

## DRAWDOWN MEASURE IN PORTFOLIO OPTIMIZATION

ALEXEI CHEKHLOV

*Thor Asset Management, Inc., 551 Fifth Ave.  
Suite 601, 6th Floor, New York, NY 10017, USA  
achekhlov@growth.com*

STANISLAV URYASEV

*Department of Industrial and Systems Engineering, University of Florida  
P.O. Box 116595, 303 Weil Hall, Gainesville, FL 32611-6595, USA  
uryasev@ufl.edu*

MICHAEL ZABARANKIN

*Department of Industrial and Systems Engineering, University of Florida  
P.O. Box 116595, 303 Weil Hall, Gainesville, FL 32611-6595, USA  
zabarank@ufl.edu*

Received 2 March 2004

Accepted 25 May 2004

A new one-parameter family of risk measures called Conditional Drawdown (CDD) has been proposed. These measures of risk are functionals of the portfolio drawdown (underwater) curve considered in active portfolio management. For some value of the tolerance parameter  $\alpha$ , in the case of a single sample path, drawdown functional is defined as the mean of the worst  $(1 - \alpha) * 100\%$  drawdowns. The CDD measure generalizes the notion of the drawdown functional to a multi-scenario case and can be considered as a generalization of deviation measure to a dynamic case. The CDD measure includes the Maximal Drawdown and Average Drawdown as its limiting cases. Mathematical properties of the CDD measure have been studied and efficient optimization techniques for CDD computation and solving asset-allocation problems with a CDD measure have been developed. The CDD family of risk functionals is similar to Conditional Value-at-Risk (CVaR), which is also called Mean Shortfall, Mean Excess Loss, or Tail Value-at-Risk. Some recommendations on how to select the optimal risk functionals for getting practically stable portfolios have been provided. A real-life asset-allocation problem has been solved using the proposed measures. For this particular example, the optimal portfolios for cases of Maximal Drawdown, Average Drawdown, and several intermediate cases between these two have been found.

*Keywords:* Equity drawdown; drawdown measure; conditional value-at-risk; portfolio optimization; stochastic optimization.

### 1. Introduction

Optimal portfolio allocation is a longstanding issue in both practical portfolio management and academic research on portfolio theory. Various methods have been proposed and studied by Grinold [11]. All of them, as a starting point, assume

some measure of portfolio performance, which consists of at least two components: evaluating expected portfolio reward; and assessing expected portfolio risk. From theoretical perspective, there are two well-known approaches to manage portfolio performance: *Expected Utility Theory* and *Risk Management*, which are usually considered within a framework of a one-period or multi-period model.

If we are interested in Risk Management approach to portfolio optimization within a long term, what are the functionals for assessing portfolio risk that account for different sequences of portfolio losses? Let portfolio be optimized within time interval  $[0, T]$ , and let  $W(t)$  be portfolio value at time moment  $t \in [0, T]$ . One of the functionals that we are looking for is portfolio *drawdown* defined by  $(\max_{\tau \in [0, t]} W(\tau) - W(t))/W(t)$ , which, indeed, accounts for a sequence of portfolio losses. What are the advantages to formulate a portfolio optimization problem with a constraint on portfolio drawdown? To answer to this question, drawdown regulations in real trading strategies and drawdown theoretical aspects should be addressed first.

### **1.1. Drawdown regulations in real trading strategies**

From a standpoint of a fund manager, who trades clients' or bank's proprietary capital, and for whom the clients' accounts are the only source of income coming in the form of management and incentive fees, losing these accounts is equivalent to the death of his/her business. This is true with no regard to whether the employed strategy is long-term valid and has very attractive expected return characteristics. Such fund manager's primary concern is to keep the existing accounts and to attract the new ones in order to increase his/her revenues. Commodity Trading Advisor (CTA) determines the following rules regarding magnitude and duration of their clients' accounts drawdowns:

- Highly unlikely to tolerate a 50% drawdown in an account with an average- or small-risk CTA.
- An account may be shut down if a 20% drawdown is breached.
- A warning is issued if an account in a 15% drawdown.
- An account will be closed if it is in a drawdown, even of small magnitude, for longer than two years.
- Time to get out of a drawdown should not be longer than a year.

### **1.2. Drawdown notion in theoretical framework**

Several studies discussed portfolio optimization with drawdown constraints. Grossman and Zhou [12] obtained an exact analytical solution to portfolio optimization with constraint on maximal drawdown based on the following model:

- Continuous setup
- One-dimensional case — allocating current capital between one risky and one risk-free assets

- An assumption of log-normality of the risky asset
- Use of dynamic programming approach — finding a time-dependent fraction of the current capital invested into the risky asset

Cvitanic and Karatzas [7] generalized this model [12] to multi-dimension case (several risky assets). In contrast to Grossman and Zhou [12] and Cvitanic and Karatzas [7], Chekhlov *et al.* [6] defined portfolio drawdown to be the drop of the current portfolio value comparing to its maximum achieved in the past up to current moment  $t$ , i.e.,  $\max_{\tau \in [0, t]} W(\tau) - W(t)$ , and introduced one-parameter family of drawdown functionals, entitled Conditional Drawdown (CDD). Moreover, Chekhlov *et al.* [6] considered portfolio optimization with a constraint on drawdown functionals in a setup similar to the index tracking problem [8], where an index historical performance is replicated by a portfolio with constant weights. Chekhlov *et al.* [6] proposed the following setup:

- Discrete formulation
- Multi-dimensional case — several risky assets (markets and futures)
- A static set of portfolio weights satisfying a certain risk condition over the whole interval  $[0, T]$
- No assumption about the underlying probability distribution, which allows considering variety of practical applications — use of the historical sample paths of assets' rates of return over  $[0, T]$
- Use of linear programming approach — reduction of portfolio optimization to linear programming (LP) problem

The CDD is related to Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) measures studied by Rockafellar and Uryasev [20,21]. By definition, with respect to a specified probability level  $\alpha$ , the  $\alpha$ -VaR of a portfolio is the lowest amount  $\zeta_\alpha$  such that, with probability  $\alpha$ , the loss will not exceed  $\zeta_\alpha$  in a specified time  $\tau$ , whereas the  $\alpha$ -CVaR is the conditional expectation of losses above that amount  $\zeta_\alpha$ . Various issues about VaR methodology were discussed by Jorion [10]. The CDD is similar to CVaR and can be viewed as a modification of the CVaR to the case when the loss-function is defined as a drawdown. CDD and CVaR are conceptually related percentile-based risk performance functionals. Optimization approaches developed for CVaR are directly extended to CDD. The CDD includes the average drawdown and maximal drawdown as its limiting cases. It takes into account both the magnitude and duration of the drawdowns, whereas the maximal drawdown concentrates on a single event — maximal account's loss from its previous peak.

However, Chekhlov *et al.* [6] only tested the suggested approach to portfolio optimization subject to constraints on drawdown functionals. The CDD [6] was not defined as a true risk measure and the real-life portfolio optimization example was considered based only on the historical sample paths of assets' rates of return.

This paper is focused on:

- Concept of drawdown measure — possession of all properties of a deviation measure, generalization of deviation measures to a dynamic case
- Concept of risk profiling — Mixed Conditional Drawdown (generalization of CDD)
- Optimization techniques for CDD computation — reduction to linear programming (LP) problem
- Portfolio optimization with constraint on Mixed CDD

Our study develops concept of drawdown measure by generalizing the notion of the CDD to the case of several sample paths for portfolio uncompounded rate of return. Definition of drawdown measure is essentially based on the notion of CVaR [1,20,21] and *mixed* CVaR [22] extended to a multi-scenario case. Drawdown measure uses the concept of risk profiling introduced by Rockafellar *et al.* [22], namely, drawdown measure is a “multi-scenario” *mixed* CVaR applied to drawdown loss-function.

From theoretical perspective, drawdown measure satisfies the system of axioms determining *deviation measures* [22,23,24]. Those axioms are: *nonnegativity*, *insensitivity to constant shift*, *positive homogeneity* and *convexity*. Moreover, drawdown measure is an example generalizing properties of deviation measures to a dynamic case. We develop optimization techniques for efficient computation of drawdown measure in the case when instruments’ rates of return are given.

Similar to the Markowitz mean-variance approach [14], we formulate and solve an optimization problem with the reward performance function and CDD constraints. The reward-CDD optimization is a piece-wise linear convex optimization problem [19], which can be reduced to a linear programming problem (LP) using auxiliary variables.

Linear programming allows solving large optimization problems with hundreds of thousands of instruments. The algorithm is fast, numerically stable, and provides a solution during one run (without adjusting parameters like in genetic algorithms or neural networks). Linear programming approaches are routinely used in portfolio optimization with various criteria, such as mean absolute deviation [13], maximum deviation [25], and mean regret [8]. Ziemba and Mulvey [26] discussed other applications of optimization techniques in the finance area.

## 2. Model Development

Suppose a given time interval  $[0, T]$  is partitioned into  $N$  subintervals  $[t_{k-1}, t_k]$ ,  $k = \overline{1, N}$ , by the set of points  $\{t_0 = 0, t_1, t_2, \dots, t_N = T\}$ , and suppose there are  $m$  risky assets with rates of return determined by *random* vector  $r(t_k) = (r_1(t_k), r_2(t_k), \dots, r_m(t_k))$  at time moments  $t_k$  for  $k = \overline{1, N}$ . We also assume that the risk-free instrument (or cash) with the constant rate of return  $r_0$  is available. The  $i$ th asset’s rate of return at time moment  $t_k$  is defined by  $r_i(t_k) = \frac{p_i(t_k)}{p_i(t_{k-1})} - 1$ , where  $p_i(t_k)$  and  $p_i(t_{k-1})$  are the  $i$ th asset’s prices per share at moments  $t_k$  and

$t_{k-1}$ , respectively. Let  $C$  denote an initial capital at  $t_0 = 0$  and let values  $x_i(t_k)$  for  $i = \overline{1, m}$  and  $x_0(t_k)$  define the proportion of the current capital invested in the  $i$ th risky asset and risk-free instrument at  $t_k$ , respectively. Consequently, a portfolio formed of the  $m$  risky assets and the risk-free instrument is determined by the vector of weights  $x(t_k) = (x_0(t_k), x_1(t_k), x_2(t_k), \dots, x_m(t_k))$ . The components of  $x(t_k)$  satisfy the budget constraint

$$\sum_{i=0}^m x_i(t_k) = 1. \tag{2.1}$$

By definition, the rate of return of the portfolio at time moment  $t_k$  is

$$r_k^{(p)}(x(t_k)) = r(t_k) \cdot x(t_k) = \sum_{i=0}^m r_i(t_k)x_i(t_k). \tag{2.2}$$

Portfolio optimization can be considered within a framework of a one-period or multi-period model. A *one-period* model in portfolio optimization assumes the  $i$ th asset's rates of return for all  $t_k$ ,  $k = \overline{1, N}$ , to be independent observations of a random variable  $r_i$ . In this case, the vector of portfolio weights is constant and portfolio rate of return is a random variable  $r^{(p)}$  presented by a linear combination of random assets' rates of return  $r_i$ ,  $i = \overline{1, m}$ , and constant  $r_0$ , i.e.,  $r^{(p)} = \sum_{i=0}^m r_i x_i$ . A traditional setup for a one-period portfolio optimization problem from Risk Management point of view is maximizing portfolio expected rate of return subject to the budget constraint and a constraint on the risk

$$\begin{aligned} & \max_x E(r^{(p)}) \\ & \text{s.t. Risk}(r^{(p)}) \leq d, \\ & \sum_{i=0}^m x_i = 1. \end{aligned} \tag{2.3}$$

Risk of the portfolio can be measured by different performance functionals, depending on investor's risk preferences. Variance, VaR, CVaR and Mean Absolute Deviation (MAD) are examples of risk functionals used in portfolio Risk Management [22]. Certainly, solving optimization problem (2.3) with different risk measures will lead to different optimal portfolios. However, all of them are based on a one-period model, which does not take into account the sequence of the asset's rates of return within time interval  $[0, T]$ .

A *multi-period* model in portfolio optimization is intended for controlling and optimizing portfolio wealth over a long term. It is essentially based on how the asset's rates of return evolve within the whole time interval. Moreover, in each time moment  $t_k$ ,  $k = \overline{0, N}$ , there might be a capital inflow or outflow into or from the portfolio, and portfolio weights  $x_i(t_k)$ ,  $i = \overline{1, m}$ , might be re-balanced. In this case, the portfolio wealth at  $t_k$  for  $k = \overline{1, N}$  is defined

$$W_k(x(t_k)) = (W_{k-1}(x(t_{k-1})) + Y(t_{k-1}))(1 + r_k^{(p)}(x(t_k))), \tag{2.4}$$

where  $Y(t_{k-1}) = F_+(t_{k-1}) - F_-(t_{k-1})$  is the resulting capital flow at  $t_{k-1}$  (inflow  $F_+(t_{k-1})$  minus outflow  $F_-(t_{k-1})$ ), which can be positive or negative.

A portfolio optimization problem can also be formulated based on the Expected Utility Theory (EUT). According to the EUT, an investor with additively separable concave utility function  $U(\cdot)$  chooses a consumption stream  $\{C_0, C_1, \dots, C_{N-1}\}$  and portfolio to maximize

$$E \left( \sum_{k=0}^{N-1} U(C(t_k), t_k) + B(W(t_N), x(t_N)) \right),$$

where  $B(\cdot)$  is the concave utility of bequest. Note that the EUT is focused on maximization of investor's consumption. However, a risk manager who runs a hedge fund and wishes to increase capital inflow by attracting new investors would be more interested in maximizing portfolio wealth at the final moment  $t_N = T$  and decreasing portfolio drops over the whole time interval  $[0, T]$ . In this case, Risk Management approach is more adequate to formulate a portfolio optimization problem

$$\begin{aligned} \max_x \quad & \mathcal{P}(W) \\ \text{s.t.} \quad & \mathcal{R}(W) \leq d, \\ & \sum_{i=0}^m x_i(t_k) = 1, \quad k = \overline{0, N}, \end{aligned} \tag{2.5}$$

where  $\mathcal{P}(W)$  and  $\mathcal{R}(W)$  are performance and risk functionals, respectively, depending on stream  $W = (W_1, W_2, \dots, W_N)$ . Suppose the optimization problem (2.5) is considered under the following conditions:

- A manager cannot affect a stream of  $Y(t_k)$  (if the portfolio value increases it is likely that capital inflow will also increase and vice-versa).
- The manager can only allocate resources among different instruments (investment strategies) in the portfolio at every moment  $t_k$ ,  $k = \overline{0, N}$ , i.e., he/she can only optimize portfolio rate of return by choosing portfolio weights  $x_i(t_k)$ .

Accounting for these conditions, how can the manager evaluate portfolio performance over  $[0, T]$  and efficiently solve (2.5)? Before to answer to this question, the following legitimate issues regarding problem formulation (2.5) should be addressed

- How the risk is measured within  $[0, T]$ .
- How the assets' rates of return are modeled within  $[0, T]$ .
- What optimization approach is chosen to solve (2.5).

This paper considers several integral characteristics of portfolio performance, which distinguish different sequences of  $W_k$  in a stream  $(W_1, W_2, \dots, W_N)$ . These characteristics are based on the notion of portfolio drawdown dealing with the drop in portfolio wealth at time moment  $t_k$  with respect to the wealth's maximum preceding  $t_k$ . Pflug and Ruszczyński [16,18] discuss alternative formulations for (2.5) as well as some approaches for defining risk of multi-period income streams.

### 3. Absolute Drawdown for a Single Sample Path

This section presents the notion of the Absolute Drawdown ( $\mathcal{AD}$ ) and considers three functionals based on this notion. The  $\mathcal{AD}$  is applied to a sample path of the uncompounded cumulative portfolio rate of return. Note that the  $\mathcal{AD}$  is applied not to the compounded cumulative portfolio rate of return  $W_k(x(t_k))$ . If the values of  $r_k^{(p)}(x(t_k))$  for  $k = \overline{1, N}$  determine a sample path (time series) of the portfolio's rate of return, then, by definition, the *uncompounded cumulative* portfolio rate of return at time moment  $t_k$  is

$$w_k(x(t_k)) = \begin{cases} 0, & k = 0, \\ \sum_{l=1}^k r_l^{(p)}(x(t_l)), & k = \overline{1, N}. \end{cases} \quad (3.1)$$

To simplify notations, we use  $w_k$  instead of  $w_k(x(t_k))$ , assuming that  $w_k$  is always a function of vector  $x(t_k)$ . Further in this section, we consider only a single sample path of  $w_k$ ,  $k = \overline{1, N}$ , which we denote by vector  $w$ , i.e.  $w = (w_1, \dots, w_N)$ .

**Definition 3.1.** The  $\mathcal{AD}$  is a vector-variable functional depending on the sample path  $w$

$$\mathcal{AD}(w) = \xi = (\xi_1, \dots, \xi_N), \quad \xi_k = \max_{0 \leq j \leq k} \{w_j\} - w_k. \quad (3.2)$$

Note that components  $(w_1, \dots, w_N)$  and  $(\xi_1, \dots, \xi_N)$  of vectors  $w$  and  $\xi$ , are, in fact, time series  $w_1, \dots, w_N$  and  $\xi_1, \dots, \xi_N$ , respectively, where the  $k$ th components of  $w$  and  $\xi$  correspond to time moment  $t_k$ . Since  $\xi_0$  is always zero, we do not include it into drawdown time series  $\xi$ . Moreover, although  $\mathcal{AD}(w)$  and  $\xi$  are the same drawdown time series, we refer to notation  $\mathcal{AD}(w)$  to emphasize its dependence on  $w$  and to notation  $\xi$  whenever we use drawdown time series just as vector of numbers.

Figure 1 illustrates an example of the absolute drawdown  $\xi$  and a corresponding sample path of uncompounded cumulative rate of return  $w$ . Starting from  $t_0 = 0$ ,

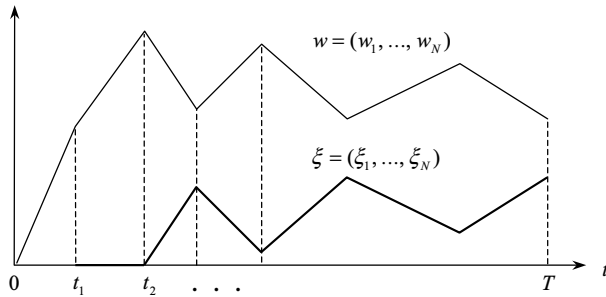


Fig. 1. Time series of uncompounded cumulative rate of return  $w$  and corresponding absolute drawdown  $\xi$ .

uncompounded cumulative rate of return  $w$  goes up and the first component of  $\xi$  equals zero. When  $w$  decreases,  $\xi$  goes up. When time series  $w$  achieves its local minimum, absolute drawdown achieves its local maximum. This process continues until  $t_N = T$ .

**Proposition 3.1.** *Defining vectorial operations:  $w + \text{const} = (w_1 + \text{const}, \dots, w_N + \text{const})$  and  $\lambda w = (\lambda w_1, \dots, \lambda w_N)$ , the  $\mathcal{AD}(w)$  satisfies the following properties*

- (1) *Nonnegativity:  $\mathcal{AD}(w) \geq 0$ .*
- (2) *Insensitivity to constant shift:  $\mathcal{AD}(w + \text{const}) = \mathcal{AD}(w)$ .*
- (3) *Positive homogeneity:  $\mathcal{AD}(\lambda w) = \lambda \mathcal{AD}(w)$ ,  $\forall \lambda \geq 0$ .*
- (4) *Convexity: if  $w_\lambda = \lambda w_a + (1 - \lambda)w_b$  is a linear combination of any two sample paths of uncompounded cumulative rates of return,  $w_a$  and  $w_b$ , with  $\lambda \in [0, 1]$ , then  $\mathcal{AD}(w_\lambda) \leq \lambda \mathcal{AD}(w_a) + (1 - \lambda)\mathcal{AD}(w_b)$ .*

**Proof.** Properties 1–3 are direct consequences of (3.2). Property 4 is proved based on  $\max_{0 \leq \tau \leq t} \{\lambda w_a + (1 - \lambda)w_b\} \leq \lambda \max_{0 \leq \tau \leq t} \{w_a\} + (1 - \lambda) \max_{0 \leq \tau \leq t} \{w_b\}$ ,  $\lambda \in [0, 1]$ .  $\square$

Note that  $\mathcal{DD}$  does not satisfy the properties which  $\mathcal{AD}$  does (advantage of  $\mathcal{AD}$ ). The difference between the  $\mathcal{AD}$  and  $\mathcal{DD}$  is similar to the difference between absolute and relative errors in a measurement. The  $\mathcal{AD}$  and  $\mathcal{DD}$  functionals can be used in Risk Management and Statistics to control absolute and relative drops in a realization of a stochastic process. However, in this paper we are focused on applications of drawdown functionals in portfolio optimization. Since further in this paper, we deal only with the absolute drawdown functional,  $\mathcal{AD}$ , the word “absolute” can be omitted without confusion.

### 3.1. Maximum, average and conditional drawdowns

We consider three functionals based on the notion of drawdown: (i) Maximum Drawdown (MaxDD), (ii) Average Drawdown (AvDD), and (iii) CDD. The last risk functional is actually a family of performance functions depending upon parameter  $\alpha$ . It is defined similar to CVaR [21] and, as special cases, includes the MaxDD and AvDD.

**Definition 3.2.** For given time interval  $[0, T]$ , partitioned into  $N$  subintervals  $[t_{k-1}, t_k]$ ,  $k = \overline{1, N}$ , with  $t_0 = 0$  and  $t_N = T$ , AvDD and MaxDD functionals are defined, respectively

$$\text{MaxDD}(w) = \max_{1 \leq k \leq N} \{\xi_k\}, \quad (3.3)$$

$$\text{AvDD}(w) = \frac{1}{N} \sum_{k=1}^N \xi_k. \quad (3.4)$$



To define Conditional Value-@-Risk (CV@R) and CDD, we introduce a function  $\pi_\xi(s)$  such that

$$\pi_\xi(s) = \frac{1}{N} \sum_{k=1}^N I_{\{\xi_k \leq s\}}, \tag{3.5}$$

where  $I_{\{\xi_k \leq s\}}$  is an indicator equal to 1, if the condition in curly brackets is true, and equal to zero, if the condition is false, i.e.,

$$I_{\{c \leq s\}} = \begin{cases} 1, & c \leq s, \\ 0, & c > s, \end{cases} \quad c \in \mathbf{R}.$$

Figure 2 explains definition of function  $\pi_\xi(s)$ . For the threshold  $s$  shown on the figure, function  $\pi_\xi(s)$  equals  $\frac{5}{8}$ , since  $\xi_k \leq s$  for five values of  $k$ , namely,  $k = 2, 3, 4, 7, 8$ .

The inverse function to (3.5) is defined

$$\pi_\xi^{-1}(\alpha) = \begin{cases} \inf\{s | \pi_\xi(s) \geq \alpha\}, & \alpha \in (0, 1], \\ 0, & \alpha = 0. \end{cases} \tag{3.6}$$

**Remark 3.1.** Since all  $\xi_k, k = \overline{1, N}$ , are nonnegative, we define  $\pi_\xi^{-1}(0)$  to be zero.

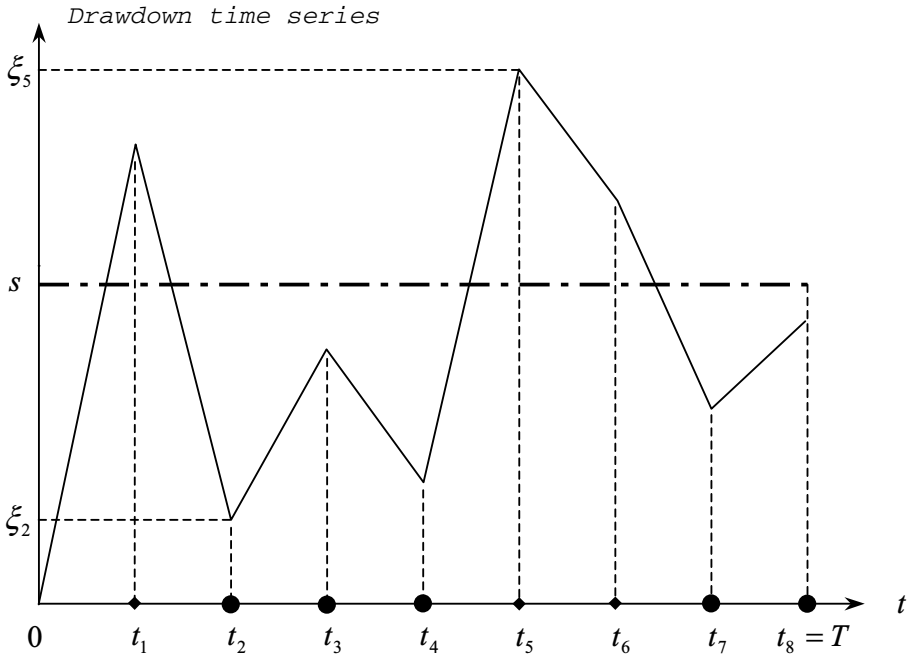


Fig. 2. Drawdown time series  $\xi$  and indicator function  $I_{\{c \leq s\}}$ .

**Remark 3.2.** In fact,  $\forall \alpha \in (0, 1]$ ,  $s = \pi_\xi^{-1}(\alpha)$  is the unique solution to two inequalities

$$\pi_\xi(s - 0) < \alpha \leq \pi_\xi(s + 0). \tag{3.7}$$

Figures 3 and 4 illustrate left and right continuous step functions  $\pi_\xi(s)$  and  $\pi_\xi^{-1}(\alpha)$ , respectively, which correspond to drawdown time series  $\xi$  shown on Fig. 2.

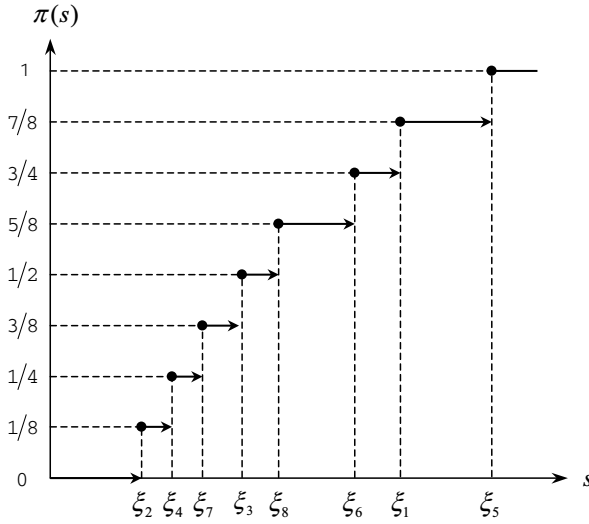


Fig. 3. Function  $\pi_\xi(s)$ .

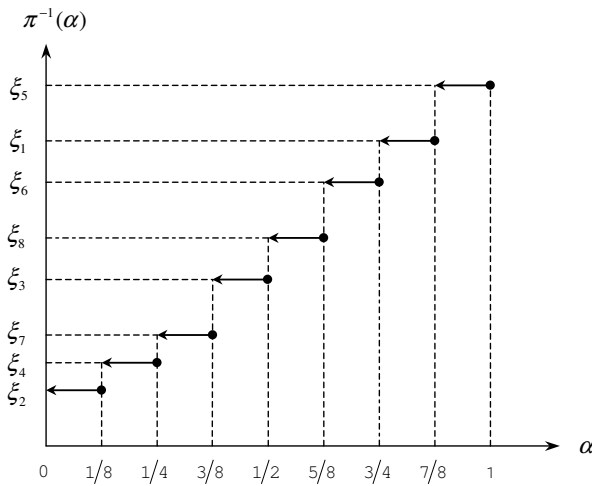


Fig. 4. Inverse function  $\pi_\xi^{-1}(\alpha)$ .

Let  $\zeta(\alpha)$  be a threshold such that  $(1 - \alpha) * 100\%$  of drawdowns exceed this threshold. By definition,

$$\zeta(\alpha) = \pi_{\xi}^{-1}(\alpha). \tag{3.8}$$

If we are able to precisely count  $(1 - \alpha) * 100\%$  of the worst drawdowns, then  $\pi_{\xi}(\zeta(\alpha)) = \pi_{\xi}(\pi_{\xi}^{-1}(\alpha)) = \alpha$ . For such a value of the parameter  $\alpha$ , the CV@R of  $\xi_k$ ,  $k = \overline{1, N}$ , is defined as the mean of the worst  $(1 - \alpha) * 100\%$  drawdowns. For instance, if  $\alpha = 0$ , then CV@R is the average drawdown, and if  $\alpha = 0.95$ , then CV@R is the average of the worst 5% drawdowns. However, in a general case,  $\pi_{\xi}(\zeta(\alpha)) = \pi_{\xi}(\pi_{\xi}^{-1}(\alpha)) \geq \alpha$ , followed from (3.6). It means that, in general, we are not able to precisely count  $(1 - \alpha) * 100\%$  of the worst drawdowns. In this case, the CV@R becomes a weighted average of the threshold  $\zeta(\alpha)$  and the mean of the worst drawdowns strictly exceeding  $\zeta(\alpha)$ .

**Definition 3.3.** For a given sequence of  $\xi_k$ ,  $k = \overline{1, N}$ , CV@R is formally defined by

$$\text{CV@R}_{\alpha}(\xi) = \left( \frac{\pi_{\xi}(\zeta(\alpha)) - \alpha}{1 - \alpha} \right) \zeta(\alpha) + \frac{1}{(1 - \alpha)N} \sum_{\xi_k \in \Xi_{\alpha}} \xi_k, \tag{3.9}$$

where  $\Xi_{\alpha} = \{\xi_k | \xi_k > \zeta(\alpha), k = \overline{1, N}\}$ .

Note that the first term in the right-hand side of (3.9) appears because of inequality  $\pi_{\xi}(\pi_{\xi}^{-1}(\alpha)) \geq \alpha$ . If  $(1 - \alpha) * 100\%$  of the worst drawdowns can be counted precisely, then  $\pi_{\xi}(\pi_{\xi}^{-1}(\alpha)) = \alpha$  and the first term in the right-hand side of (3.9) disappears. Equation (3.9) follows from the framework of the CVaR methodology [20,21]. Close relation between the CVaR and CV@R is discussed in the following remark.

**Remark 3.3.** CV@R $_{\alpha}$ , given by (3.9), and functional CVaR $_{\alpha}$  [22, p. 7, example 4], are linearly dependent, i.e., if  $X$  is an arbitrary random variable then

$$\text{CV@R}_{\alpha}(X) = \frac{1}{1 - \alpha} (E(X) + \alpha \text{CVaR}_{\alpha}(X)). \tag{3.10}$$

Thus, use of the CV@R or CVaR is only the matter of convenience.

**Definition 3.4.** In a single scenario case, the CDD with tolerance level  $\alpha \in [0, 1]$  is the CV@R applied to the drawdown functional,  $\mathcal{AD}(w)$ ,

$$\Delta_{\alpha}(w) = \text{CV@R}_{\alpha}(\mathcal{AD}(w)). \tag{3.11}$$

Equivalently, interpreting  $\xi_k$ ,  $k = \overline{1, N}$ , to be observations of a “random variable”  $\xi$ ,  $\alpha$ -CDD is the CV@R $_{\alpha}$  of a loss function  $\mathcal{AD}(w)$ .

### 3.2. Conditional Value-at-Risk and Conditional Drawdown properties

CDD is an example of a functional generalizing properties of deviation measures to a dynamic case. However, since CDD is closely related to CVaR, which properties were studied in detail by Rockafellar and Uryasev [20,21], it is useful to discuss CDD properties based on properties of CVaR. Because of linear relation (3.10), we can replace CVaR by CV@R.

**Proposition 3.2.** *CV@R $_{\alpha}(\xi)$  satisfies the following properties*

- (1) *Constant translation:*  $\text{CV@R}_{\alpha}(\xi + \text{const}) = \text{CV@R}_{\alpha}(\xi) + \text{const}$ ,  $\forall \alpha \in [0, 1]$ .
- (2) *Positive homogeneity:*  $\text{CV@R}_{\alpha}(\lambda\xi) = \lambda \text{CV@R}_{\alpha}(\xi)$ ,  $\forall \lambda \geq 0$  and  $\forall \alpha \in [0, 1]$ .
- (3) *Monotonicity:* if  $\xi_k \leq \eta_k$ ,  $1 \leq k \leq N$ , then  $\text{CV@R}_{\alpha}(\xi) \leq \text{CV@R}_{\alpha}(\eta)$ ,  $\forall \alpha \in [0, 1]$ .
- (4) *Convexity:* if  $\xi_{\lambda} = \lambda\xi_a + (1 - \lambda)\xi_b$  is a linear combination of any two drawdown sample paths  $\xi_a$  and  $\xi_b$  with  $\lambda \in [0, 1]$ , then  $\text{CV@R}_{\alpha}(\xi_{\lambda}) \leq \lambda \text{CV@R}_{\alpha}(\xi_a) + (1 - \lambda) \text{CV@R}_{\alpha}(\xi_b)$ .

**Proof.** Based on linear relation between CV@R $_{\alpha}$  and CVaR $_{\alpha}$ , given by (3.10), properties 1–4 are direct consequence of CVaR $_{\alpha}$  properties [22].  $\square$

**Proposition 3.3.** *The CDD =  $\Delta_{\alpha}(w)$  satisfies the properties of deviation measures, i.e.,*

- (1) *Nonnegativity:*  $\Delta_{\alpha}(w) \geq 0$ ,  $\forall \alpha \in [0, 1]$ .
- (2) *Insensitivity to constant shift:*  $\Delta_{\alpha}(w + \text{const}) = \Delta_{\alpha}(w)$ ,  $\forall \alpha \in [0, 1]$ .
- (3) *Positive homogeneity:*  $\Delta_{\alpha}(\lambda w) = \lambda \Delta_{\alpha}(w)$ ,  $\forall \lambda \geq 0$  and  $\forall \alpha \in [0, 1]$ .
- (4) *Convexity:* if  $w_{\lambda} = \lambda w_a + (1 - \lambda)w_b$  is a linear combination of any two sample paths of uncompounded cumulative rate of returns  $w_a$  and  $w_b$  with  $\lambda \in [0, 1]$ , then  $\Delta_{\alpha}(w_{\lambda}) \leq \lambda \Delta_{\alpha}(w_a) + (1 - \lambda) \Delta_{\alpha}(w_b)$ .

**Proof.** Properties 1–4 follow from Propositions 3.1 and 3.2. Indeed, based on the relation between the CDD and CV@R, i.e.,  $\Delta_{\alpha}(w) = \text{CV@R}_{\alpha}(\mathcal{AD}(w))$ , the first property is a direct consequence of  $\mathcal{AD}(w)$  nonnegativity. Properties 2–4 are proved, respectively,

$$\Delta_{\alpha}(w + c) = \text{CV@R}_{\alpha}(\mathcal{AD}(w + c)) = \text{CV@R}_{\alpha}(\mathcal{AD}(w)) = \Delta_{\alpha}(w),$$

$$\begin{aligned} \Delta_{\alpha}(\lambda w) &= \text{CV@R}_{\alpha}(\mathcal{AD}(\lambda w)) = \text{CV@R}_{\alpha}(\lambda \mathcal{AD}(w)) \\ &= \lambda \text{CV@R}_{\alpha}(\mathcal{AD}(w)) = \lambda \Delta_{\alpha}(w), \end{aligned}$$

$$\begin{aligned} \Delta_{\alpha}(w_{\lambda}) &= \text{CV@R}_{\alpha}(\mathcal{AD}(\lambda w_a + (1 - \lambda)w_b)) \\ &\leq \text{CV@R}_{\alpha}(\lambda \mathcal{AD}(w_a) + (1 - \lambda)\mathcal{AD}(w_b)) \end{aligned}$$

$$\begin{aligned} &\leq \lambda \text{CV@R}_\alpha(\mathcal{AD}(w_a)) + (1 - \lambda)\text{CV@R}_\alpha(\mathcal{AD}(w_b)) \\ &= \lambda\Delta_\alpha(w_a) + (1 - \lambda)\Delta_\alpha(w_b). \end{aligned}$$

Note that the monotonicity property of  $\text{CV@R}$  is used in the first line of the proof of CDD convexity.  $\square$

**Proposition 3.4.** *MaxDD (3.3) and AvDD (3.4) are the special cases of the  $\alpha$ -CDD functional (this notation is used to emphasize CDD dependence on  $\alpha$ ), namely,*

$$\text{MaxDD}(w) = \Delta_1(w), \quad \text{AvDD}(w) = \Delta_0(w). \quad (3.12)$$

**Proof.** To prove the first formula of (3.12), we assume that  $\zeta(1) < \infty$ . Based on this assumption, in the case of  $\alpha = 1$ , we have  $\zeta(1-) = \pi_\xi^{-1}(1-) = \pi_\xi^{-1}(1) = \zeta(1)$ , i.e., function  $\zeta(\alpha)$  is constant in the left vicinity of 1. Hence,  $\pi_\xi(\zeta(1-)) = \pi_\xi(\zeta(1)) = 1$ ,  $\Xi_1 = \emptyset$  and

$$\Delta_1(w) = \zeta(1) \lim_{\alpha \rightarrow 1^-} \left( \frac{\pi_\xi(\zeta(\alpha)) - \alpha}{1 - \alpha} \right) = \zeta(1) \lim_{\alpha \rightarrow 1^-} \left( \frac{1 - \alpha}{1 - \alpha} \right) = \zeta(1) = \text{MaxDD}(w).$$

When  $\alpha = 0$ , according to (3.6),  $\zeta(0) = 0$ ,  $\Xi_0 = \{\xi_k | k = \overline{1, N}\}$  and, consequently,

$$\Delta_0(w) = \frac{1}{N} \sum_{t_k \in \Xi_0} \xi_k = \frac{1}{N} \sum_{k=1}^N \xi_k = \text{AvDD}(w). \quad \square$$

**Theorem 3.1.**  *$\text{CV@R}_\alpha(\xi)$  can be presented in the alternative form*

$$\text{CV@R}_\alpha(\xi) = \frac{1}{1 - \alpha} \int_\alpha^1 \pi_\xi^{-1}(q) dq, \quad (3.13)$$

which is mathematically equivalent to (3.9).

**Proof.** Let  $\{s_j | j = \overline{1, J}\}$  be the set of the ordered values of  $\xi_k$ ,  $k = \overline{1, N}$ , where  $J$  is the number of different values of  $\xi_k$ ,  $k = \overline{1, N}$ , such that  $s_1 < s_2 < \dots < s_J$  and  $n_j \geq 1$  is the multiplicity of  $s_j$ , i.e.,  $n_j = \sum_{k=1}^N I_{\{\xi_k = s_j\}}$  and  $\sum_{j=1}^J n_j = N$ . Defining  $q_j = \frac{1}{N} \sum_{l=1}^j n_l$ , step functions  $\pi_\xi$  and  $\pi_\xi^{-1}$  are determined by the set of  $(s_j, q_j)$ ,  $j = \overline{1, J}$ , i.e.,

$$\pi_\xi(s_j) = q_j, \quad \pi_\xi^{-1}(q_j) = s_j. \quad (3.14)$$

Let  $s_0 = 0$  and  $q_0 = 0$ , then since  $\bigcap_{j=1}^J (q_{j-1}, q_j] = \emptyset$  and  $\bigcup_{j=1}^J (q_{j-1}, q_j] = (0, 1]$ , for any value of  $\alpha \in (0, 1]$ , there exists  $j^*$  from  $\overline{1, J}$  such that  $\alpha \in (q_{j^*-1}, q_{j^*}]$ . Using (3.14) and condition  $\alpha \in (q_{j^*-1}, q_{j^*}]$ , we obtain

$$\zeta(\alpha) = s_{j^*}, \quad \pi_\xi(\zeta(\alpha)) = q_{j^*},$$

and, consequently,

$$\frac{1}{N} \sum_{t_k \in \Xi_\alpha} \xi_k = \frac{1}{N} \sum_{j=j^*+1}^J s_j n_j = \sum_{j=j^*+1}^J \pi_\xi^{-1}(q_j)(q_j - q_{j-1}) = \int_{q_{j^*}}^1 \pi_\xi^{-1}(q) dq.$$

Taking the last relations into account, for any  $\alpha \in (0, 1)$ , the integral in the right-hand side of (3.13) is presented

$$\int_{\alpha}^1 \pi_{\xi}^{-1}(q) dq = (q_{j^*} - \alpha) s_{j^*} + \int_{q_{j^*}}^1 \pi_{\xi}^{-1}(q) dq = (\pi_{\xi}(\zeta(\alpha)) - \alpha) \zeta(\alpha) + \frac{1}{N} \sum_{t_k \in \Xi_{\alpha}} \xi_k,$$

which coincides with the expression (3.9) with accuracy of multiplier  $(1 - \alpha)^{-1}$ .

Only two cases are left to consider, namely, when  $\alpha = 0$  and  $\alpha = 1$ . Assuming  $\pi_{\xi}^{-1}(1) < \infty$ , we have, respectively,

$$\Delta_0(w) = \int_0^1 \pi_{\xi}^{-1}(q) dq = \frac{1}{N} \sum_{j=1}^J n_j s_j = \frac{1}{N} \sum_{k=1}^N \xi_k = \text{AvDD}(w),$$

$$\Delta_1(w) = \lim_{\alpha \rightarrow 1} \left( \frac{1}{1 - \alpha} \int_{\alpha}^1 \pi_{\xi}^{-1}(q) dq \right) = \pi_{\xi}^{-1}(1) = \text{MaxDD}(w). \quad \square$$

**Remark 3.4.** Let  $X$  be an arbitrary random variable with the cumulative distribution function  $F_X(t) = \Pr\{X \leq t\}$ . Assuming  $F_X^{-1}(\alpha)$  to be the inverse function of  $F_X(t)$ , functionals  $\text{CV@R}_{\alpha}$  and  $\text{CVaR}_{\alpha}$  are expressed, respectively,

$$\text{CV@R}_{\alpha}(X) = \frac{1}{1 - \alpha} \int_{\alpha}^1 F_X^{-1}(q) dq, \quad \text{CVaR}_{\alpha}(X) = -\frac{1}{\alpha} \int_0^{\alpha} F_X^{-1}(q) dq. \quad (3.15)$$

Relation (3.10) can be verified based on (3.15). CVaR methodology was thoroughly developed by Rockafellar and Uryasev [20,21].

**Example 3.1.** To illustrate the concept of the CV@R, let us calculate  $\text{CV@R}_{0.7}(\xi)$  for drawdown time series  $\xi$  shown on Fig. 2. According to Fig. 4,  $\zeta(0.7) = \pi_{\xi}^{-1}(0.7) = \xi_6$ , and, consequently, from Fig. 3,  $\pi_{\xi}(\zeta(0.7)) = \pi_{\xi}(\xi_6) = 0.75$ . Using formula (3.9), we obtain  $\text{CV@R}_{0.7}(\xi) = \frac{(0.75 - 0.7)}{1 - 0.7} \xi_6 + \frac{1}{1 - 0.7} \frac{(\xi_1 + \xi_5)}{8} = \frac{1}{6} \xi_6 + \frac{5}{12} (\xi_1 + \xi_5)$ . To verify this result, we can calculate  $\text{CV@R}_{0.7}(\xi)$  based on (3.13). Namely, following Fig. 4, we have  $\text{CV@R}_{0.7}(\xi) = \frac{1}{1 - 0.7} ((0.75 - 0.7) \xi_6 + (0.875 - 0.75) \xi_1 + (1 - 0.875) \xi_5) = \frac{1}{6} \xi_6 + \frac{5}{12} \xi_1 + \frac{5}{12} \xi_5$ .

**Example 3.2.** For the drawdown time series shown on Fig. 2,  $\text{MaxDD}(w) = \xi_5$  and  $\text{AvDD}(w) = \frac{1}{8} \sum_{k=1}^8 \xi_k$ .

### 3.3. Mixed conditional drawdown

The notion of CDD can be generalized by considering convex combinations of the CDDs corresponding to different confidence levels. This idea is essentially based on risk profiling, i.e., assignment of specific weights for CDDs with predetermined confidence levels.

**Definition 3.5.** Given a risk profile  $\chi(\alpha)$  such that

1.  $d\chi(\alpha) \geq 0$ ;

$$2. \int_0^1 d\chi(\alpha) = 1.$$

Mixed CDD, is defined by

$$\Delta_{\chi}^+(w) = \int_0^1 \Delta_{\alpha}(w) d\chi(\alpha). \quad (3.16)$$

Obviously, the mixed CDD preserves all properties of  $\Delta_{\alpha}(w)$  stated in Proposition 3.4. A fund manager can flexibly express his or her risk preferences by shaping  $\chi(\alpha)$ .

**Proposition 3.5.** *The mixed CDD can be presented in the alternative form*

$$\Delta_{\chi}^+(w) = \int_0^1 \pi_{\xi}^{-1}(\alpha) \mu(\alpha) d\alpha, \quad (3.17)$$

with “spectrum”  $\mu(\alpha)$  to be:

- (1) nonnegative on  $[0, 1]$ ;
- (2) nondecreasing on  $[0, 1]$ ;
- (3)  $\int_0^1 \mu(\alpha) d\alpha = 1$ .

The relation between  $\chi(\alpha)$  in (3.16) and  $\mu(\alpha)$  in (3.17) is

$$d\mu(\alpha) = \frac{1}{1-\alpha} d\chi(\alpha).$$

**Proof.** Expressing  $\Delta_{\alpha}(w)$  in the form of (3.13), consider

$$\begin{aligned} \Delta_{\chi}^+(w) &= \int_0^1 \left( \frac{1}{1-\alpha} \int_{\alpha}^1 \pi_{\xi}^{-1}(q) dq \right) d\chi(\alpha) \\ &= \int_0^1 \left( \frac{1}{1-\alpha} \int_0^1 \pi_{\xi}^{-1}(q) I_{\{q \geq \alpha\}} dq \right) d\chi(\alpha) \\ &= \int_0^1 \pi_{\xi}^{-1}(q) \left( \int_0^1 \frac{1}{1-\alpha} I_{\{q \geq \alpha\}} d\chi(\alpha) \right) dq \\ &= \int_0^1 \pi_{\xi}^{-1}(q) \left( \int_0^q \frac{1}{1-\alpha} d\chi(\alpha) \right) dq \\ &= \int_0^1 \pi_{\xi}^{-1}(q) \mu(q) dq, \end{aligned}$$

where  $\mu(\alpha) = \int_0^{\alpha} \frac{1}{1-q} d\chi(q)$  satisfies all properties 1–3. Indeed,  $\mu(\alpha)$  is nonnegative and nondecreasing, since  $d\mu(\alpha) = \frac{1}{1-\alpha} d\chi(\alpha) \geq 0$ . Moreover,  $\int_0^1 \mu(\alpha) d\alpha = \int_0^1 \int_0^1 \frac{1}{1-q} I_{\{\alpha \geq q\}} d\chi(q) d\alpha = 1$ . Obviously, conditions 1–3 are necessarily satisfied by function  $\mu(\alpha)$ , since they are derived from the properties of function  $\chi(\alpha)$ . However, if function  $\mu(\alpha)$  satisfies conditions 1–3 then it is sufficient for (3.17) to be *constant translating, positively homogeneous, monotonic and convex* with respect to  $\xi$ . The last fact comes from a direct verification of those properties.  $\square$

**Corollary 3.1.** *The non-decrease property of “spectrum,”  $\mu(\alpha)$ , is a necessary condition for the mixed CDD to be convex. This property has an obvious but important interpretation, namely, the greater drawdown quantile,  $\pi_\xi^{-1}$ , is, the greater penalty coefficient,  $\mu$ , should be assigned. A similar conclusion regarding risk spectrum in coherent risk measures was made by Acerbi and Tasche [1]. This conclusion is a consequence of a general coherency principle, stating: the greater risk is, the more it should be penalized [2].*

**Example 3.3.** MaxDD and AvDD are mixed CDDs with risk profiles  $\chi(\alpha) = I_{\{\alpha \geq 1\}}$  and  $\chi(\alpha) = I_{\{\alpha > 0\}}$ , respectively.

**Discrete risk profile.** An important case is when risk profile,  $\chi(\alpha)$ , is specified by the discrete set of points  $\chi_i = d\chi(\alpha_i)$ ,  $i = \overline{1, L}$ . In this case, the mixed CDD is expressed by

$$\Delta_\chi^+(w) = \sum_{i=1}^L \chi_i \Delta_{\alpha_i}(w), \quad (3.18)$$

where  $\sum_{i=1}^L \chi_i = 1$  and  $\chi_i \geq 0$ . Consequently, “spectrum” function is presented by

$$\mu(\alpha) = \sum_{i=1}^L \frac{\chi_i}{1 - \alpha_i} I_{\{\alpha \geq \alpha_i\}}. \quad (3.19)$$

**Detail.** Interchanging summation and integration operations in  $\Delta_\chi^+(w)$ , the result follows

$$\begin{aligned} \Delta_\chi^+(w) &= \sum_{i=1}^L \chi_i \Delta_{\alpha_i}(w) = \sum_{i=1}^L \frac{\chi_i}{1 - \alpha_i} \int_{\alpha_i}^1 \pi_\xi^{-1}(q) dq \\ &= \int_0^1 \left( \sum_{i=1}^L \frac{\chi_i}{1 - \alpha_i} I_{\{\alpha \geq \alpha_i\}} \right) \pi_\xi^{-1}(q) dq. \end{aligned}$$

Obviously, (3.19) is a positive nondecreasing function.

#### 4. Optimization Techniques for Conditional Drawdown Computation

This section develops optimization techniques for CDD efficient computation. Formulas (3.9) and (3.13) require to calculate the value of  $\zeta(\alpha)$  first, which doubles computational time. However, there is an optimization procedure that obtains the values of threshold  $\zeta(\alpha)$  and CDD simultaneously. This procedure is especially important in a large scale optimization.

In the case when a time series of drawdowns is given, computation of the  $\alpha$ -CDD is reduced to computation of  $\text{CV@R}_\alpha(\xi)$ .



**Theorem 4.1.** *Given a time series of instrument's drawdowns  $\xi = (\xi_1, \dots, \xi_N)$ , corresponding to time moments  $\{t_1, \dots, t_N\}$ , the CDD functional is presented by  $\text{CV@R}_\alpha(\xi)$ , which computation is reduced to the following linear programming procedure*

$$\begin{aligned} \text{CV@R}_\alpha(\xi) = \min_{y, z} \quad & y + \frac{1}{(1-\alpha)N} \sum_{k=1}^N z_k \\ \text{s.t.} \quad & z_k \geq \xi_k - y, \quad z_k \geq 0, \quad k = \overline{1, N}, \end{aligned} \quad (4.1)$$

leading to a single optimal value of  $y$  equal to  $\zeta(\alpha)$  if  $\pi_\xi(\zeta(\alpha)) > \alpha$ , and to a closed interval of optimal  $y$  with the left endpoint of  $\zeta(\alpha)$  if  $\pi_\xi(\zeta(\alpha)) = \alpha$ .

**Proof.** We introduce a piece-wise function

$$h(y) = y + \frac{1}{(1-\alpha)N} \sum_{k=1}^N [\xi_k - y]^+, \quad (4.2)$$

where  $[\xi_k - y]^+ = \max\{\xi_k - y, 0\}$ , and establish the following relation

$$\text{CV@R}_\alpha(\xi) = \min_y h(y). \quad (4.3)$$

The derivative of  $h(y)$  with respect to  $y$  is presented by

$$\begin{aligned} \frac{d}{dy} h(y) &= 1 - \frac{1}{(1-\alpha)N} \sum_{k=1}^N I_{\{y < \xi_k\}} = 1 - \frac{1}{(1-\alpha)N} \sum_{k=1}^N (1 - I_{\{\xi_k \leq y\}}) \\ &= \frac{\pi_\xi(y) - \alpha}{1 - \alpha}. \end{aligned} \quad (4.4)$$

Note that  $\frac{d}{dy} h(y)$  is continuous for all values of  $y$ , except the set of points  $y = \{\xi_k | k = \overline{1, N}\}$ . The necessary condition for function  $h(y)$  to attain an extremum is

$$\frac{d^-}{dy} h(y) \leq 0 \leq \frac{d^+}{dy} h(y), \quad (4.5)$$

where  $\frac{d^-}{dy} h(y) = \frac{1}{(1-\alpha)}(\pi_\xi(y-0) - \alpha)$  and  $\frac{d^+}{dy} h(y) = \frac{1}{(1-\alpha)}(\pi_\xi(y+0) - \alpha)$  are left and right derivatives, respectively, which coincide with each other for all  $y$  except  $y = \{\xi_k | k = \overline{1, N}\}$ . According to (4.4) and (4.5), an optimal value  $y^*$  should satisfy inequalities

$$\pi_\xi(y^* - 0) \leq \alpha \leq \pi_\xi(y^* + 0),$$

which have a unique solution  $y^* = \zeta(\alpha)$  if  $\pi_\xi(\zeta(\alpha)) > \alpha$  (see Remark 3.2), i.e., if  $y^* \neq \{\xi_k | k = \overline{1, N}\}$ . However, if  $\pi_\xi(\zeta(\alpha)) = \alpha$ , then there is a closed interval of optimal values  $y^*$ , with the left endpoint of  $\zeta(\alpha)$ , namely,  $y^* \in [\zeta(\alpha), \zeta(\alpha + 0)]$ , where  $\pi_\xi(\zeta(\alpha + 0)) > \alpha$ . Hence, two cases are considered:

(a)  $y^* = \zeta(\alpha)$  if  $\pi_\xi(\zeta(\alpha)) > \alpha$ ;

(b)  $y^* \in [\zeta(\alpha), \zeta(\alpha + 0)]$  if  $\pi_\xi(\zeta(\alpha)) = \alpha$ .

In both cases, equality  $[\xi_k - y^*]^+ = (\xi_k - y^*)I_{\{\xi_k \geq y^*\}} = (\xi_k - y^*)I_{\{\xi_k > \zeta(\alpha)\}}$  holds with respect to all  $\xi_k$ ,  $k = \overline{1, N}$ , for any fixed  $y^*$ . Thus, based on this fact, we obtain

$$\begin{aligned} \min_y h(y) &= h(y^*) = y^* + \frac{1}{(1-\alpha)N} \sum_{k=1}^N [\xi_k - y^*]^+ \\ &= \frac{1}{1-\alpha} \left( 1 - \alpha - \frac{1}{N} \sum_{k=1}^N I_{\{\xi_k > \zeta(\alpha)\}} \right) y^* + \frac{1}{(1-\alpha)N} \sum_{k=1}^N \xi_k I_{\{\xi_k > \zeta(\alpha)\}} \\ &= \frac{(\pi_\xi(\zeta(\alpha)) - \alpha)}{1-\alpha} y^* + \frac{1}{(1-\alpha)N} \sum_{t_k \in \Xi_\alpha} \xi_k, \end{aligned}$$

where  $\frac{(\pi_\xi(\zeta(\alpha)) - \alpha)}{1-\alpha} y^* = \frac{(\pi_\xi(\zeta(\alpha)) - \alpha)}{1-\alpha} \zeta(\alpha)$  in the case of (a), and  $\frac{(\pi_\xi(\zeta(\alpha)) - \alpha)}{1-\alpha} y^* = 0$  in the case of (b). Consequently,  $\min_y h(y)$  coincides with the definition of the CDD.

Since expression  $\sum_{k=1}^N [\xi_k - y]^+$  is minimized, it can equivalently be presented by the sum of nonnegative auxiliary variables  $z_k \geq 0$ ,  $k = \overline{1, N}$ , satisfying additional constraints  $z_k \geq \xi_k - y$ ,  $k = \overline{1, N}$ .  $\square$

**Corollary 4.1.** *CV@R $_\alpha$ ( $\xi$ ) is an optimal value for the objective function of the following knapsack problem*

$$\begin{aligned} \text{CV@R}_\alpha(\xi) &= \max_q \sum_{k=1}^N \xi_k q_k \\ \text{s.t.} \quad &\sum_{k=1}^N q_k = 1, \quad 0 \leq q_k \leq \frac{1}{(1-\alpha)N}, \quad k = \overline{1, N}. \end{aligned} \tag{4.6}$$

*The value of CV@R $_\alpha$ ( $\xi$ ) can be found in  $O(n \log_2 n)$  time.*

**Proof.** It is enough to observe that knapsack problem (4.6) is *dual* to linear programming problem (4.1). Based on duality theory, optimal values of the objective functions in (4.1) and (4.6) should coincide. Problem (4.6) can be solved by the standard *greedy* algorithm in  $O(n \log_2 n)$  time. The algorithm sorts items according to their “costs”  $\{\xi_k | k = \overline{1, N}\}$ . Let  $\lfloor a \rfloor$  denote the integer part of real number  $a$ . Obviously,  $q$ -variables, corresponding to the largest  $\lfloor (1-\alpha)N \rfloor$  “costs,” have optimal values equal to  $\frac{1}{(1-\alpha)N}$ , and the  $q$ -variable, corresponding to the  $(\lfloor (1-\alpha)N \rfloor + 1)$ th “cost” in the sorted order, has optimal value equal to  $1 - \frac{\lfloor (1-\alpha)N \rfloor}{(1-\alpha)N}$ . The rest of  $q$ -variables equal 0. In this case, the complexity of the algorithm is mainly determined by a sorting procedure, which, in this case, requires at least  $O(n \log_2 n)$  operations.  $\square$

Formulation (4.6) is closely related to the presentation of CV@R based on the concept of a *risk envelope*, which is a closed, convex set of probabilities containing 1. Risk envelope theory was developed by Rockafellar *et al.* [22–24].

Suppose, a sample path of instrument's rates of return  $(r_1, \dots, r_N)$ , corresponding to time moments  $\{t_1, \dots, t_N\}$ , is given. In this case, un compounded cumulative instrument's rate of return at  $t_k$  is  $w_k = \sum_{l=1}^k r_l$ , and the CDD is presented in the form of  $\Delta_\alpha(w)$ .

**Proposition 4.1.** *Given a sample path of instrument's rates of return  $(r_1, \dots, r_N)$ , the CDD functional,  $\Delta_\alpha(w)$ , is computed by the following optimization procedure*

$$\begin{aligned} \Delta_\alpha(w) = \min_{u, y, z} \quad & y + \frac{1}{(1-\alpha)N} \sum_{k=1}^N z_k \\ \text{s.t.} \quad & z_k \geq u_k - y, \\ & u_k \geq u_{k-1} - r_k, \quad u_0 = 0, \\ & z_k \geq 0, \quad u_k \geq 0, \quad k = \overline{1, N}, \end{aligned} \quad (4.7)$$

which leads to a single optimal value of  $y$  equal to  $\zeta(\alpha)$  if  $\pi_\xi(\zeta(\alpha)) > \alpha$ , and to a closed interval of optimal  $y$  with the left endpoint of  $\zeta(\alpha)$  if  $\pi_\xi(\zeta(\alpha)) = \alpha$ .

**Proof.** By virtue of relation  $\Delta_\alpha(w) = \text{CV@R}_\alpha(\mathcal{AD}(w)) = \text{CV@R}_\alpha(\xi)$ , optimization problem (4.7) is a direct consequence of (4.1). Using recursive formula  $\xi_k = [\xi_{k-1} - r_k]^+$ , constraint  $z_k \geq \xi_k - y$  in (4.1) is reduced to  $z_k \geq u_k - y$ , where nonnegative auxiliary variables  $u_k$  satisfy additional constraints  $u_k \geq \xi_{k-1} - r_k$ ,  $k = \overline{1, N}$ , with  $u_0 = 0$ .  $\square$

**Corollary 4.2.** *Given a sample path of instrument's rates of return  $(r_1, \dots, r_N)$ , the CDD functional,  $\Delta_\alpha(w)$ , is computed by the following optimization procedure*

$$\begin{aligned} \Delta_\alpha(w) = \max_{q, \eta} \quad & - \sum_{k=1}^N r_k \eta_k \\ \text{s.t.} \quad & \sum_{k=1}^N q_k = 1, \quad \eta_k - \eta_{k+1} \leq q_k \leq \frac{1}{(1-\alpha)N}, \\ & q_k \geq 0, \quad \eta_k \geq 0, \quad \eta_{N+1} = 0, \quad k = \overline{1, N}. \end{aligned} \quad (4.8)$$

**Proof.** Problem (4.8) is *dual* to linear programming program (4.7).  $\square$

Theorem 4.1 and all its corollaries can be easily generalized to the case of mixed CDD.

**Proposition 4.2.** *Given a sample path of instrument's rates of return  $\{r_k | k = \overline{1, N}\}$  and discrete risk profile  $\chi_i = d\chi(\alpha_i)$ ,  $i = \overline{1, L}$ , the mixed CDD,  $\Delta_\chi^+(w)$ , is*

computed by

$$\begin{aligned} \Delta_{\chi}^+(w) = \min_{u,y,z} & \sum_{i=1}^L \chi_i \left( y_i + \frac{1}{(1-\alpha)N} \sum_{k=1}^N z_{ik} \right) \\ \text{s.t.} & z_{ik} \geq u_k - y_i, \\ & u_k \geq u_{k-1} - r_k, \quad u_0 = 0, \\ & z_{ik} \geq 0, \quad u_k \geq 0, \quad i = \overline{1, L}, \quad k = \overline{1, N}. \end{aligned} \quad (4.9)$$

**Proof.** Formulation (4.9) is a direct consequence of mixed CDD definition (3.18) and optimization problem (4.7). Notice that auxiliary variables  $u_k$  do not have index  $i$ , since they determine the drawdown sequence same for all  $\alpha_i$ .  $\square$

## 5. Multi-Scenario Conditional Value-at-Risk and Drawdown Measure

This section presents concept of the ‘‘Multi-scenario’’ CV@R and drawdown measure, which, in fact, are the CV@R and CDD defined in the case of several sample paths for uncompounded cumulative portfolio rate of return. We generalize results obtained for the CDD under assumption of a single sample path to the case of several sample paths.

Let  $\Omega$  denote a discrete set of random events, i.e.,  $\Omega = \{\omega_j | j = \overline{1, K}\}$ , and let  $p_j$  be the probability of event  $\omega_j$  ( $\forall j : p_j \geq 0$ , and  $\sum_{j=1}^K p_j = 1$ ). Suppose  $r_j(t_k) = (r_{1j}(t_k), r_{2j}(t_k), \dots, r_{mj}(t_k))$ ,  $k = \overline{1, N}$ , is the  $j$ th sample path for the random vector of risky assets’ rates of return, corresponding to random event  $\omega_j \in \Omega$  and time interval  $[0, T]$  presented by the discrete set of time moments  $\{t_0 = 0, t_1, t_2, \dots, t_N = T\}$ . Consequently, the  $j$ th sample path for the rate of return and uncompounded cumulative rate of return of a portfolio with capital weights  $x(t_k) = (x_0(t_k), x_1(t_k), x_2(t_k), \dots, x_m(t_k))$  are defined, respectively,

$$r_{jk}^{(p)}(x(t_k)) = r_j(t_k) \cdot x(t_k) = \sum_{i=1}^m r_{ij}(t_k) x_i(t_k), \quad (5.1)$$

$$w_{jk}(x(t_k)) = \begin{cases} 0, & k = 0, \\ \sum_{l=1}^k r_{jl}^{(p)}(x(t_l)), & k = \overline{1, N}. \end{cases} \quad (5.2)$$

To simplify notations, we use  $w_{jk}$  instead of  $w_{jk}(x(t_k))$  implying that  $w_{jk}$  is always a function of  $x(t_k)$ . In a multi-scenario case,  $w$  denotes matrix  $\{w_{jk}\}$ ,  $j = \overline{1, K}$ ,  $k = \overline{0, N}$ .

### 5.1. Multi-scenario Conditional Value-at-Risk

**Definition 5.1.** In a multi-scenario case, the  $\mathcal{AD}(w)$  is a matrix-variable functional defined on  $\Omega \times [0, T]$

$$\mathcal{AD}(w) = \xi = \{\xi_{jk}\}, \quad \xi_{jk} = \max_{0 \leq l \leq k} \{w_{jl}\} - w_{jk}, \quad j = \overline{1, K}, \quad k = \overline{1, N}. \quad (5.3)$$

All  $\mathcal{AD}$  properties stated in Proposition 3.1 hold in a multi-scenario case. Indeed, based on (5.3), properties 1–4 in Proposition 3.1 can be verified directly. Matrix  $\mathcal{AD}(w)$  is interpreted to be drawdown surface  $\xi_{jk}$ ,  $(\omega_j, t_k) \in \Omega \times [0, T]$ .

**Definition 5.2.** Similar to definitions of MaxDD and AvDD in single scenario case, MaxDD and AvDD are defined on  $\Omega \times [0, T]$ , respectively,

$$\text{MaxDD}(w) = \max_{1 \leq j \leq K, 1 \leq k \leq N} \{\xi_{jk}\}, \quad (5.4)$$

$$\text{AvDD}(w) = \frac{1}{N} \sum_{k=1}^N \sum_{j=1}^K p_j \xi_{jk}. \quad (5.5)$$

**Definition 5.3.** Indicator function for drawdown surface, its inverse function and threshold plane,  $\zeta(\alpha)$ , are defined, respectively,

$$\pi_\xi(s) = \frac{1}{N} \sum_{k=1}^N \sum_{j=1}^K p_j I_{\{\xi_{jk} \leq s\}}, \quad (5.6)$$

$$\pi_\xi^{-1}(\alpha) = \begin{cases} \inf\{s | \pi_\xi(s) \geq \alpha\}, & \alpha \in (0, 1], \\ 0, & \alpha = 0, \end{cases} \quad (5.7)$$

$$\zeta(\alpha) = \pi_\xi^{-1}(\alpha). \quad (5.8)$$

Figure 5 illustrates drawdown surface  $\xi_{jk}$  and threshold plane  $\zeta(\alpha)$ .

**Definition 5.4.** Multi-scenario CV@R may be defined similar to a single period CV@R, namely,

$$\text{CV@R}(\xi) = \left( \frac{\pi_\xi(\zeta(\alpha)) - \alpha}{1 - \alpha} \right) \zeta(\alpha) + \frac{1}{(1 - \alpha)N} \sum_{\xi_{jk} \in \Xi_\alpha} p_j \xi_{jk}, \quad (5.9)$$

where  $\Xi_\alpha = \{\xi_{jk} | \xi_{jk} > \zeta(\alpha), k = \overline{1, N}\}$ .

**Proposition 5.1.** Multi-scenario CV@R, given by (5.9), can be presented in the alternative form

$$\text{CV@R}(\xi) = \frac{1}{1 - \alpha} \int_\alpha^1 \pi_\xi^{-1}(q) dq, \quad (5.10)$$

where  $\pi_\xi^{-1}(q)$  is the inverse function given by (5.7).

**Proof.** Similar to the proof of Theorem 3.1. □

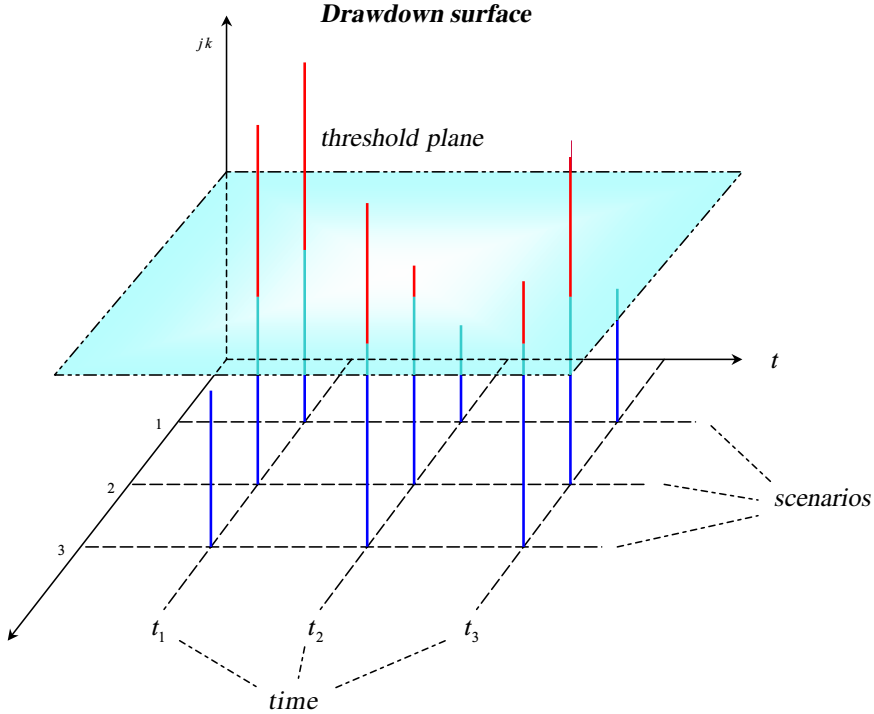


Fig. 5. Drawdown surface and threshold plane.

**Remark 5.1.** Let  $X$  be an arbitrary random variable. Suppose we are given  $K$  sample paths  $X(t_k, \omega_j)$ ,  $k = \overline{1, N}$ , corresponding to random events  $\omega_j \in \Omega$  with probabilities  $p_j$  such that  $\sum_{j=1}^K p_j = 1$ . Defining an indicator function for  $X$  to be  $\pi_X(s) = \frac{1}{N} \sum_{k=1}^N \sum_{j=1}^K p_j I_{\{X(t_k, \omega_j) \leq s\}}$  (where the inverse function  $\pi_X^{-1}$  is defined similar to (5.7)), *multi-scenario* CV@R may be determined similar to a *single period* CV@R, namely,  $CV@R_\alpha(X) = \frac{1}{1-\alpha} \int_\alpha^1 \pi_X^{-1}(q) dq$ .

**5.2. Drawdown measure**

In a multi-scenario case, CDD with tolerance level  $\alpha$  is interpreted as

- The average of the worst  $(1 - \alpha) * 100\%$  drawdowns on drawdown surface, if the worst  $(1 - \alpha) * 100\%$  drawdowns can be counted precisely.
- The linear combination of  $\zeta(\alpha)$  and the average of the drawdowns strictly exceeding threshold plane  $\zeta(\alpha)$ , if we are unable to precisely count of  $(1 - \alpha) * 100\%$  drawdowns.

A strict mathematical definition of the drawdown measure is given below.

**Definition 5.5.** In a multi-scenario case, the CDD, with tolerance level  $\alpha \in [0, 1]$ , is the multi-scenario  $\text{CV@R}_\alpha$  applied to drawdown surface,  $\mathcal{AD}(w)$ ,

$$\Delta_\alpha(w) = \text{CV@R}_\alpha(\mathcal{AD}(w)), \quad (5.11)$$

and drawdown measure is the mixed CDD with risk profile  $\chi(\alpha)$

$$\Delta_\chi^+(w) = \int_0^1 \Delta_\alpha(w) d\chi(\alpha), \quad (5.12)$$

where  $\Delta_\alpha(w)$  is given by (5.11).

**Proposition 5.2.** *Defining matrix operations:  $w + \text{const} = \{w_{jk} + \text{const}\}$  and  $\lambda w = \{\lambda w_{jk}\}$ , drawdown measure  $\Delta_\chi^+(w)$  satisfies the following properties*

- (1) *Nonnegativity:*  $\Delta_\chi^+(w) \geq 0, \forall \alpha \in [0, 1]$ .
- (2) *Insensitivity to constant shift:*  $\Delta_\chi^+(w + \text{const}) = \Delta_\chi^+(w), \forall \alpha \in [0, 1]$ .
- (3) *Positive homogeneity:*  $\Delta_\chi^+(\lambda w) = \lambda \Delta_\chi^+(w), \forall \lambda \geq 0$  and  $\forall \alpha \in [0, 1]$ .
- (4) *Convexity:* if  $w_\lambda = \lambda w_1 + (1 - \lambda)w_2$  is a linear combination of any  $w_1$  and  $w_2$  with  $\lambda \in [0, 1]$ , then  $\Delta_\chi^+(w_\lambda) \leq \lambda \Delta_\chi^+(w_1) + (1 - \lambda) \Delta_\chi^+(w_2)$ .

**Proof.** Properties 1–4 are direct generalization of CDD properties stated in Proposition 3.4.  $\square$

**Proposition 5.3.** *In the case of discrete risk profile, drawdown measure is computed by*

$$\begin{aligned} \Delta_\chi^+(w) = \min_{u, y, z} \quad & \sum_{i=1}^L \chi_i \left( y_i + \frac{1}{(1 - \alpha_i)N} \sum_{k=1}^N \sum_{j=1}^K p_j z_{ijk} \right) \\ \text{s.t.} \quad & z_{ijk} \geq u_{jk} - y_i, \\ & u_{jk} \geq u_{j(k-1)} - r_{jk}^{(p)}, \\ & u_{jk} \geq 0, \quad u_{j0} = 0, \quad z_{ijk} \geq 0, \\ & i = \overline{1, L}, j = \overline{1, K}, k = \overline{1, N}. \end{aligned} \quad (5.13)$$

**Proof.** Introducing intermediate optimization problems

$$\begin{aligned} \sum_{i=1}^L \chi_i \text{CV@R}_{\alpha_i}(\xi) &= \min_{y_i} \sum_{i=1}^L \chi_i \left( y_i + \frac{1}{(1 - \alpha_i)N} \sum_{k=1}^N \sum_{j=1}^K p_j [\xi_{jk} - y_i]^+ \right), \\ \sum_{i=1}^L \chi_i \text{CV@R}_{\alpha_i}(\xi) &= \min_{y_i, z_{ijk}} \sum_{i=1}^L \chi_i \left( y_i + \frac{1}{(1 - \alpha_i)N} \sum_{k=1}^N \sum_{j=1}^K p_j z_{ijk} \right) \\ \text{s.t.} \quad & z_{ijk} \geq \xi_{jk} - y_i, z_{ijk} \geq 0, \\ & i = \overline{1, L}, j = \overline{1, K}, k = \overline{1, N}, \end{aligned}$$

the proof is conducted similar to the proof of Theorem 4.1.  $\square$

## 6. Portfolio Optimization with Drawdown Measure

This section formulates a portfolio optimization problem with drawdown risk measure and suggests efficient optimization techniques for its solving. Optimal asset allocation considers:

- Generation of sample paths for the assets' rates of return.
- Uncompounded cumulative portfolio rate of return rather than compounded one.

In this case, optimal asset allocation maximizes the expected value of uncompounded cumulative portfolio rate of return at the final time moment  $t_N = T$  subject to a constraint on drawdown measure

$$\begin{aligned} \max_{x \in X} \quad & E_\omega(w(T, \omega, x)) = \sum_{j=1}^K p_j w_{jN}(x) \\ \text{s.t.} \quad & \Delta_X^+(w(x)) \leq \gamma, \end{aligned} \quad (6.1)$$

where  $X$  is the set of linear "technological" constraints and  $\gamma \in [0, 1]$  is a proportion of the initial capital allowed to loose.

In contrast to Grossman and Zhou [12] and Cvitanic and Karatzas [7], who considered vector of portfolio weights to be a function of time within  $[0, T]$ , we assume portfolio weights  $x(t_k)$  to be static for all  $t_k$ ,  $k = \overline{0, N}$ . This special strategy can be achieved by portfolio rebalancing at every  $t_k$ ,  $k = \overline{0, N}$ . Justification of this assumption depends on a particular case study. Based on the assumption made, uncompounded cumulative portfolio rate of return  $w$  is rewritten

$$w_{jk}(x) = \sum_{l=1}^k r_{jl}^{(p)}(x) = \sum_{i=1}^m \sum_{l=1}^k r_{ij}(t_l) x_i. \quad (6.2)$$

### 6.1. Reduction to linear programming problem

**Theorem 6.1.** *Problem (6.1) is reduced to linear programming (LP) problem*

$$\begin{aligned} \max_{u, x \in X, y, z} \quad & \sum_{j=1}^K p_j w_{jN}(x) \\ \text{s.t.} \quad & \sum_{i=1}^L \chi_i \left( y_i + \frac{1}{(1 - \alpha_i)N} \sum_{k=1}^N \sum_{j=1}^K p_j z_{ijk} \right) \leq \gamma, \\ & z_{ijk} \geq u_{jk} - y_i, \\ & u_{jk} \geq u_{j(k-1)} - r_{jk}^{(p)}(x), \\ & u_{jk} \geq 0, \quad u_{j0} = 0, \quad z_{ijk} \geq 0, \\ & i = \overline{1, L}, j = \overline{1, K}, k = \overline{1, N}, \end{aligned} \quad (6.3)$$

where  $u_{jk}$ ,  $y_i$  and  $z_{ijk}$  are auxiliary variables.



**Proof.** Consider piece-wise function  $H(x, y)$

$$H(x, y) = \sum_{i=1}^L \chi_i \left( y_i + \frac{1}{(1 - \alpha_i)N} \sum_{k=1}^N \sum_{j=1}^K p_j [\xi_{jk}(x) - y_i]^+ \right). \quad (6.4)$$

According to Proposition 5.3, drawdown measure may be presented by

$$\Delta_{\chi}^+(w(x)) = \sum_{i=1}^L \chi_i \text{CV@R}_{\alpha_i}(\xi(x)) = \min_y H(x, y). \quad (6.5)$$

Consequently, problem (6.1) is reduced to

$$\begin{aligned} \max_{x \in X} \quad & \sum_{j=1}^K p_j w_{jN}(x) \\ \text{s.t.} \quad & \min_y H(x, y) \leq \gamma, \end{aligned} \quad (6.6)$$

The key point of the proof is to show that minimum in the constraint of (6.6) may be relaxed, i.e., to show that problem (6.6) is equivalent to

$$\begin{aligned} \max_{x \in X, y} \quad & \sum_{j=1}^K p_j w_{jN}(x) \\ \text{s.t.} \quad & H(x, y) \leq \gamma, \end{aligned} \quad (6.7)$$

The proof of this fact is conducted by relaxing constraint  $\min_y H(x, y) \leq \gamma$  in (6.6), namely, problem (6.6) is equivalently rewritten

$$\begin{aligned} \min_{\lambda \geq 0} \max_{x \in X} \quad & \left( \sum_{j=1}^K p_j w_{jN}(x) + \lambda(\gamma - \min_y H(x, y)) \right), \\ \min_{\lambda \geq 0} \max_{x \in X, y} \quad & \left( \sum_{j=1}^K p_j w_{jN}(x) + \lambda(\gamma - H(x, y)) \right). \end{aligned} \quad (6.8)$$

However, problem (6.8) is the *Lagrange relaxation* of (6.7). Hence, (6.7) is equivalent to (6.6). According to Theorem 6.1 and Proposition 5.3, LP (6.3) is a direct consequence of (6.7).  $\square$

**Corollary 6.1.** *In the cases of MaxDD( $w$ ) and AvDD( $w$ ), corresponding to the mixed CDD with risk profiles of  $\chi(\alpha) = I_{\{\alpha > 0\}}$  and  $\chi(\alpha) = I_{\{\alpha \geq 1\}}$ , LP (6.3) is simplified, respectively,*

$$\begin{aligned} \max_{u, x \in X} \quad & \sum_{j=1}^K p_j w_{jN}(x) \\ \text{s.t.} \quad & u_{jk} \geq u_{j(k-1)} - r_{jk}^{(p)}(x), \\ & \gamma \geq u_{jk} \geq 0, u_{j0} = 0, \\ & j = \overline{1, K}, k = \overline{1, N}, \end{aligned} \quad (6.9)$$

$$\begin{aligned}
& \max_{u, x \in X} \sum_{j=1}^K p_j w_{jN}(x) \\
& \text{s.t.} \quad \frac{1}{N} \sum_{k=1}^N \sum_{j=1}^K p_j u_{jk} \leq \gamma, \\
& \quad u_{jk} \geq u_{j(k-1)} - r_{jk}^{(p)}(x), \\
& \quad u_{jk} \geq 0, u_{j0} = 0, \\
& \quad j = \overline{1, K}, k = \overline{1, N}.
\end{aligned} \tag{6.10}$$

## 6.2. Efficient frontier

Efficient frontier is a central concept in Risk Management methodology. Suppose for every value of  $\gamma$  and risk profile  $\chi$ ,  $x_\chi^*(\gamma)$  is an optimal solution to (6.3). In this case, efficient frontier is a curve expressing dependence of optimal portfolio expected reward  $\sum_{j=1}^K p_j w_{jN}(x_\chi^*(\gamma))$  on portfolio risk  $\gamma$ .

**Proposition 6.1.** *Efficient frontier  $(\gamma, \sum_{j=1}^K p_j w_{jN}(x_\chi^*(\gamma)))$  is a concave curve.*

**Proof.** Denoting  $g(x) = \sum_{j=1}^K p_j w_{jN}(x)$ , we show that for any  $\gamma_{1,2} \in [0, 1]$  and  $\tau \in [0, 1]$

$$g(x_\chi^*(\tau\gamma_1 + (1-\tau)\gamma_2)) \geq \tau g(x_\chi^*(\gamma_1)) + (1-\tau)g(x_\chi^*(\gamma_2)).$$

According to the proof of Theorem 6.1, we have

$$\begin{aligned}
g(x_\chi^*(\gamma)) &= \max_{x \in X, y} g(x) \\
& \text{s.t.} \quad H(x, y) \leq \gamma,
\end{aligned}$$

and using notation  $G_\lambda(x, y) = g(x) - \lambda H(x, y)$ , we obtain

$$g(x_\chi^*(\gamma)) = \min_{\lambda \geq 0} \max_{x \in X, y} (G_\lambda(x, y) + \lambda\gamma) = \min_{\lambda \geq 0} (G_\lambda(x(\lambda), y(\lambda)) + \lambda\gamma).$$

Since expression  $G_\lambda(x(\lambda), y(\lambda)) + \lambda\gamma$  is linear with respect to  $\gamma$ ,  $\min_{\lambda \geq 0} (G_\lambda(x(\lambda), y(\lambda)) + \lambda\gamma)$  is a concave function of  $\gamma$ . Indeed,

$$\begin{aligned}
& \min_{\lambda \geq 0} (G_\lambda(x(\lambda), y(\lambda)) + \lambda(\tau\gamma_1 + (1-\tau)\gamma_2)) \\
&= \min_{\lambda \geq 0} (\tau(G_\lambda(x(\lambda), y(\lambda)) + \lambda\gamma_1) + (1-\tau)(G_\lambda(x(\lambda), y(\lambda)) + \lambda\gamma_2)) \\
&\geq \tau \min_{\lambda \geq 0} (G_\lambda(x(\lambda), y(\lambda)) + \lambda\gamma_1) + (1-\tau) \min_{\lambda \geq 0} (G_\lambda(x(\lambda), y(\lambda)) + \lambda\gamma_2).
\end{aligned}$$

This fact proves the proposition.  $\square$

*Risk-adjusted return* is an important characteristic for choosing an optimal portfolio on an efficient frontier that evaluates the ratio of the portfolio reward to the portfolio risk

$$\rho_\chi(\gamma) = \gamma^{-1} \sum_{j=1}^K p_j w_{jN}(x_\chi^*(\gamma)). \quad (6.11)$$

A fund manager is interested in such a value of  $\gamma \in [0, 1]$ , for which the risk-adjusted return  $\rho_\chi(\gamma)$  is maximal. It is interpreted to be *the best balance between the risk accepted and the rate of return achieved*. According to Proposition 6.1,  $\sum_{j=1}^K p_j w_{jN}(x_\chi^*(\gamma))$  is concave, hence, when this function achieves its maximum at  $\gamma > 0$ , ratio  $\rho_\chi(\gamma)$  has a finite global maximum. Although  $\rho_\chi(\gamma)$  is a nonlinear function with respect to  $\gamma$ , a problem for finding  $\rho_\chi(\gamma)$  maximum and corresponding optimal  $\gamma$  is reduced to an LP.

**Proposition 6.2.** *The optimization problem  $\max_{\gamma \in [0,1]} \rho_\chi(\gamma)$  is reduced to LP*

$$\begin{aligned} \max_{\tilde{u}, v, \tilde{x} \in \tilde{X}, \tilde{y}, \tilde{z}} \quad & \sum_{j=1}^K p_j w_{iN}(\tilde{x}) \\ \text{s.t.} \quad & \sum_{i=1}^L \chi_i \left( \tilde{y}_i + \frac{1}{(1 - \alpha_i)N} \sum_{k=1}^N \sum_{j=1}^K p_j \tilde{z}_{ijk} \right) \leq 1, \\ & \tilde{z}_{ijk} \geq \tilde{u}_{jk} - \tilde{y}_i, \\ & \tilde{u}_{jk} \geq \tilde{u}_{j(k-1)} - r_{jk}^{(p)}(\tilde{x}), \\ & \tilde{u}_{jk} \geq 0, \tilde{u}_{j0} = 0, \tilde{z}_{ijk} \geq 0, \\ & i = \overline{1, L}, j = \overline{1, K}, k = \overline{1, N}. \end{aligned} \quad (6.12)$$

If  $\tilde{x}^*$  is an optimal solution to (6.12) then  $\rho_\chi(\gamma^*) = \max_{\gamma \in [0,1]} \rho_\chi(\gamma) = \sum_{j=1}^K p_j w_{jN}(\tilde{x}^*)$ , with optimal value  $\gamma^* = 1 / \sum_{l=0}^m \tilde{x}_l^*$  and corresponding optimal portfolio  $x_l^* = \tilde{x}_l^* \gamma^*$ ,  $l = \overline{0, m}$ .

**Proof.** Since

$$\max_{\gamma \in [0,1]} \rho_\chi(\gamma) = \max_{\gamma \in [0,1]} \gamma^{-1} \sum_{j=1}^K p_j w_{jN}(x_\chi^*(\gamma)) = \max_{\gamma \in [0,1]} \max_{x \in X_\chi} \gamma^{-1} \sum_{j=1}^K p_j w_{jN}(x),$$

where  $X_\chi$  is the set of constraints in problem (6.3), the problem of  $\max_{\gamma \in [0,1]} \max_{x \in X_\chi} \gamma^{-1} \sum_{j=1}^K p_j w_{jN}(x)$  is reduced to LP (6.12) by changing variables  $\tilde{x}_l = x_l / \gamma$ ,  $\tilde{y}_i = y_i / \gamma$ ,  $\tilde{u}_{kj} = u_{kj} / \gamma$ ,  $\tilde{z}_{ijk} = z_{ijk} / \gamma$ ,  $l = \overline{0, m}$ ,  $i = \overline{1, L}$ ,  $j = \overline{1, K}$ ,  $k = \overline{1, N}$ . Set  $\tilde{X}$  may include additional variable  $v = 1/\gamma$ . For instance, a box constraint  $x_{\min} \leq x_l \leq x_{\max}$  from the set  $X$  is transformed to  $x_{\min} v \leq \tilde{x}_l \leq x_{\max} v$ , which is an element of  $\tilde{X}$ .  $\square$

## 7. Drawdown Measure in Real-life Portfolio Optimization

### 7.1. *Static asset allocation*

This section formulates and solves a real-life portfolio optimization problem with a static set of weights using drawdown measure. A problem of dynamic weight allocation when asset (or a set of assets) is log-Brownian under a constraint on the worst equity drawdown was considered in several papers. First, a 1-dimensional case was solved by Grossman and Zhou [12] as a mathematical programming problem. Then, the problem was generalized to a multi-dimensional case by Cvitanic and Karatzas [7].

In contrast to Grossman and Zhou [12] and Cvitanic and Karatzas [7], we are interested in a constant set of weights that optimizes a certain portfolio of assets, which are not assumed to have a log-brownian dynamics. This problem is stimulated by several important practical financial applications, particularly related to the so-called hedge-fund business.

A Commodity Trading Advisor (CTA) company is a hedge fund that normally trades several (sometimes, more than a 100) futures markets simultaneously using some mathematical strategies that it believes have certain edge. Such a company manages substantial assets as a part of all hedge funds, by some estimates, close to \$100 BN. Most of the CTA community trades the, so-called, long-term trend-following systems, but there are now multiple examples of short-term mean-reverting trading systems as well. These systems may be viewed as some functions of the individual futures market price realized prior to the present time. These strategies normally have a substantial smoothing-out effect on the futures prices and have close to stationary properties. Every CTA, then, has to allocate a certain portion of overall risk (or overall capital that it manages) to each and every “market”. Due to a substantial level of stationarity of the strategies, each CTA calculates the weights according to a certain internal proprietary weight allocation procedure. Normally, this set remains fixed and does not change unless a certain market gets added or removed from the set, which normally happens when a new system is introduced, when a certain market disappears (like Deutsche Mark or French Franc in 1999), or a new market is being added. A standard practice in the CTA community is to use some version of the classical Markowitz mean-variance approach.

Another important example of static asset allocation comes from the so-called, Fund of Fund (FoF) business. In the recent several years this sector of hedge funds has experienced a substantial growth. A typical FoF manager gives allocations of its clients’ capital to a set of pre-selected managers, normally between five and 25. It does so fairly infrequently, because of liquidity constraints imposed by managers themselves, but this is not the only reason. FoF views equity return streams as fairly stationary time series with some attractive return, risk, and correlation properties, which need some time to present themselves. Unless some unexpected event happens, the allocations are given for a substantial period of time, on average of

two years or more. A group of analysts in a typical FoF is responsible for finding a constant set of weights, which makes a total portfolio of the FoF to be attractive to its clients.

Both of these typical cases are faced with a problem of finding a constant set of weights, which optimize their portfolios in a certain sense. The practical goal of this paper is to facilitate this process with a clear and statistically sound algorithm, which utilizes a newly designed set of risk-measures based on a notion of equity drawdown.

Despite their known potential drawbacks, it is a well-accepted and, moreover, recommended practice [27], is to study historical back-tested strategy results of a hedge fund and, based on these results, obtain an estimate of the inherent risk using some risk measures. The only popular quantitative risk measure is VaR [27]. Various insufficiencies of the VaR measure are also widely known. We believe that the results developed in our study would facilitate understanding of how this can be achieved.

## **7.2. Historical data and scenario generation**

Even though scientists and engineers used certain simple versions of re-sampling procedures since 1930s, it was namely B. Efron [9] who unified the disconnected ideas; and re-sampling emerged as a robust method of estimating confidence intervals of some measurable functions over a statistical sample of data. Method is particularly useful for the time series where obtaining other realizations of the data may be difficult or even impossible.

Bootstrap is a form of re-sampling the original data set bootstrap, which “re-samples with replacement.” Sometimes, the simplest version of it is called “non-parametric bootstrap.” The method originally was applied to some sociological and biological applications, staying in the shade for statistical, engineering and financial applications up until the 1990s. Due to their intrinsic “one realization only”-nature, the financial time series could be one of the best applications for re-sampling methods.

Within financial applications, a strong particular interest in obtaining estimates of certain measurable quantities (such as rate of return, or standard deviation), comes from the development of trading systems. It is well known, that a problem of actual using over-fitted trading systems can possibly lead to substantial financial losses. Therefore, it is hard to underestimate the importance of a problem of discovering how over-fitted a particular trading system is. Among a few examples, one can mention a single asset trading system, for example, a system which trades a back-adjusted continuous 10-year US Government Note futures contract, or, a more general portfolio optimization problem such as allocation of weights between several assets in a portfolio subject to certain constraints.

Our study considers a particular example of optimal portfolio-allocation problem. This example could be very relevant for global CTA managers, who apply

certain trading systems (very frequently, long-term trend-following systems) across a wide set of global futures markets attempting to take advantage of price movements occurring in these markets. Normally, after they are content with their trading system, they have to make a decision of allocating their portfolio risk between various markets.

In this example, we are given a set of sample paths of certain futures trading systems (in this particular case, some long-term trend-following system) as applied to a set of 32 different global futures markets. The system includes long, short or flat markets, and always trades the same number of contracts with the average trade length from one to two months.

Here is a list of the markets with their corresponding exchanges that were traded by the system. Ticker symbols of FutureSource are used for their abbreviation. In alphabetical order of ticker symbol:

1. AAO — The Australian All Ordinaries Index (OTC).
2. AD — Australian Dollar Currency Futures (CME).
3. AXB — Australian 10-Year Bond Futures (SFE).
4. BD — US Long (30-Year) Treasury Bond Futures (CBT).
5. BP — British Pound Sterling Currency Futures (CME).
6. CD — Canadian Dollar Currency Futures (CME).
7. CP — Copper Futures (COMEX).
8. DGB — German 10-Year Bond (Bund) Futures (LIFFE).
9. DX — US Dollar Index Currency Futures (FNX).
10. ED — 90-Day Euro Dollar Futures (CME).
11. EU — Euro Currency Futures (CME).
12. FV — US 5-Year Treasury Note Futures (CBT).
13. FXADJY — Australian Dollar vs. Japanese Yen Cross Currency Forward (OTC).
14. FXBPJY — British Pound Sterling vs. Japanese Yen Cross Currency Forward (OTC).
15. FXEUBP — Euro vs. British Pound Sterling Cross Currency Forward (OTC).
16. FXEUJY — Euro vs. Japanese Yen Cross Currency Forward (OTC).
17. FXEUSF — Euro vs. Swiss Franc Cross Currency Forward (OTC).
18. FXNZUS — New Zealand Dollar Currency Forward (OTC).
19. FXUSSG — Singaporean Dollar Currency Forward (OTC).
20. FXUSSK — Swedish Krona Currency Forward (OTC).
21. GC — Gold 100 Oz. Futures (COMEX).
22. JY — Japanese Yen Currency Futures (CME).
23. LBT — Italian 10-Year Bond Forward (OTC).
24. LFT — FTSE-100 Index Futures (LIFFE).
25. LGL — Long Gilt (UK 10-Year Bond) Futures (LIFFE).
26. LML — Aluminum Futures (COMEX).
27. MNN — French Notional Bond Futures (MATIF).

28. SF — Swiss Franc Currency Futures (CME).
29. SI — Silver Futures (COMEX).
30. SJB — JGB (Japanese 10-Year Government Bond) Futures (TSE).
31. SNI — NIKKEI-225 Index Futures (SIMEX).
32. TY — 10-Year US Government Bond Futures (CBT).

These markets include most major asset classes traded through futures: fixed-income (short-term and long-term, both domestic and international), international equity indices, currencies and cross-currencies, and metals. Given set of 32 time series with daily rates of return covers a period of time between 6/12/1995 and 12/13/1999. Time is measured in trading days only, with a convention of five work-days per week, with adding previous day closing data for holidays with missing data.

A basic version of non-parametric bootstrap re-sampling generates “children” samples from the original “father” sample in the following way: it fills out a “child” with father’s daily rates of return in random order “with replacement”, i.e., when the same daily rate of return can be pulled out twice or more. Well-known difficulty in obtaining a re-sampled probability distribution function by such a procedure is that if the original “father” time series has certain auto-correlation structure, it will be totally lost in the “children”-re-samples because of random mixing in re-sample generation. At the same time, namely those auto-correlation properties of the time series, if present, should be responsible for the trend-following systems having positive rate of return. To remedy the situation, we will use a modification of a simple bootstrap re-sampling, which is called block-bootstrap re-sampling. Here is a brief description of the procedure.

First, we need to empirically study the correlation properties of the time series involved. For all data series, we have numerically calculated their auto-correlation coefficients  $C(n)$  for the period of 200 days, where  $n = 200$  is number of days in a period. The cut-off of 200 trading days was chosen in such a way that the measurements of correlation coefficient would still have some statistical accuracy on a sample length of 1076 days used.

Next, we empirically found a threshold for the absolute value of the auto-correlation coefficients equal to 2.5%, above which the values of coefficient larger than this threshold are statistically significant. Then, reducing number of days  $n$  in the period from 200 to 0, for all time series, we found the first value  $n^*$  that violates condition  $C(n) \leq 2.5\%$ . In this case, the value of  $n^*$ , which provides the statistically significant correlation lengths for all considered time series, is 100 trading days.

Now, instead of randomly picking an individual daily return from the original data series, we pick un-interchanged blocks of daily returns of length 100 trading days, starting from a random starting point. To ensure consistency across all time series, and preserve the cross-market correlation structure, we choose the same starting point for all 32 time series. That is, we use the same random starting point for all markets, then draw another starting point, and use it across all markets again, etc., until the necessary number of “children” re-samples will be filled-in.

### 7.3. Numerical results

We consider additional (“technological”) box constraints on portfolio weights  $0.2 \leq x_i \leq 0.8, i = \overline{1, 32}$ . This choice was dictated by the need to have the resultant margin-to-equity ratio in the account within admissible bounds, which are specific for a particular portfolio. In futures trading setup, “technological” constraints are analogous to the “fully-invested” condition from classical Sharpe–Markowitz theory [14], which make the efficient frontier strictly concave. In the absence of these constraints, the efficient frontier would be a straight line passing through (0,0), due to the virtually infinite leverage of these types of strategies. If all positions are equal to the lower bound 0.2, then the sum of the positions is  $0.2 \times 32 = 6.4$  and the minimal leverage is 6.4. However, if all positions are equal to the upper bound 0.8, then the sum of the positions is  $0.8 \times 32 = 25.6$  and the maximal leverage becomes 25.6. The optimal allocation of weights picks both the optimal leverage and proportions between instruments. Another subtle issue has to do with the stability of the optimal portfolios if the constraints are “too lax”. It is a matter of empirical evidence that the more lax the constraints are, the better portfolio equity curve you can get through optimal mixing, and the less stable with respect to walk-forward analysis these results would be. The above set of constraints was empirically found to be both leading to sufficiently stable portfolios and providing enough mixing of the individual equity curves.

We solved optimization problems (6.9), (6.10) and (6.3) with MaxDD, AvDD and 0.8-CDD ( $\alpha = 0.8$ ) measures, respectively, in the cases of 1-historical, 100 and

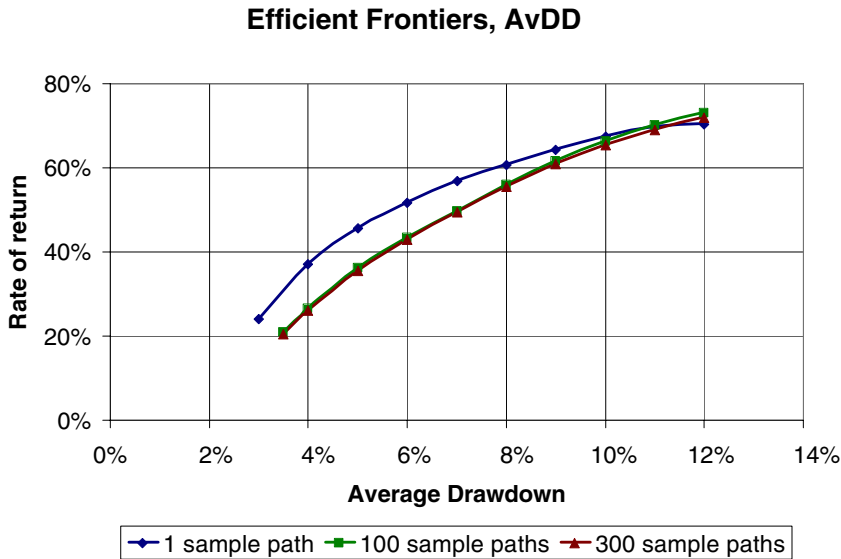


Fig. 6. Efficient frontiers: Average Drawdown.



300 sample paths generated for all 32 instruments. All optimization problems were solved using CPLEX package. The graphs of efficient frontiers and tables with optimal portfolio configurations for optimization problems with MaxDD, AvDD, and 0.8-CDD in all three cases: 1, 100 and 300 sample paths are presented by Figs. 6,

### Optimal risk-adjusted returns, AvDD

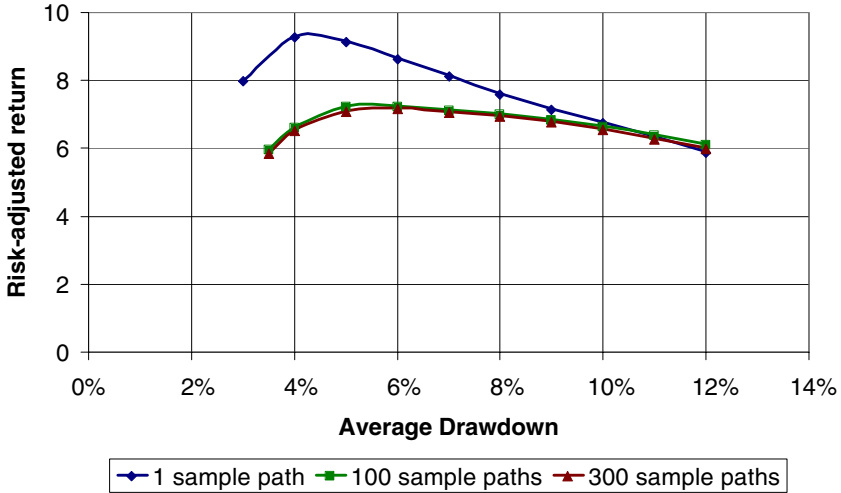


Fig. 7. Optimal risk-adjusted returns: Average Drawdown.

### Efficient Frontiers, 0.8-CDD

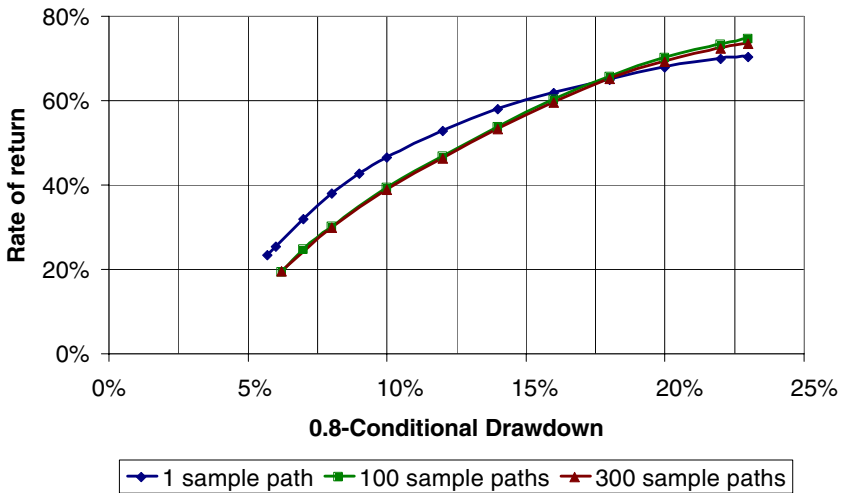


Fig. 8. Efficient frontiers: 0.8-Conditional Drawdown.

8, 10 and Tables 1–9, respectively. We, also, enclose the risk-adjusted returns (annualized rate of return divided by the corresponding value of a drawdown measure) for each of these cases, see Figs. 7, 9 and 11. The solutions achieving maximal risk-adjusted returns are boldfaced, see Tables 1–9.

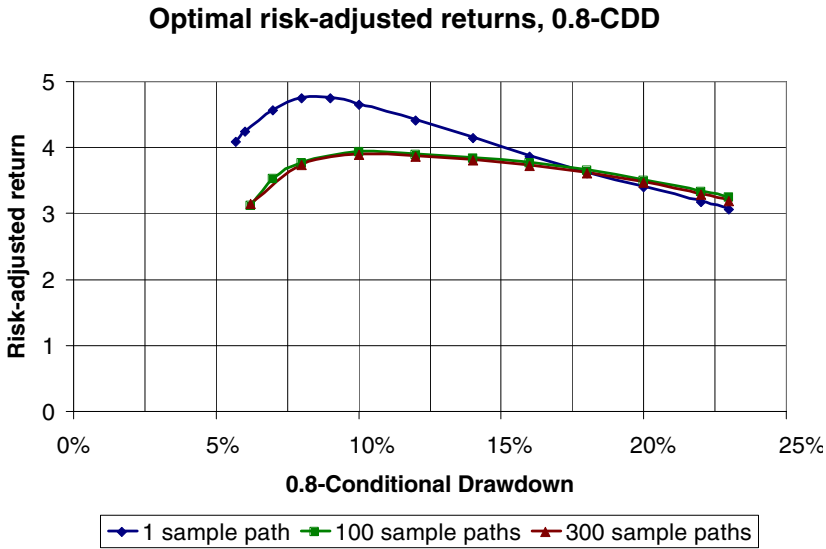


Fig. 9. Optimal risk-adjusted returns: 0.8-Conditional Drawdown.

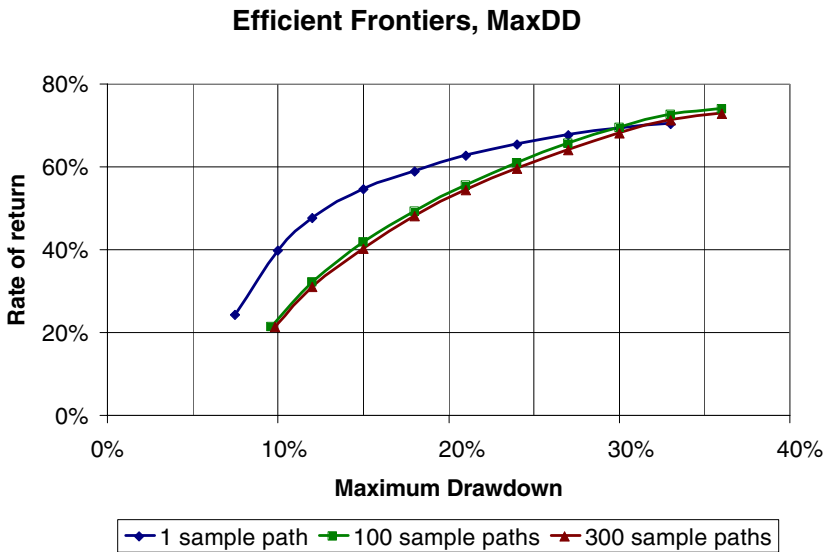


Fig. 10. Efficient frontiers: Maximum Drawdown.



Table 2. Average Drawdown optimization: 100 sample paths.

<i>Rate of return, %</i>	20.8	26.4	36.1	<b>43.4</b>	49.7	55.8	61.5	66.3	70.1	73.1
<i>AvDD, %</i>	3.5	4.0	5.0	<b>6.0</b>	7.0	8.0	9.0	10.0	11.0	12.0
<i>Risk-adj. return</i>	5.94	6.61	7.22	<b>7.23</b>	7.11	6.98	6.83	6.63	6.38	6.09
AD	0.20	0.20	0.20	<b>0.22</b>	0.30	0.36	0.77	0.80	0.80	0.80
BD	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.25	0.80	0.80
BP	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.20	0.20
CD	0.44	0.38	0.77	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80
CP	0.20	0.20	0.36	<b>0.50</b>	0.53	0.80	0.80	0.80	0.80	0.80
DX	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.20	0.48
ED	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.20	0.20
EU	0.20	0.35	0.80	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80
FXADJY	0.20	0.20	0.20	<b>0.26</b>	0.46	0.54	0.64	0.70	0.79	0.80
FXBPJY	0.20	0.29	0.69	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80
FXEUBP	0.20	0.20	0.22	<b>0.45</b>	0.75	0.80	0.80	0.80	0.80	0.80
FXEUJY	0.48	0.45	0.20	<b>0.20</b>	0.21	0.51	0.80	0.80	0.80	0.80
FXEUSF	0.24	0.47	0.80	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80
FXNZUS	0.20	0.20	0.20	<b>0.20</b>	0.44	0.74	0.80	0.80	0.80	0.80
FXUSSG	0.26	0.36	0.74	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80
FXUSSK	0.56	0.73	0.72	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80
FY	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.38	0.80	0.80	0.80
GC	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.34	0.78	0.80
JY	0.20	0.20	0.20	<b>0.20</b>	0.36	0.62	0.80	0.80	0.80	0.80
LIFT	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.20	0.20
LIGI	0.20	0.20	0.20	<b>0.20</b>	0.24	0.51	0.68	0.80	0.80	0.80
LIIB	0.20	0.20	0.20	<b>0.20</b>	0.28	0.44	0.72	0.80	0.80	0.80
LMAL	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.23	0.38	0.42	0.80
MANB	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.80	0.80	0.80
SF	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.64	0.80
SFAO	0.20	0.31	0.42	<b>0.77</b>	0.80	0.80	0.80	0.80	0.80	0.80
SFBD	0.20	0.20	0.20	<b>0.45</b>	0.58	0.80	0.80	0.80	0.80	0.80
SI	0.21	0.46	0.70	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80
SIJB	0.20	0.20	0.20	<b>0.24</b>	0.37	0.32	0.46	0.64	0.80	0.80
SINI	0.20	0.20	0.20	<b>0.51</b>	0.77	0.80	0.80	0.80	0.80	0.80
TY	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.20	0.60
UXBU	0.21	0.58	0.80	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80

Table 3. Average Drawdown optimization: 300 sample paths.

<i>Rate of return, %</i>	20.5	26.1	35.4	<b>42.9</b>	49.4	55.4	60.8	65.3	69.0	71.9
<i>AvDD, %</i>	3.5	4.0	5.0	<b>6.0</b>	7.0	8.0	9.0	10.0	11.0	12.0
<i>Risk-adj. return</i>	5.85	6.52	7.08	<b>7.16</b>	7.06	6.93	6.76	6.53	6.27	5.99
AD	0.20	0.20	0.20	<b>0.29</b>	0.26	0.51	0.80	0.80	0.80	0.80
BD	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.80	0.80
BP	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.20	0.20
CD	0.24	0.32	0.72	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80
CP	0.20	0.20	0.47	<b>0.50</b>	0.80	0.80	0.80	0.80	0.80	0.80
DX	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.23	0.20	0.63
ED	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.20	0.20
EU	0.20	0.37	0.80	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80
FXADJY	0.20	0.20	0.20	<b>0.20</b>	0.36	0.42	0.49	0.60	0.69	0.80
FXBPJY	0.20	0.20	0.66	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80
FXEUBP	0.20	0.20	0.27	<b>0.55</b>	0.77	0.80	0.80	0.80	0.80	0.80
FXEUJY	0.62	0.73	0.29	<b>0.33</b>	0.37	0.69	0.80	0.80	0.80	0.80
FXEUSF	0.23	0.52	0.80	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80
FXNZUS	0.20	0.20	0.20	<b>0.20</b>	0.46	0.62	0.80	0.80	0.80	0.80
FXUSSG	0.22	0.33	0.74	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80
FXUSSK	0.59	0.69	0.65	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80
FY	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.24	0.80	0.80	0.80
GC	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.22	0.62	0.72
JY	0.20	0.20	0.20	<b>0.20</b>	0.45	0.70	0.80	0.80	0.80	0.80
LIFT	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.20	0.20
LIGI	0.20	0.20	0.20	<b>0.20</b>	0.27	0.56	0.57	0.80	0.80	0.80
LIIB	0.20	0.20	0.20	<b>0.20</b>	0.20	0.45	0.72	0.80	0.80	0.80
LMAL	0.20	0.20	0.20	<b>0.22</b>	0.22	0.20	0.20	0.41	0.46	0.80
MANB	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.70	0.80	0.80	0.80
SF	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.35	0.80	0.80
SFAO	0.20	0.25	0.38	<b>0.63</b>	0.78	0.80	0.80	0.80	0.80	0.80
SFBD	0.20	0.20	0.20	<b>0.40</b>	0.63	0.80	0.80	0.80	0.80	0.80
SI	0.21	0.37	0.52	<b>0.72</b>	0.80	0.80	0.80	0.80	0.80	0.80
SIJB	0.20	0.20	0.20	<b>0.27</b>	0.32	0.38	0.50	0.69	0.80	0.80
SINI	0.20	0.20	0.22	<b>0.67</b>	0.80	0.80	0.80	0.80	0.80	0.80
TY	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.20	0.50
UXBU	0.29	0.69	0.80	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80

Table 4. 0.8-Conditional Drawdown optimization: 1 sample path.

<i>Rate of return, %</i>	23.3	25.5	31.9	<b>38.0</b>	46.6	52.9	57.9	61.8	65.0	70.3
<i>0.8-CDD, %</i>	5.7	6.0	7.0	<b>8.0</b>	10.0	12.0	14.0	16.0	18.0	23.0
<i>Risk-adj. return</i>	4.09	4.24	4.56	<b>4.75</b>	4.66	4.41	4.14	3.86	3.61	3.06
AD	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.42	0.55	0.33	0.80
BD	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.44	0.80	0.80
BP	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.20	0.56
CD	0.20	0.20	0.20	<b>0.20</b>	0.27	0.80	0.80	0.80	0.80	0.80
CP	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.80	0.80	0.80	0.80
DX	0.20	0.20	0.20	<b>0.20</b>	0.20	0.72	0.80	0.80	0.80	0.80
ED	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.20	0.20
EU	0.20	0.20	0.20	<b>0.50</b>	0.80	0.80	0.80	0.80	0.80	0.80
FXADJY	0.20	0.25	0.26	<b>0.25</b>	0.30	0.37	0.29	0.36	0.80	0.80
FXBPJY	0.77	0.80	0.80	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80
FXEUBP	0.20	0.33	0.39	<b>0.53</b>	0.76	0.80	0.80	0.80	0.80	0.80
FXEUIY	0.80	0.80	0.80	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80
FXEUSF	0.20	0.20	0.20	<b>0.52</b>	0.27	0.80	0.80	0.80	0.80	0.80
FXNZUS	0.20	0.25	0.40	<b>0.68</b>	0.80	0.80	0.80	0.80	0.80	0.80
FXUSSG	0.42	0.47	0.65	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80
FXUSSK	0.20	0.20	0.33	<b>0.43</b>	0.80	0.80	0.80	0.80	0.80	0.80
FY	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.20	0.80
GC	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.20	0.80
JY	0.20	0.25	0.60	<b>0.77</b>	0.80	0.80	0.80	0.80	0.80	0.80
LIFT	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.20	0.20
LIGI	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.80	0.80	0.80	0.80
LIIB	0.20	0.20	0.49	<b>0.44</b>	0.74	0.80	0.80	0.80	0.80	0.80
LMAL	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.26	0.59	0.48	0.80
MANB	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.80	0.80	0.80
SF	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.61	0.80
SFAO	0.20	0.20	0.20	<b>0.20</b>	0.61	0.80	0.80	0.80	0.80	0.80
SFBD	0.20	0.20	0.20	<b>0.20</b>	0.63	0.80	0.80	0.80	0.80	0.80
SI	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.20	0.80
SIJB	0.20	0.21	0.20	<b>0.29</b>	0.29	0.38	0.60	0.80	0.80	0.80
SINI	0.47	0.58	0.60	<b>0.67</b>	0.80	0.80	0.80	0.80	0.80	0.80
TY	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.20	0.80
UXBU	0.20	0.23	0.52	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80

Table 5. 0.8-Conditional Drawdown optimization: 100 sample paths.

<i>Rate of return, %</i>	19.4	24.7	30.1	<b>39.3</b>	46.7	53.6	60.1	65.7	70.0	74.5
<i>0.8-CDD, %</i>	6.2	7.0	8.0	<b>10.0</b>	12.0	14.0	16.0	18.0	20.0	23.0
<i>Risk-adj. return</i>	3.12	3.53	3.76	<b>3.93</b>	3.89	3.83	3.76	3.65	3.50	3.24
AD	0.20	0.20	0.20	<b>0.20</b>	0.21	0.29	0.62	0.80	0.80	0.80
BD	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.74	0.80
BP	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.20	0.55
CD	0.20	0.20	0.20	<b>0.33</b>	0.47	0.58	0.75	0.80	0.80	0.80
CP	0.20	0.20	0.20	<b>0.41</b>	0.46	0.62	0.80	0.80	0.80	0.80
DX	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.28	0.80
ED	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.20	0.20
EU	0.20	0.24	0.50	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80
FXADJY	0.20	0.20	0.20	<b>0.20</b>	0.23	0.36	0.45	0.60	0.79	0.80
FXBPJY	0.20	0.40	0.51	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80
FXEUBP	0.20	0.20	0.20	<b>0.43</b>	0.62	0.79	0.80	0.80	0.80	0.80
FXEUJY	0.39	0.54	0.49	<b>0.47</b>	0.77	0.80	0.80	0.80	0.80	0.80
FXEUSF	0.20	0.43	0.67	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80
FXNZUS	0.20	0.20	0.20	<b>0.20</b>	0.40	0.75	0.80	0.80	0.80	0.80
FXUSSG	0.20	0.39	0.52	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80
FXUSSK	0.51	0.74	0.73	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80
FY	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.24	0.67	0.80	0.80
GC	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.24	0.80
JY	0.20	0.20	0.20	<b>0.20</b>	0.25	0.41	0.70	0.80	0.80	0.80
LIFT	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.20	0.20
LIGI	0.20	0.20	0.20	<b>0.20</b>	0.20	0.35	0.72	0.80	0.80	0.80
LIIB	0.20	0.20	0.20	<b>0.23</b>	0.43	0.57	0.75	0.80	0.80	0.80
LMAL	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.37	0.80
MANB	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.77	0.80	0.80
SF	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.24	0.80	0.80
SFAO	0.20	0.20	0.32	<b>0.53</b>	0.78	0.80	0.80	0.80	0.80	0.80
SFBD	0.20	0.20	0.20	<b>0.20</b>	0.26	0.39	0.49	0.74	0.80	0.80
SI	0.20	0.20	0.26	<b>0.55</b>	0.72	0.80	0.80	0.80	0.80	0.80
SIJB	0.20	0.20	0.20	<b>0.40</b>	0.48	0.58	0.69	0.80	0.80	0.80
SINI	0.20	0.20	0.20	<b>0.24</b>	0.58	0.80	0.80	0.80	0.80	0.80
TY	0.20	0.20	0.20	<b>0.20</b>	0.20	0.20	0.20	0.20	0.20	0.80
UXBU	0.20	0.47	0.78	<b>0.80</b>	0.80	0.80	0.80	0.80	0.80	0.80











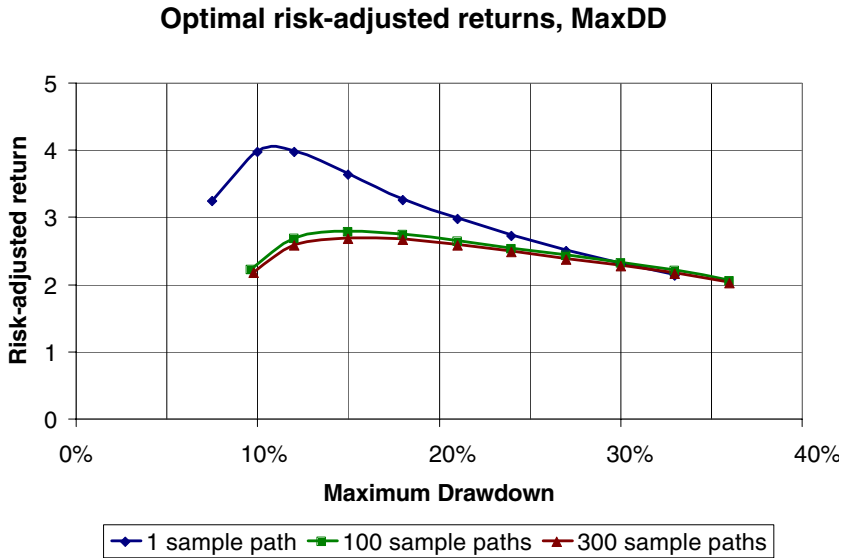


Fig. 11. Optimal risk-adjusted returns: Maximum Drawdown.

## 8. Conclusions

We introduced drawdown measure, which, we believe, is useful for practical portfolio management. This measure is similar to CVaR, includes the MaxDD and AvDD measures as its limiting cases and possesses all properties of a deviation measure. Moreover, it may be considered as a generalization of deviation measure to a dynamic case. We developed the optimization techniques that efficiently solve an asset-allocation problem with CDD, MaxDD and AvDD measures. We formulated and, for a real-life example, solved a portfolio optimization problem. These techniques, if implemented in a managed accounts' environment, will allow a trading or risk manager to allocate risk according to his/her personal assessment of extreme drawdowns and their duration on his/her portfolio equity.

We believe that however attractive the MaxDD measure is, the solutions produced using this measure in portfolio optimization may have a significant statistical error because the decision is based on a single observation of the maximal loss. Whereas CDD controls the worst  $(1 - \alpha) * 100\%$  of drawdowns, and due to statistical averaging within that range, obtains a better predictive power for the risk in the future, leading to a more stable portfolio. Our study indicates that the CDD with an appropriate level ( $\alpha = 0.8$ , i.e., optimizing over the 20% of the worst drawdowns) generates a more stable weights allocation than that produced using MaxDD measure.

Numerical results of the considered real-life asset-allocation problem with drawdown measure draw the following conclusions:

- The statistical accuracy is already sufficient for the case of 100 sample paths, i.e., difference between 100-sample path solutions and 300-sample paths solutions is negligible.
- For most of the allowable risk values (across all risk measures considered), the efficient frontier for stochastic (re-sampled) solutions lies below and is less concave than the so-called historical, or 1-scenario, efficient frontier. Only at the riskiest end of the efficient frontier, the efficient frontiers either converge to one another or intersect. This means that only for the riskiest portfolios, the stochastic, or, re-sampled solutions, provide an improvement to the risk-adjusted returns.
- The risk adjusted returns, especially at the optimal (maximal risk-adjusted return) point on the efficient frontier are, however, uniformly smaller than the re-sampled, or, stochastic solutions. On average, re-sampled optimal risk-adjusted returns solutions are 20% to 30% worse than those predicted by 1-path historical solutions. This result supports the wide-spread idea that using only one historical price path may lead (and probably does) to overstated and over-fitted results, which may not realize on average in the future. Though the results of the re-sampled or stochastic optimization lead to worse optimal solutions, those solutions are more trustworthy.
- Analyzing 32-dimensional vectors of instruments weights for the optimal historical and stochastic solutions, we found that they are substantially different: for example, the Euclidian norm of the stochastic optimal solution is, on average 50%, smaller than that for the historical optimal solution, and the angle between these vectors in our particular case is 50 degrees.

## Acknowledgments

Authors are grateful to Professor R. Tyrrell Rockafellar for valuable comments, which helped to improve the paper. We also thank Kyrylo Shmatov for programming help and stimulating discussions.

## References

- [1] C. Acerbi and D. Tasche, On the coherence of expected shortfall, *Journal of Banking and Finance* **26** (2002) 1487–1503.
- [2] P. Artzner, F. Delbaen, J.-M. Eber and D. Heath, Coherent measures of risk, *Mathematical Finance* **9** (1999) 203–227.
- [3] P. Artzner, F. Delbaen, J.-M. Eber, D. Heath and H. Ku, Multiperiod risk and coherent multiperiod risk measurement, preprint.
- [4] F. Delbaen, Coherent risk measures on general probability spaces, preprint.
- [5] F. Delbaen, Draft: Coherent risk measures, lecture notes, Pisa (2000).
- [6] A. Chekhlov, S. Uryasev and M. Zabarankin, Portfolio optimization with drawdown constraints, in *Asset and Liability Management Tools*, ed. B. Scherer (Risk Books, London, 2003) pp. 263–278.
- [7] J. Cvitanic and I. Karatzas, On portfolio optimization under “drawdown” constraints, *IMA Lecture Notes in Mathematics & Applications* **65** (1995) 77–88.

- [8] R. S. Dembo and A. J. King, Tracking models and the optimal regret distribution in asset allocation, *Applied Stochastic Models and Data Analysis* **8** (1992) 151–157.
- [9] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap* (Chapman & Hall/CRC, New York, 1993).
- [10] P. Jorion, *Value at Risk: A New Benchmark for Measuring Derivatives Risk* (Irwin Professional Publisher, 1996).
- [11] R. C. Grinold and R. N. Kahn, *Active Portfolio Management* (McGraw-Hill, New York, 1999).
- [12] S. J. Grossman and Z. Zhou, Optimal investment strategies for controlling draw-downs, *Mathematical Finance* **3**(3) (1993) 241–276.
- [13] H. Konno and H. Yamazaki, Mean absolute deviation portfolio optimization model and its application to tokyo stock market, *Management Science* **37** (1991) 519–531.
- [14] H. M. Markowitz, Portfolio selection, *Journal of Finance* **7**(1) (1952) 77–91.
- [15] H. Mausser and D. Rosen, Beyond VaR: From measuring risk to managing risk, *ALGO Research Quarterly* **1**(2) (1999) 5–20.
- [16] G. C. Pflug, Optimal scenario tree generation for multiperiod financial planning, *Mathematical Programming Ser. B* **89** (2001) 251–271.
- [17] G. C. Pflug, Some remarks on the Value-at-Risk and the conditional Value-at-Risk, in *Probabilistic Constrained Optimization: Methodology and Applications*, ed. S. P. Uryasev (Kluwer, 2000) pp. 278–287.
- [18] G. C. Pflug and A. Ruszczyński, A risk measure for income streams, in *Risk Measures for the 21st Century*, ed. G. Szego (Wiley, 2004).
- [19] R. T. Rockafellar, *Convex Analysis* (Princeton University Press, 1970).
- [20] R. T. Rockafellar and S. P. Uryasev, Optimization of conditional Value-at-Risk, *Journal of Risk* **2** (2000) 21–42.
- [21] R. T. Rockafellar and S. P. Uryasev, Conditional Value-at-Risk for general loss distributions, *Journal of Banking and Finance* **26** (2002) 1443–1471.
- [22] R. T. Rockafellar, S. Uryasev and M. Zabarankin, Deviation measures in risk analysis and optimization, research report 2002–7, ISE Dept., University of Florida (2002).
- [23] R. T. Rockafellar, S. Uryasev and M. Zabarankin, Deviation measures in generalized linear regression, research report 2002–9, ISE Dept., University of Florida (2002).
- [24] R. T. Rockafellar, S. Uryasev and M. Zabarankin, Portfolio analysis with general deviation measures, research report 2003–8, ISE Dept., University of Florida (2003).
- [25] M. R. Young, A minimax portfolio selection rule with linear programming solution, *Management Science* **44**(5) (1998) 673–683.
- [26] W. T. Ziemba and J. M. Mulvey (eds.), *Worldwide Asset and Liability Modeling* (Cambridge University Press, 1998).
- [27] Sound practices for hedge fund managers, A response to the President’s Working Group (PWG) report on LTCM crisis by the leading hedge funds: Caxton Corporation, Kingdon Capital LLC, Moore Capital Management Inc, Soros Fund Management LLC, Tudor Investment Corporation (February 2000).