a new context recently with the demonstration of the fact that the estimation error actually diverges for a critical value of the ratio $r = N/T$ [7] and with the recognition of this divergence as a manifestation of an algorithmic phase transition [9]. This phase transition exhibits remarkable universality properties: the exponent of the quantity q_0, introduced in [7] as a measure of the estimation error, is largely independent of the details of the various models considered, such as the covariance structure of the market [8], the risk measure used [9], [10], [11], or the nature of the underlying stochastic process [12]. A related and equally surprising phenomenon observed in [9] is that for some risk measures (expected shortfall, maximal loss) even the existence of an optimum becomes a probabilistic issue: for a certain fraction of the samples, depending again on the critical parameter $r = N/T$, the optimization problem will not have a solution. In the case of expected shortfall and for asymptotically large portfolios this leads to a sharp phase boundary on the plane spanned by r and the cutoff beyond which the conditional expectation value of loss is calculated under that risk measure. This way the phase transition concept is able to unify very different aspects of the problem into a single coherent picture.

Divergent fluctuations can only occur in an infinite volume, however, and, indeed, the problem considered in [7], [9] corresponds to seeking the minimal risk portfolio with absolutely no restriction on the portfolio weights other than that they sum to unity (the budget constraint). In such an, admittedly rather unrealistic, situation the portfolio weights can be of either sign and can take

* Corresponding author: Nándor Gulyás
e-mail: ngulyas@colbud.hu

arbitrarily large absolute values. A ban on short selling (i.e. stipulating that all the weights be non-negative), or any other constraint that restricts the domain of the optimization problem to a finite volume (e.g. limits on various assets, industrial sectors, or geographic regions, but even limits of a purely technical nature, built into the portfolio management software in order to avoid infinite cycles, that the analyst may not even be aware of) will evidently prevent infinitely large fluctuations, thereby eliminating the divergence of q_0. Numerical experience tells us, however, that the portfolio weights will continue to display strong sample to sample fluctuations even in a finite volume, and the purpose of this paper is to study and quantitatively characterize these residual fluctuations.

The ultimate cause of the instability described above is lack of sufficient information. In order to have a reliable estimate of the risk measure, we would need a situation where the dimension of the problem N is much smaller than the length T of the time series. Institutional portfolios are large, containing hundreds of items, whereas the choice of the length and frequency of the time series is dictated by considerations of stationarity and transaction costs, respectively. In practice T is never longer than $T = 1000$ (corresponding to four years worth of daily data), and often much shorter. (E.g. the EWMA method advocated by Riskmetrics [13] starts to cut off around a three-months-long time horizon. The use of high-frequency, intraday data would be utterly meaningless in a portfolio optimization context: one does not restructure portfolios every 5 seconds.) Therefore, the inequality $N/T << 1$ almost never holds in practice, and we have to live with the consequences of this information deficit.

In order to avoid the problems inherent in real-life time series (fat tails, non-stationarity), we follow the same strategy here as in [9], that is we generate very long simulated time series, and cut out segments of length T from them, as if making observations on a market. This way we will have a total control over the underlying stochastic process, and can have as many samples as needed for the study of the statistics of sample to sample fluctuations.

The linear constraints under which we consider the fluctuations of the portfolio weights define a polyhedron in the space of weights. We are going to study the behaviour of the weights optimized under two different risk measures (variance and maximal loss, respectively) as the ratio $r = N/T$ crosses its critical value (1 for variance, and 1/2 for maximal loss [9]) from below, and goes deep into the $N/T >> 1$ region. The main effect we find is that with N/T increasing a larger and larger fraction of the weights gets stuck to the boundaries of the allowed domain, that is to the faces of the polyhedron defined by the constraints. In the case of a ban on short selling this means that a large number of weights becomes zero, as it were, the portfolio spontaneously reduces its diversification. This effect has been described earlier, see [14], where a brief discussion can be found in Chapters 1 and 2, without a systematic study of the evolution of the fraction of zero weights as function of N/T.

On the basis of our phase transition picture it is very easy to understand how this spontaneous reduction of portfolio size comes about. Were it not for the constraints, with N/T increasing a larger and larger fraction of the weights would run away to very large values, some to infinity. It is these runaway solutions that bump into and get stuck to the walls represented by the constraints.

Although the presence of the constraints prevents the global estimation error from blowing up for any finite r, the quantity q_0 is monotonically increasing with r and, in the case of excluded short selling, approaches its upper bound $\sqrt{N}$. As will be shown below, this corresponds to a single weight becoming unity, with all the rest vanishing.

As we repeat the experiment several times, we notice that the set of weights sticking to the walls changes each time, as if the solution were jumping around on the walls. This is just another manifestation of the remanent fluctuations in a finite volume. We are not aware of any previous description of this effect.

In order to quantitatively grasp the effects described above, we introduce two new measures, the fraction of weights sticking to the boundaries, and an angle variable, which measures the average deviation of the solution vector from its true direction. Although both these new measures are intimately related to q_0, they provide a more visually appealing characterization of the solution in the $r > 1$ region. The rest of the paper gives details about the behaviour of these parameters.

2 Mathematical setup

In this section we set up our mathematical framework and define the two new measures of fluctuations just mentioned.

(1) The original Markovitz problem.
Here the risk measure is the standard deviation of the portfolio, and the optimisation problem is a quadratic programming task [1]. For simplicity, we look for the minimal risk portfolio, without any constraint on the expected return.

$$min \quad \sigma^2$$
$$\sigma^2 = \sum_{i,j} w_i \cdot \sigma_{ij} \cdot w_j$$
$$\sum_i w_i = 1$$

$$a_i \leq w_i \leq b_i \qquad \forall i = 1..N, \tag{1}$$

where σ_{ij} is the empirical covariance matrix and a_i and b_i represent the linear constraints. Together with the budget constraint, the latter define a polyhedron.

(2) The case of Maximal Loss as a risk measure.

$$\begin{aligned} &min \quad u \\ u = -max_t &\sum_i w_i \cdot x_{it} \\ &\sum_i w_i = 1 \\ a_i \leq w_i \leq b_i \qquad &\forall i = 1..N, \end{aligned} \tag{2}$$

where x_{it} are the portfolio returns. Note that Maximal Loss is a special case of Expected Shortfall (see [15],[16]), corresponding to the limit when for each asset only the single worst return is considered. The optimization of a linear combination of the worst outcomes over the weights is a linearly programmable minimax problem.

In the simplest case the constraint parameters are $a_i = 0$ and $b_i = 1$ $\forall$ $i = 1 \ldots N$, and the returns are uncorrelated Gaussian random variables with zero mean and unit standard deviation. This simple setup provides a convenient framework in which to display the finite volume effects. A slightly more complicated case, with a simple correlation pattern, will be considered in a later section. (Sec. 3.2)

2.1 Measures

The quantity q_0, introduced in [7], [9] is a measure of the relative estimation error due to the finiteness of the samples.

$$q_0 = \left\langle \sqrt{\frac{\sum_{ij} \sigma_{ij} \cdot w_i^{meas.} \cdot w_j^{meas.}}{\sum_{ij} \sigma_{ij} \cdot w_i^{true} \cdot w_j^{true}}} \right\rangle_{x_{it}} \tag{3}$$

where $\left\langle \cdot \right\rangle_{x_{it}}$ means the average over the random variables x_{it}.

When no other constraint is imposed except the budget constraint, q_0 diverges at a critical value of the ratio $r = N/T$ that depends on the risk measure [9]. In the present study the optimum is sought in a finite volume that is bordered by a polyhedron; therefore q_0 cannot blow up for any finite value of r. It will be seen, however, that q_0 is monotonically increasing with r. An easily derived upper bound gives an idea about its asymptotic behaviour. For the sake of simplicity, let us consider iid normal variables (whose covariance matrix is just the unit matrix). Then q_0 is proportional to the length of the portfolio vector [9]:

$$q_0^* = \sqrt{N \cdot \sum_{i=1}^{N} w_i^2} \tag{4}$$

$$q_0 = \left\langle q_0^* \right\rangle_{x_{it}} \tag{5}$$

As the weights sum to unity, this expression reaches its maximum when all the weights, except one, become zero. So $\sqrt{N}$ is an upper bound for q_0. If the weights are more or less uniformly distributed over the assets, q_0 is of order unity. When the weights become concentrated on a few assets, however, q_0 becomes of the order of magnitude $\sqrt{N}$, and reaches its upper bound when a single asset carries the total weight.

As a result of measurement noise, the portfolio vector will randomly rotate from sample to sample, according to the random fluctuations of its components. The parameter ϕ we introduce now is a measure of this random rotation.

$$\phi = \left\langle \frac{(\overrightarrow{w}_{true} \cdot \overrightarrow{w}_{meas})}{|\overrightarrow{w}_{meas}| \cdot |\overrightarrow{w}_{true}|} \right\rangle_{x_{it}} \tag{6}$$

Geometrically, ϕ is the normalised and averaged scalar product of the true portfolio vector and the measured one. (Remember that as the underlying process is generated by ourselves, the true portfolio vector is exactly known.) When ϕ is close to 1 the true portfolio vector and the measured one nearly coincide for most samples which corresponds to low estimation error. Conversely, when ϕ is close to 0, the true portfolio vector and the measured one are typically almost perpendicular, and the estimation error is large.

For iid variables the true normalized solution vector is

$$\overrightarrow{w}_{true} = \frac{1}{\sqrt{N}} \cdot (1, 1, 1..., 1) \tag{7}$$

The normalized measured solution vector is

$$\overrightarrow{w}_{meas} = \frac{(w_1, w_2, w_3..., w_N)}{\sqrt{\sum_{i=1}^{N} w_i^2}} \tag{8}$$

By virtue of Eq. (4);

we have

$$(\overrightarrow{w}_{meas} \cdot \overrightarrow{w}_{true}) = \frac{\sum_{i=1}^{N} w_i}{q_0^*} = \frac{1}{q_0^*} \tag{9}$$

Since in the limit $N \to \infty$ the fluctuations of q_0^* disappear, the average of this scalar product is

$$\phi = \Big\langle (\overrightarrow{w}_{meas} \cdot \overrightarrow{w}_{true}) \Big\rangle_{x_{it}} = \frac{1}{q_0} \tag{10}$$

As mentioned already in the introduction, the probability that the solution is found on the boundary of the allowed region increases with $r = N/T$. Without the constraints the solutions would run away. Besides, for different samples these solutions are found on different faces of the constraint polyhedron.

In order to measure the fraction of these type of solutions, we introduce the following measure:

$$\zeta = \left\langle \frac{100}{N} \sum_i \Big(H_\delta(w_i - a_i) + H_\delta(w_i - b_i) \Big) \right\rangle_{x_{it}} \tag{11}$$

where

$$H_\delta(x) = \begin{cases} 1 \text{ if } x < \delta \\ 0 \text{ otherwise,} \end{cases}$$

Thus, the measure ζ tells us what percentage of the portfolio weights falls into the neighbourhood of radius δ of the limits (a_i, b_i) imposed on the weights. In our simulations the value of δ was chosen to be $\frac{1}{N*100} = 0.0001$.

3 Results

3.1 Independent, identically distributed normal assets

In this subsection we consider iid normal assets, with no short selling allowed, that is with $a_i = 0$ and $b_i = 1$, $\forall i = 1...N$. If we had a very large amount of information for a fixed number of such assets, that is in the limit $T \to \infty$, $r \to 0$), by symmetry we would evidently find:

$$\overrightarrow{w} = \frac{1}{N}(1, ..., 1) \tag{12}$$

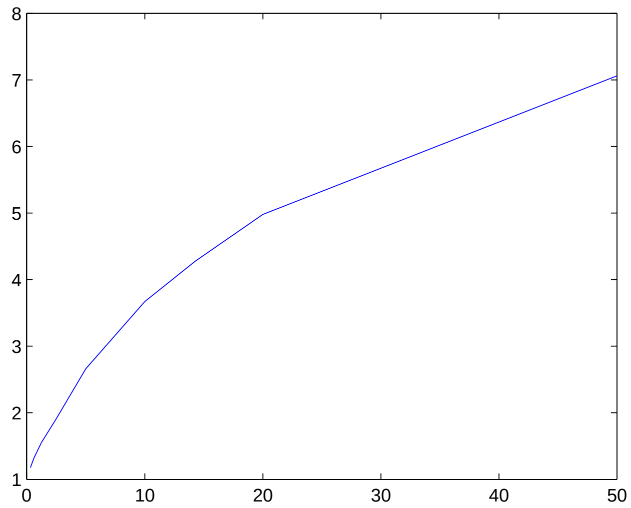

Fig. 1. q_0 as function of r for the minimal variance portfolio, with no short selling. (N=100, T=2,...,320), the figure was obtained by averaging over 5000 samples.

The vector Eq. (12); is then the true solution. For finite T we will find a different result that fluctuates from sample to sample.

3.1.1 Results for the minimum variance portfolio

The results of the numerical solution of the quadratic programming task described in Sec. 2 are summarized in the three subsequent figures. Fig. 1 shows the behaviour of q_0. As can be seen, q_0 is monotonically increasing with r which means that the portfolio weights are more and more concentrated on a few assets.

Fig. 3 displays a different aspect of the problem. It can be seen that with increasing $r = N/T$ the value of ϕ decreases and tends to zero, as it should, in view of Eq. (10) and q_0 increasing. But the vanishing of ϕ has a direct meaning: if the average scalar product of the true solution vector and the empirical one converges to 0, this means that the optimal solutions obtained for the different samples point in increasingly random directions.

Still another feature of the solutions is displayed in Fig. 3. With increasing r the percentage ζ of the solutions sitting on the walls tends to 100, which means that almost all portfolio weights are located close to a face of the polyhedron.

3.1.2 Results for Maximal Loss

The solution of the linear programming problem described in Sec. 2. is summarized in the next three figures. These figures tell the same story as the

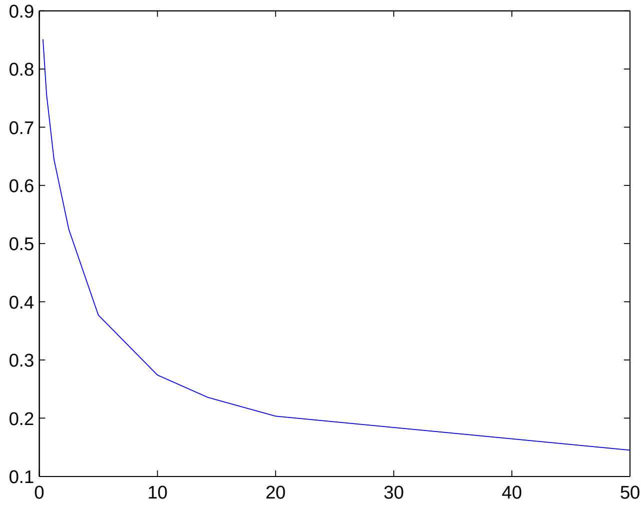

Fig. 2. The sample averaged scalar product ϕ of the true solution and the empirical one as function of r for the minimal variance, with no short selling allowed. (N=100, T=2,...,320), the plots were obtained by averaging over 5000 samples.

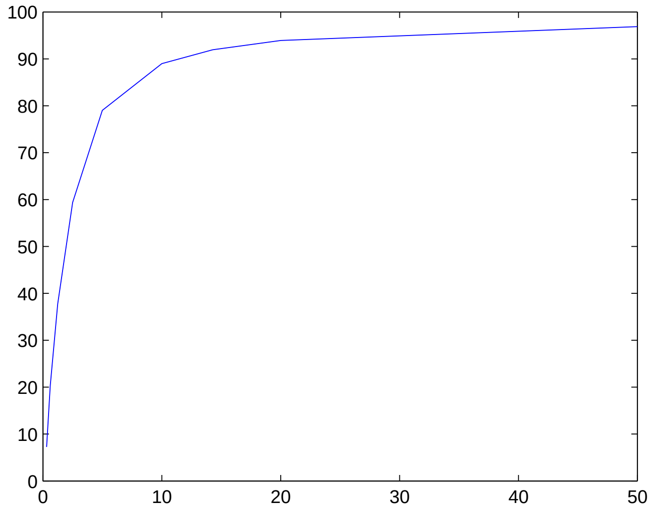

Fig. 3. The sample averaged fraction ζ of the solutions sticking to the boundary as function of r for the minimal variance, with no short selling allowed. (N=100, T=2,...,320), the plots were obtained by averaging over 5000 samples.

previous ones: an increasing number of weights stick to the walls and the average scalar product between the measured solution vector and the true one vanishes. The scaling behaviour of the solutions also seems similar, although we have not been able to push the simulations far enough into the asymptotic region to allow us to reliably deduce the exponent of the divergence of q_0, or the decay of ϕ

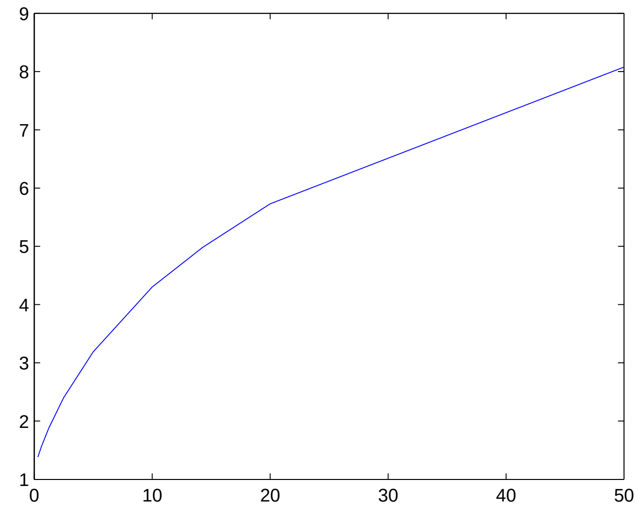

Fig. 4. q_0 as a function of r for Maximal Loss, with no short selling. (N=100, T=2,...,320, average over 5000 samples)

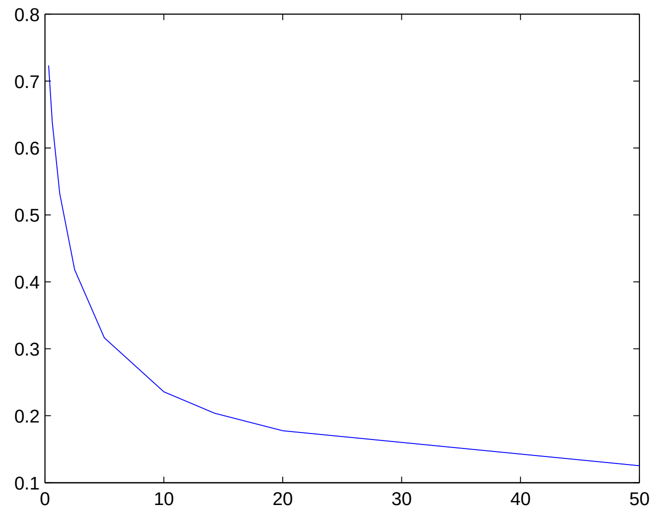

Fig. 5. The measure ϕ as functions of r for Maximal Loss, with no short selling. (N=100, T=2,...,320, average over 5000 samples.)

3.2 A toy model with correlations

In this section we go a little beyond the above iid case, and consider a toy model for a market with four industrial sectors, and different limits imposed on short positions in the assets belonging to these sectors.

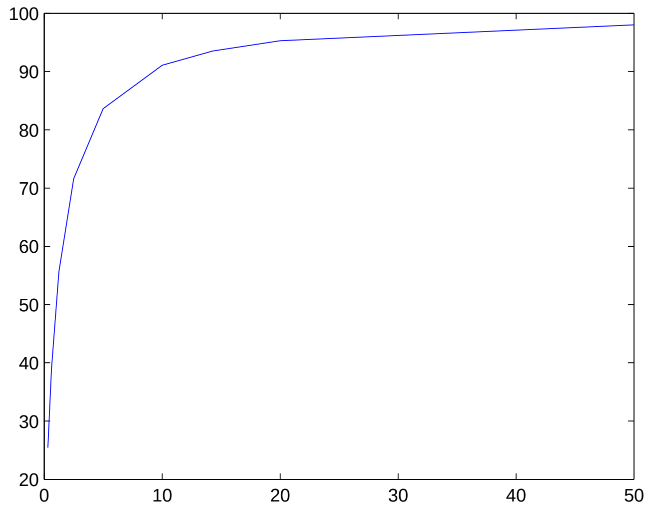

Fig. 6. The measures ζ as functions of r for Maximal Loss, with no short selling. (N=100, T=2,...,320, average over 5000 samples.)

3.2.1 Model description

The structure of the market will be represented by assuming a particular structure for the covariance matrix. For simplicity, we choose four sectors and assume that the covariances are constant within each sector, but different between the different sectors. We also assume that there are different correlations between different pairs of sectors. Thus the "true" covariance matrix we choose is the following:

$$C = \begin{pmatrix} \overline{\rho_1} & \rho & \rho & \rho \\ \rho & \overline{\rho_1} & \overline{\rho_2} & \rho \\ \rho & \overline{\rho_2} & \overline{\rho_1} & \rho \\ \rho & \rho & \rho & \overline{\rho_1} \end{pmatrix} \qquad \overline{\rho_n} = \begin{pmatrix} 1 & \rho_n & \dots & \rho_n \\ \rho_n & 1 & \dots & \rho_n \\ \dots & \dots & \dots & \dots \\ \rho_n & \rho_1 & \rho_n & 1 \end{pmatrix}$$

and $n = 1, 2$

We solve the problem with parameter values of $\rho_1 = 0.5$, $\rho_2 = -0.5$ $\rho = 0.1$, and optimize the portfolio under the Maximal Loss risk measure.

Let S_1, S_2, S_3, S_4 be the index set of sectors 1, 2, 3, 4, respectively, and let us have $N = 100$ assets, distributed over the the four sectors evenly, so that each of them have 25 assets belonging to them. The linear programming task is then:

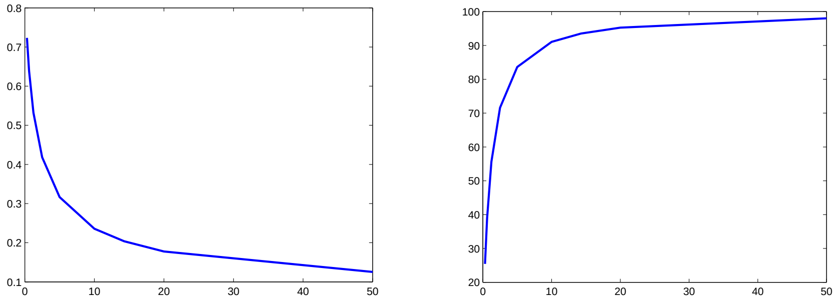

Fig. 7. Maximal Loss optimisation problem for the model defined in 13. The figures shows how ϕ(left hand side) and ζ(right hand side) depend on r. The figures been obtained by averaging over 5000 samples. (N=100, T=2,...,320)

$$\begin{aligned} &min \quad u \\ u = -max_t \sum_i & w_i \cdot x_{it} \\ \sum_{i \in \{1 \ldots N\}} & w_i = 1 \\ -1 \leq & w_{i \in S_1} \\ -3 \leq & w_{i \in S_3} \\ 0 \leq & w_{i \in S_3} \\ -0.5 \leq & w_{i \in S_4}, \end{aligned} \tag{13}$$

The solution of the problem for these constraints leads to the result that for sufficiently high r the parameters ϕ and ζ behave the same way as in the iid problem. The results are displayed in Fig. 7. They show that most of the portfolio weights stick to the walls of the constraint polyhedron, regardless of the fact that there are significant correlations in the model, and that this polyhedron differs from the one defined in Sec. 3.1.

4 Conclusion

Let us first summarize the results of this study: We have found that portfolio optimization under various linear constraints, risk measures and market models leads to solutions that are highly unstable, show strong sample to sample fluctuations, and unless a truly huge amount of information is available (that is unless r is very small), tend to stick to the walls defined by the constraints. In order to bring out the effects, we were considering also very high r values, which is clearly absurd from a portfolio management point of view, but we would like to stress that these effects start to show up already way below the critical value of r. In the limit of large r values our measurements clearly suggest that the various quantities considered here follow universal power laws, but we feel that we have not been able to push the simulations far enough into the asymptotically high N (or r) region to allow us to draw definitive conclusions as to the precise values of these exponents (which would be irrelevant for portfolio management anyhow).

Jagannathan and Ma [17] argue that linear constraints (the ban on short selling, in particular) act as a kind of shrinkage or filtering and lead to better behaved portfolios. We have seen that linear constraints do actually help taming the large fluctuations, especially by restricting the volume in the space of weights where these fluctuations take place. If we do not have a sufficient amount of information, however, these constraints will not be able to prevent the weights from fluctuating from sample to sample inside the allowed region, producing huge variations in the composition of the "optimal" portfolio, and sticking to the walls in large numbers. This means that the solutions will only partially reflect the covariance structure of the assets, they will, to a large extent, be determined by the constraints, in fact, by the portfolio managers who devised the limit system. This can be a dangerous state of affairs, especially if one attributes an undeserved significance to the optimal portfolios produced by a software package under those limits.

Acknowledgements

One of us (I. K.) is grateful to the Santa Fe Institute for the hospitality extended to him during the period of writing up this paper. This work has been supported by the "Cooperative Center for Communication Networks Data Analysis", a NAP project sponsored by the National Office of Research and Technology under grant No. KCKHA005.

References

[1] H. Markovitz, *Portfolio Selection: Efficient Diversification of Investments*, J.Wiley & Sons, New York (1959).

[2] E. J. Elton, M. J. Gruber, *Modern Portfolio Theory and Investment Analysis*, Wiley, New York (1995).

[3] P. Jorion, *Bayes-Stein Estimation for Portfolio Analysis*, Journal of Financial and Quantitative Analysis, **21**, 279-292, (1986).

[4] P. Frost and J. Savarino, *An Empirical Bayes Approach to Efficient Portfolio Selection*, Journal of Financial and Quantitative Analysis, **21**, 293-305 (1986).

[5] O. Ledoit and M. Wolf, *Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection*, Journal of Empirical Finance **10**, 603-621 (2003).

[6] L. Laloux, P. Cizeau, J.-P. Bouchaud and M. Potters, Int. J. Theor. Appl. Finance **3**, 391 (2000).

[7] S. Pafka and I. Kondor, *Noisy Covariance Matrices and Portfolio Optimization II*, Physica **A319C**, 487-494 (2003).

[8] S. Pafka and I. Kondor, *Estimated Correlation Matrices and Portfolio Optimization*, Physica **A343**, 623-634 (2004).

[9] I. Kondor, S. Pafka and G. Nagy, *Noise sensitivity of portfolio selection under various risk measures*, Journal of Banking and Finance, **31**, 1545-1573 (2007).

[10] S. Ciliberti, I. Kondor, M. Mezard, *On the Feasibility of Portfolio Optimization under Expected Shortfall*, Quantitative Finance, in print.

[11] S. Ciliberti, M. Mezard, *Risk Minimization through Portfolio Replication*, Eur Phys J B **57**, 175-180 (2007).

[12] I. Varga-Haszonits, I. Kondor, *Covariance Matrix Estimation and Portfolio Optimization in Constant Conditional Correlation GARCH models*, Physica A, in print

[13] Riskmetrics, Technical Document, J.P. Morgan and Reuters, *http://www.riskmetrics.com*

[14] B. Scherer, R. D. Martin, *Introduction to Modern Portflio Optimization With NUOPT and S-PLUS*, Springer, New York (2205)

[15] C. Acerbi, *Coherent representations of subjective risk-aversion*, In: Szegö, G. (Ed.), Risk measures for the 21st century. John Wiley & Sons (2004).

[16] C. Acerbi, D. Tasche, *On the coherence of expected shortfall*, Journal of Banking and Finance **26**, 1487-1503 (2002).

[17] R. Jagannathan, T. Ma, *Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps*, Journal of Finance, **58**, 1651-1684 (2003).