Tri-Criteria Optimization for Scenario-Based Risk Measures

August 31, 2015

by

Henry Clausen

ETH Zurich
Institute for Theoretical Physics
Econophysics Group
Supervised by Prof. Dr. Diethelm Würtz

Seminar Thesis Submitted in Partial fulfillment of the
Requirements for the Degree of

MSc. Computational Science and Engineering
Field of Specialization: *Theoretical Physics*

in the Departments of Mathematics and Physics

ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

ETH ZURICH, AUGUST 2015

Abstract

In modern portfolio theory an optimal portfolio is defined as a portfolio that optimizes both the expected future return and the risk of the investment. Established methods in portfolio optimization quantify the estimated risk of an asset by a chosen risk measure and optimize the portfolio using this risk measure. Currently, portfolios are only optimized for a single chosen risk measure. Since every risk measure has flaws, one might desire to optimize a combination of different risk measures. The task of portfolio optimization is then transformed from a bi-criteria to a multi-criteria problem. In this work the requirements on such criteria in the context of portfolio optimization is discussed and a simple algorithm to solve three-criteria portfolio optimization is presented.

Zusammenfassung

In der modernen Portfoliotheorie ist ein Portfolio effizient, wenn es sowohl den erwarteten Profit als auch das totale Risiko der Investition optimiert. Etablierte Optimierungsmethoden messen das Risiko einer Investition durch Risikomaße und optimieren ein Portfolio mithilfe dieser. Bis jetzt existieren ausschließlich Methoden, die Portfolios mithilfe eines einzigen Risikomaßes optimieren. Da jedes Risikomaß Schwächen besitzt liegt es nahe, eine Kombination von mehreren Risikomaßen zu optimieren. Das traditionelle Portfolio-Problem ändert sich dann von einem Bi-Criteria zu einem Multi-Criteria Problem. In dieser Arbeit werden die Anforderungen an zusätzliche Kriterien besprochen und ein simpler Algorithmus zur Multi-Criteria Optimierung wird vorgestellt.

Acknowledgement

First of all, I would like to thank Prof. Diethelm Würtz for giving me the opportunity to write my thesis at his group, for his guidance and for sharing his experience and expertise with me. He supervised the project in an excellent way. It was a very pleasant experience to work with him.
Further thanks are addressed to Tobias Setz for his help and patience during the project. His input was often of great value.
Finally, I would like to thank the Institute for Theoretical Physics for providing the possibility to realize the project and for helping me enrich my knowledge.

Contents

1 Introduction

1.1 Background

A field of big interest for quantitative research departments is the development of portfolio[1] optimization approaches. The aim of such optimizations is to allocate portfolios in an optimal position in order to maximize the future return of the portfolio while simultaneously minimizing the chance of a potential capital loss. The latter is also called the risk of the portfolio and is the main subject of interest in portfolio optimization. The risk of a portfolio cannot be measured precisely. Its evaluation is subject to the field of *Risk assessment* and there exist many approaches to estimate the risk of a portfolio.
The most famous portfolio optimization approach was introduced by Harry Markowitz in his paper *Portfolio Selection* [20] in 1952. His approach quantifies the portfolio risk by the volatility of its preceding returns. However this approach has been subject to criticism and is said to be outdated [16].
Currently, the tendency is to use downside risk measures such as Average Drawdown at Risk (CDaR) or Conditional Value-at-Risk (CVaR) for portfolio optimization. These risk measures belong to the group of *scenario-based risk measures* [3] and describe the risk of a portfolio as the worst average loss under a set of weighted scenarios.
Scenario-based risk measures proved to comply better with investors perception of risk and are less error-prone to actual return distributions. They belong to the class of coherent risk measures that was defined by Artzner et al. [3] and are furthermore convenient to use in optimization.
Nowadays a different approach to risk reduction received attention: Risk budgeting and tail dependence analysis seek to diversify the risk dependencies of a portfolio [4]. This can help reduce losses in declining market situations. However risk dependencies are normally hard to estimate.

1.2 Aim of this work

Until now, portfolio optimization approaches have normally been bi-criteria optimization problems: Maximizing return while minimizing a chosen risk measure. However there does not (yet) exist a risk measure that fully captures the riskiness of an investment. Scenario-optimized portfolios tend to have high risk concentrations while pure diversification diminishes returns.

[1] A portfolio is a collection of different economic assets.

One might therefore seek to optimize multiple criteria simultaneously such as the return, the conditional value at risk and the diversification of the portfolio. Doing this could reduce the individual deficiencies of the risk measures and improve the performance of the optimized portfolios under different scenarios.
In this thesis we will introduce a method to optimize portfolios under three criteria: The expected return, a scenario-risk measure and a quadratic diversification measure. We will show that individual risk and diversification measures can be easily substituted and implemented. An exemplary implementation of this method will be presented for the CVaR as risk measure and the Herfindahl index as diversification measure and the main characteristics of this method will be discussed.

1.3 Thesis Structure and Methodology

The thesis is structured as following:

- Chapter 2 sketches the core concepts of portfolio theory and describes the basic notion of returns, risk and diversification. It contains a description of logarithmic returns and of variance risk and its limitations. Furthermore, it answers the question of which properties are needed to fulfill the requirements for a coherent risk measure. Scenario-based risk measures are introduced and their advantage over volatility is illustrated. As an representative of scenario-based risk measures, the basic concepts of the CVaR are discussed. Finally different concepts of diversification measures are introduced and characterized.

- Chapter 3 introduces the basic definitions and concepts of multi-criteria optimization. The ϵ-constraint method is illustrated and used for the framework of bi-criteria portfolio optimization. It is then briefly described how the CVaR can be linearized in the portfolio framework so that the portfolio problem for the CVaR can then be solved computationally.

- Chapter 4 is the main chapter of this thesis. It introduces the concept of the *Weighted Sum Method* and analyzes its efficiency to find Pareto optimal solutions. It is then discussed whether and under what conditions it can be applied to basic portfolio optimization. Furthermore the traditional bi-criteria portfolio problem of optimizing expected return and risk of the portfolio is extended to a tri-criteria problem. It is then discussed how this problem can be solved using the weighted sum method and how well portfolio allocation works with this method. An implementation of a tri-criteria portfolio optimization is then presented and its benefits are illustrated.

- Chapter 5 briefly analyzes the profitability of investment strategies obtained from the implemented optimization method.
- Chapter 6 discusses the computational performance of the presented implementation.
- In chapter 7 concludes the thesis, emphasizes important results and gives an outlook for further optimization possibilities of this topic.

1.4 Datasets

For demonstrative purposes several datasets of stock returns were used in this thesis. These datasets where downloaded from `http://finance.yahoo.com/`.
For the visualizations of the concepts in chapter 3 and 4 we used a dataset with daily log-returns of assets in the Dow-Jones stock index between 2014-06-27 and 2015-06-25. Additionally a dataset of assets in the LPP2005 benchmark index[2] between 2006-03-20 and 2007-04-11 is used. This index consists of only 6 assets and interesting for our discussion in chapter 4 as it illustrates the difficulties of our implemented method.
For comparison reasons, we additionally depicted the same visualizations for a dataset of assets in the DAX-stock index between 2014-08-04 and 2015-07-27. These figures can be found in the Appendix.
For the Rolling-window analysis in chapter 5 we used datasets of assets in the Dow-Jones stock index between 2000-01-05 and 2015-07-08 and a dataset of the stock components of the DAX index between 2003-01-07 and 2015-07-27.
For the runtime comparisons in chapter 6 we used the same dataset of the Dow-Jones index as in chapter 3 and 4

[2]The LLP2005 benchmark index is published by the Pictet Group. It is an index designed to represent the investment strategies of Swiss pension funds. It consists of six independent assets.

2 Basic Theory

In this chapter the basic concepts of portfolio theory will be explained and the framework that will be used throughout this thesis will be established. The concepts of returns, risk and diversification will be established and the important characteristics of these concepts will be emphasized. The reader is advised to pay careful attention to the convexity (concavity) of the introduced measures as this is an essential characteristic for chapter 4.

2.1 Assets, portfolios and returns

A financial *asset* is an economic resource with an ownership that can be converted into cash. Both tangible objects such as gold or intangible objects like a bond can be assets as long they are assigned a positive economic value [26]. The price of an asset is defined between the owner of the asset and the issuer, a person that has announced interest in the asset. The typical place were trades of assets take place are financial markets. Common examples of assets are

- bonds,
- equities,
- derivatives,
- currencies, and
- funds.

A *portfolio* is a collection of assets investments [22]. The amount of investment into each asset may differ. The purpose of a portfolio is to generate wealth through accretion of the held assets while reducing the investment risk via diversification.
We will define the amount of investment into asset i by x_i. The portfolio can then be characterized by the vector $\bar{x}$ of all asset weights x_i.
We will normalize the total investment volume of an investor to be equal to one so that we have a consistent framework. Furthermore we will require that all available resources are invested. This normalization can then be expressed as

$$\sum_{i}^{N} x_i = 1 \tag{1}$$

Normally in portfolio selection theory, only long-term portfolios are considered which means that we forbid short selling. Consequently we will introduce the additional constraint

$$\forall i \in \{1, ..., N\}: \; x_i \geq 0. \tag{2}$$

We will work with this convention throughout this thesis.

The traditional **portfolio problem** approaches the objective of finding a portfolio that minimizes risk while maximizing expected return. As these aspirations are often not compatible, the portfolio making the ideal compromise is the object of interest [22]. To explore this problem further, we will introduce the concepts of risk and returns first.

Returns

The price of financial assets normally changes over time. To measure the appreciation of an asset, a quantity called the *return* of an asset is used. The *simple net return* R_t of an asset over a period of time is defined as

$$R^t = \frac{P^t - P^{t-1}}{P^{t-1}} \tag{3}$$

where P^t is the price of the asset at the beginning of the period and P^{t+1} is the price at the end of the period. Returns are a percentage and can be either positive or negative.
More often used than the simple net return is the *continuously compounded return* or *log return* r^t of an asset:

$$r^t = \ln(1 + R^t) = \ln(\frac{P^t}{P^{t-1}}) \tag{4}$$

An advantage of the log return over the net return is that log returns are additive under aggregation of prices: If we want the cumulative log return over k periods, we sum over all k log returns

$$r^{t[k]} = \ln(\frac{P^t}{P^{t-k}}) = \sum_{i=t-k}^{t} r^i \tag{5}$$

Thus in this thesis log returns will be used instead of simple net returns. When speaking of returns, the term should be understood as log returns.

Expected return

The future value of an asset cannot be exactly predicted. However it is of great interest for investors to make an estimation on the future behaviour of an asset. If we assume the return to behave like a random variable, the best unbiased estimator of the expectation value of the return $E(r)$ is the mean value of all prior returns [22]:

$$E(r) = \frac{1}{N} \sum_{i=1}^{N} r^i \tag{6}$$

where N is the number of time steps (i.e. trading days) that are incorporated in the calculation. Note that for log returns, $E(r)$ is just the return over all periods $r^{t[N]}$ (see equation 5).
The expected return of a portfolio of n assets is given by the sum

$$E(r_P) = \sum_i x_i E(r_i) = \bar{x} \cdot \overline{\mu} \tag{7}$$

where $r_1, ..., r_n$ are the log returns of the different assets and $\mu_i = \frac{1}{N} \sum_{t=1}^{N} r_i^t$. Note that the expected return is a **linear function** of the portfolio weights $\bar{x}$. This fact will become crucial in the later discussion.

The expected return is no guarantee for future returns. However, it is an established forecast to estimate the future value of an asset and it provides a measure of actual returns.

2.2 Risk

Trading in assets whose future values are uncertain necessarily involves risk for the investors. The management of risk is a major concern when operating in financial markets. The desire for reduction of investment risk is also the driving force in portfolio optimization. This section will discuss several concepts of risk assessment.

Variance

The most widely used risk measure is the variance. Harry Markowitz proposed 1952 in his article *Portfolio selection** [20] to identify the riskiness of an asset with the variance of its returns. It is the simplest established risk measure and for that reason the basis for the traditional portfolio theory, also

called mean-variance portfolio theory [26]. This theory treats the returns as random variables with normal distribution[3]. The variance measures the return fluctuations of an asset: The higher the variance, the greater are the return differences during the observed time period.
Assume normally distributed random variables. The variance of a random variable X is defined [14] as

$$\sigma^2(X) = E\big[(X - E[X])^2\big]$$

The covariance between two random variables X_i, X_j is defined to be

$$\sigma(X_i, X_j) = E\Big[\big(X_i - E[X_i]\big)\big(X_j - E[X_j]\big)\Big]$$

Notice that $\sigma(X, X) = \sigma^2(X)$.
The variance of a weighted sum of random variables

$$\bar{X}_{\bar{a}} = \sum_i a_i X_i$$

is then defined as

$$\begin{aligned} \sigma^2(\bar{X}_{\bar{a}}) = \sigma(\bar{X}_{\bar{a}}, \bar{X}_{\bar{a}}) &= \sum_i \sum_j a_i a_j \sigma(X_i, X_j) \\ &= \bar{a}^T \, \mathbf{\Sigma} \, \bar{a} \end{aligned} \tag{8}$$

where $\mathbf{\Sigma}$ is the covariance matrix between all random variables X_i and $\bar{a}$ is the vector of weights. The variance of a portfolio $\bar{x}$ can then be written as

$$\sigma^2(\bar{x}) = \bar{x}^T \, \mathbf{\Sigma} \, \bar{x} \tag{9}$$

Markowitz understood this quantity as the risk of a portfolio.
Note that the variance as well as the covariance of random variables is not necessarily finite for non-normally distributed random variables! We will talk about that issue in the next section.
The covariance matrix is positive semi-definite [14]. An important consequence of this is that the variance risk measure is a convex function of portfolio weights, i.e. for two sets of portfolio weights $\bar{x}_1$ and $\bar{x}_2$ and $a \in [0, 1]$ we have

$$\sigma^2(a\bar{x}_1 + (1-a)\bar{x}_2) \leq a\sigma^2(\bar{x}_1) + (1-a)\sigma^2(\bar{x}_2) \tag{10}$$

[3] This requirement can be weakened to **elliptical** distributions [5, 23].

We can refine this statement for the case where no asset is perfectly correlated to a combination of other assets (i.e. the span of all N random variables (assets) is N-dimensional): Assume a set of N random variables (assets) $X_i,\ i \in \{1, .., N\}$. If

$$\forall i \in \{1, .., N\} \ \nexists \bar{a} \in \mathbb{R}^{N-1},\ c \in \mathbb{R} :\ X_i = \sum_{k \neq i}^{N} a_k X_k + c \tag{11}$$

then the covariance matrix is positive definite and thus the variance is a **strictly convex function** of portfolio weights:
For two sets of portfolio weights $\bar{x}_1$ and $\bar{x}_2$ and $a \in (0, 1)$ we have

$$\sigma^2(a\bar{x}_1 + (1-a)\bar{x}_2) < a\sigma^2(\bar{x}_1) + (1-a)\sigma^2(\bar{x}_2) \tag{12}$$

Risk measures that are strictly convex under the condition in equation 11 are also called *strictly convex modulo translation* [6].

Disadvantages of the variance risk measure

- An enormous drawback of Markowitz's portfolio model is the requirement of multivariate normal distributions for the asset returns. This requirement is necessary in order to provide finite variance and expection values. In reality, returns are rarely normally distributed.

  In 1983, Owen & Rabinovitch and Chamberlain extended Markowitz's portfolio model such that the requirement for multivariate normal distributions can be generalized to the requirement of multivariate elliptic distributions while the mean-covariance matrix is generalized to the mean-characteristic matrix [5, 23].

  In 2012 Chicheportiche & Bouchaud published their article *The joint distribution of stock returns is not elliptical* [7], which states that in reality asset returns have bigger tails and are far from being jointly normally (nor elliptically) distributed.

  Figure 1 shows a comparison between a normal distribution and the actual distribution of returns for the Dow Jones index. The y-axis is log scaled. The heavy tails of the actual distribution are clearly visible compared to the normal distribution.

- The Markowitz portfolio model works with the true covariance of the assets. In reality, covariance estimators have to be used to obtain a covariance matrix. These estimators often generate a biased result and tend to underestimate the risk. N. E. Karoui et al. [17] showed that the

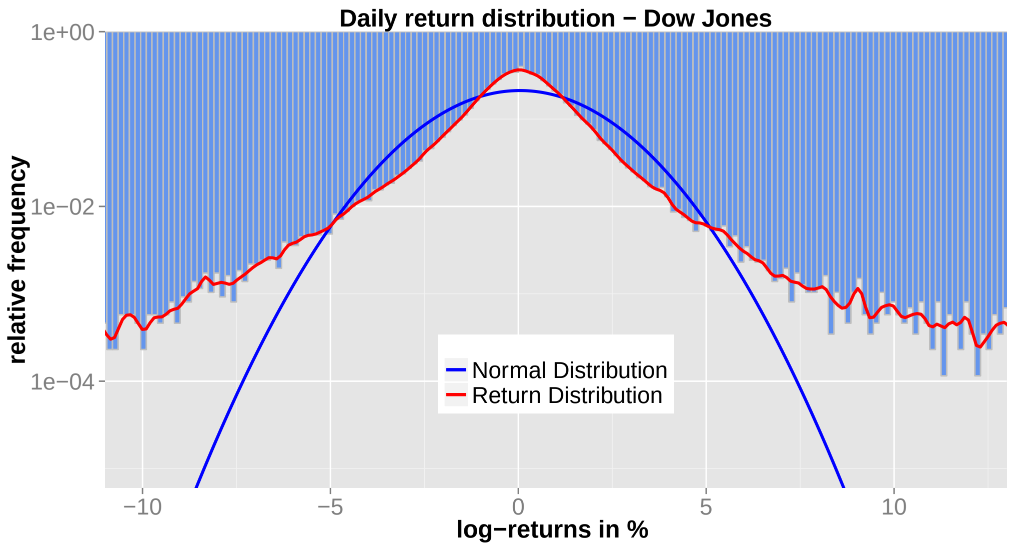

Figure 1: A comparison between a fat tailed daily return distribution (red) and a normal distribution (blue) with similar mean and variance. The histogram contains daily log returns of assets in the Dow Jones index from March 20 2008 to June 25 2015. The density distribution was calculated from kernel density estimation using the *density*-function in the R-package *stats*[5].

underestimation of the risk is minimal for i.i.d. normal distributions, but more pronounced in an elliptic and correlated scenario ([17] and [10]). For heavy-tailed distributions, it is in general difficult to estimate the variance (if finite) correctly.

- For normally distributed returns, the variance might be an adequate risk measure as it fully describes the distribution. For general distributions, other characteristics might be a more realistic representation of the way in which investors perceive risk since the variances penalizes both downside and upside extremes. In reality, risk is understood as an asymmetric measure as only downside extremes are undesired. Markowitz later favoured semi-variance and downside deviation as risk measures [13]. Another alternative risk measure that we will focus on in the next section is the *Conditional Value at Risk* (CVaR).

Figure 2 shows a comparison between a normal distribution and the actual asymmetric distribution of returns of Exxon Mobil. The actual

[5]http://stat.ethz.ch/R-manual/R-patched/library/stats/html/density.html

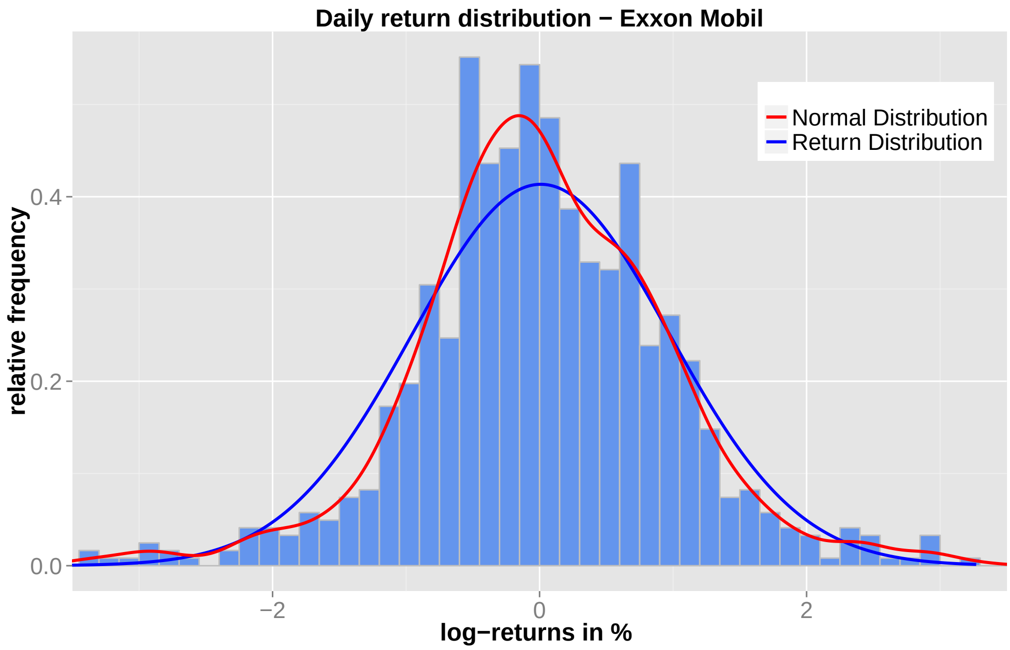

Figure 2: An example of an asymmetric daily return distribution (red) compared to a normal distribution (blue) with similar mean and variance. The histogram contains 770 daily log returns of Exxon Mobil from April 05 2012 to June 25 2015. The density distribution was calculated from kernel density estimation using the *density* function in r-stats.

return distribution shows a clear kurtosis. Due to this kurtosis the risk of a loss is different from a normal distribution.

- Despite the structural flaws of the variance risk, it can - like any other risk measure that is based on analysing posterior returns - be criticized for relying on the past of the asset. An asset can behave differently in the future than it did before. A projection of the past riskiness of an asset into the future is not always correct, especially in turbulent times.

Coherent risk measures

As the variance does not satisfy the requirements of a risk, Artzner et al. [3] proposed a framework that proper risk measures should fulfill in order to suffice the perception of a risk. These risk measures are called *coherent risk measures*. Artzner defined the conditions on a coherent risk measure as following:

Definition 1. *Suppose Ω is a linear vector space of random variables $X_1, X_2, ...$ with finite expectation value (in our context the space of daily returns). A* ***coherent risk measure*** *on Ω is a function*

$$\rho : \Omega \to \mathbb{R} \cup \{+\infty\} \tag{13}$$

that satisfies the following axioms:

Axiom 1: Normality

$$\rho(0) = 0 \tag{14}$$

In other words, there is no risk in holding no assets. Artzner additionally required that

$$\forall X \neq 0 : \ \rho(X) > 0, \tag{15}$$

i.e. there are no risk free assets.

Axiom 2: Monotonicity

$$\text{If } X_1, X_2 \in \Omega \text{ and } X_1 \leq X_2 \Rightarrow \rho(X_1) \geq \rho(X_2) \tag{16}$$

The $\leq$ is to be interpreted as a probabilistic almost all. In our context this means that if almost all returns generated by asset A are greater than all returns generated by asset B, then A cannot be riskier than B. This condition is not fulfilled by the variance as a portfolio that only generates losses can have a lower variance than a portfolio that only generates positive returns.

Axiom 3: Subadditivity

$$\forall X_1, X_2 \in \Omega : \ \rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2) \tag{17}$$

This axiom will later be important for us. It states that every combination of assets A and B cannot be riskier than the sum of the single risks of A and B, i.e. that diversification does not increase the risk.

Axiom 4: Positive homogeneity

$$a \geq 0 : \ \rho(\alpha X) = \alpha\rho(X) \tag{18}$$

This just means that the risk scales linearly with the amount of investment in an asset.

Axiom 5: Translation invariance

$$\alpha \geq 0: \ \rho(X+\alpha) = \rho(X) - \alpha \tag{19}$$

Axiom 5 states that if you add cash to your portfolio, its risk reduces. This axiom is also not fulfilled by the variance.

Artzner et al. [3] understood these conditions as the minimum requirements on a risk measure. Every risk measure that does not fulfill these requirements can not resemble the human perception of risk.

Definition 2. *Let Ω be a probability sample space. Let $\mathcal{P}$ be a set of probability measures on Ω. Let $\mathcal{G}$ be the linear vector space of random variables on Ω. A risk measure defined by the function*

$$\rho_{\mathcal{P}}(X) = \sup_{\mathbb{P}\in\mathcal{P}} \left\{ E_{\mathbb{P}}[-X] \right\} \tag{20}$$

on $\mathcal{G}$ is called a scenario-based risk measure.

Artzner et al. proofed[6] that a risk measure is coherent if and only if it is of the form in equation 20 [3]. The principle idea behind scenario-based risk measures is to measure maximum expected loss of the asset where certain scenarios can be weighted heavier than others to mitigate their effect on the result.
Scenario-based risk measures possess a property that will come very handy for us. Cheridito et al. [6] proofed the following for all scenario-based risk measures:

Theorem 1. *Let Ω be a probability sample space. Let $\rho_{\mathcal{P}}(X)$ be a scenario based risk measure. Let $X_1(\omega), ..., X_N(\omega) \in \mathcal{G}, \ \omega \in \Omega$ be a set of random variables on Ω with finite expectation value.*
Assume the following: For all subsets $\mathcal{A} \in \Omega$ with $P(\mathcal{A}) > 0$, none of the random variables $X \in \mathcal{G}$ is linearly dependent on the others, i.e. $\nexists i \in \{1,..,N\} : \forall \omega \in \mathcal{A} \ \ X_i(\omega) = \sum_{k\neq i} a_k X_k(\omega)$.

Then $\rho_{\mathcal{P}}(Y)$ is ***strictly convex*** *in portfolio weights.*

[6]The exact statement is as following: **Proposition**: *A risk measure ρ is coherent if and only if there exists a family $\mathcal{P}$ of probability measures on the set of states of nature, such that $\rho(X) = \sup_{\mathbb{P}\in\mathcal{P}} \left\{ E_{\mathbb{P}}[-X] \right\}$*

2.3 CVaR as an example for scenario-based risk measures

Already before Artzner et al.'s [3] publication of a framework for coherent risk measures, investors were aware of the downsides of the variance risk and consequently strived for alternative risk measures.
In this section we will analyze a well established risk measure, the *Conditional Value at Risk* (CVaR), as a representative for the class of scenario-based risk measures. For that, a brief overview over the *Value at Risk* (VaR), the basis for the CVaR, is needed.

Value at Risk

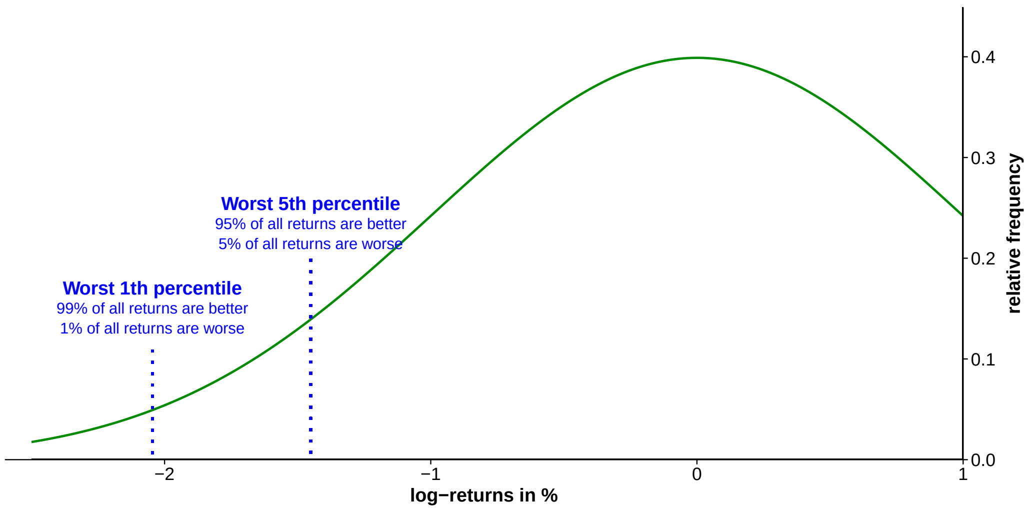

Figure 3: Visualization of the VaR at 1% and 5% for a gaussian distribution. 99% resp. 95% of all returns are on the greater than the $\mathrm{VaR}_{0.01}$ resp. the $\mathrm{VaR}_{0.05}$, only 1% resp. 5% of all returns are lower.

The concept of the VaR emerged in the late 1980s, as a consequence of the stock market crash in 1987 [16]. The events during the crash were much more extreme than the standard statistical models would allow. Hence investors desired a better risk modeling of the extreme losses of a return distribution. In 1994 JP Morgan published the first public methodology of the VaR.
Artzner et al. [3] define the VaR at α of a distribution as the threshold where the probability of a loss exceeding the VaR is α. More mathematically, if X is a random variable and $\alpha \in (0, 1)$, then

$$\mathrm{VaR}_\alpha(X) := -\inf\{l \in \mathbb{R} : P(l < X) \leq 1 - \alpha\} \tag{21}$$

or in other words, the $\mathrm{VaR}_\alpha(L)$ is the level-α-quantile of L.
For investors, the VaR signifies the threshold such that losses exceeding the threshold become unlikely.

Conditional Value at Risk

Apparently the VaR is not a coherent risk measure as it fails to satisfy the subadditivity property. Consequently the VaR can discourage diversification and can assign a higher VaR to definitely less risky portfolios [1, 2]. Another disadvantage of the VaR compared to the variance risk is that it is very difficult to use in optimization. VaR is non-linear and non-convex, thus making it very hard to compute a minimum VaR-portfolio, especially when the return distribution is not normal [21]. Therefore a risk measure based on the VaR was introduced in order to diminish these issues:

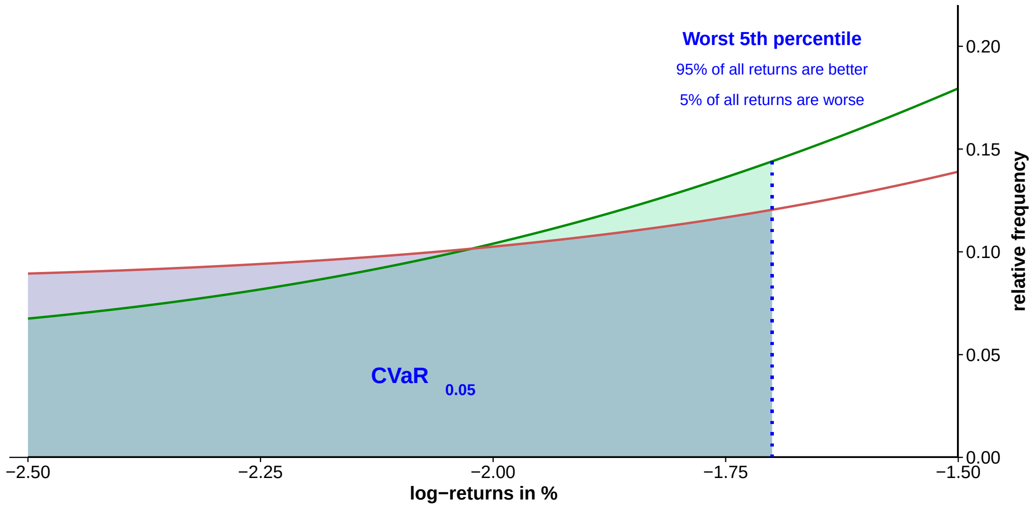

Figure 4: Visualization of the CVaR at 5% for two distributions. Although both distribution have the same $\mathrm{VaR}_{0.05}$, their $\mathrm{CVaR}_{0.05}$ is different due to their different tail behaviour.

Uryasev and Rockafeller [25] define the *conditional value at risk* (CVaR) at α, also called expected shortfall at α, of an asset as the conditional expectation of VaR_γ given that $\gamma \leq \alpha$:

Definition 3. *Given that X is a random variable with finite expectation value and $\alpha \in (0, 1)$, then the conditional value at risk $\phi_\alpha(X)$ is given by*

$$CVaR_\alpha(X) := \frac{1}{\alpha} \int_0^\alpha VaR_\gamma(X) d\gamma \tag{22}$$

If the distribution of X is continous, then the $CVaR$ is equivalent to the tail conditional expectation:

$$\mathrm{CVaR}_\alpha(X) := -\frac{1}{\alpha} E\Big[X \ \mathbf{1}_{\{X \leq -\mathrm{VaR}_\alpha(X)\}}\ \Big] \tag{23}$$

Distributions with different tails can have the same VaR_α. The behaviour of the tails is then better characterized by the CVaR_α as depicted in figure 6. As the CVaR is a scenario-based risk measure, it is a coherent risk measure [2] as well. From subadditivity and positive homogeneity it follows that CVaR is a convex function of the portfolio weights $\bar{x}$ and thus better applicable in optimization than VaR. Furthermore, the CVaR is a **strictly convex function** for a set of assets that fulfill the conditions of theorem 1.

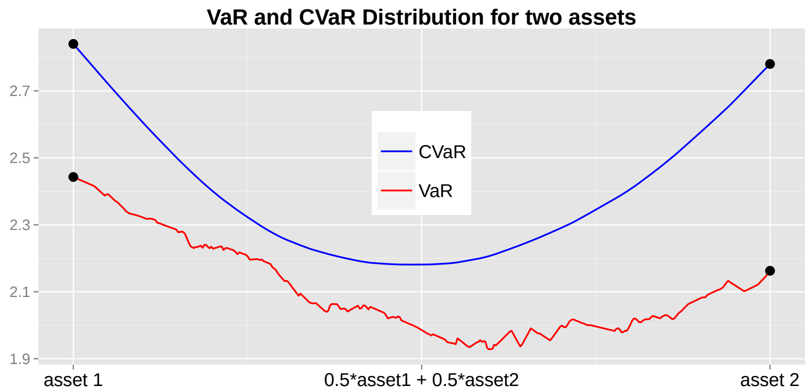

Figure 5: Comparison of the VaR and the CVaR at$\alpha = 0.05$ for combinations of two assets. Unlike the CVaR, the VaR is neither convex nor smooth.

Uryasev and Rockafeller [24] showed that the problem of finding the optimal mean-CVaR-portfolio can be linearized which makes the CVaR more applicable in portfolio optimization than the VaR. The exact framework will be discussed in section 3.2.

Advantages of CVaR

- Applicable to asymmetric fat-tailed distributions
- Coherent risk measure
- Convex and smooth
- Can be linearized
- More conservative than VaR

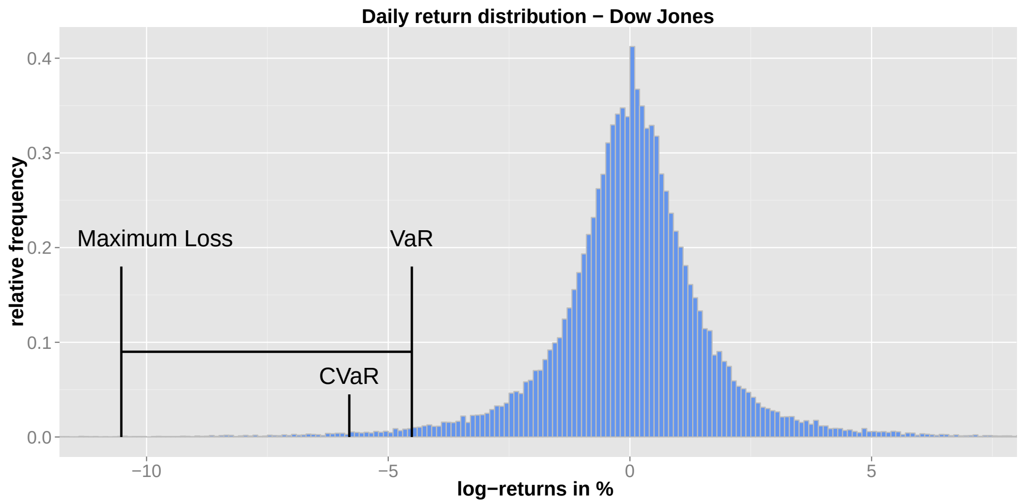

Figure 6: Visualization of the VaR and CVaR at α =1%

2.4 Diversification

In portfolio theory, diversification reduces risk and adds protection against extreme events by reducing the dependence on particular assets. In the case that an asset generates an extreme loss, a more diversified portfolio experiences a smaller loss than a portfolio that is highly invested in this asset. This effect is said to diminish with an increasing number of assets [11], the opinions on the optimal number of assets for a diversified portfolio go from 10 to 30 assets.

While diversification is an intuitive concept, there is no unique quantitative measure but a wide variety of established diversification measures [18]. Three of them are presented here:

Herfindahl index

This index uses the squared absolute value of the weight vector [18]:

$$\mathrm{Div}_H(\bar{x}) := 1 - \sum_{i=1}^{N} x_i^2 = 1 - \bar{x}^T\ \mathbf{I}\ \bar{x}^T \tag{24}$$

The Herfindahl index does not include any correlation between the assets. Note that the calculation of the Herfindahl index of a portfolio is as easy as the calculation of its variance risk and is thus convenient for portfolio optimization techniques. It is a quadratic and therefore **strictly concave** function of portfolio weights.

Mean deviation index

Instead of using the squared portfolio weights, one can measure the deviation from the *Equal-weights-portfolio* (EWP), the portfolio that is invested with $1/N$ in every available asset:

$$\mathrm{Div}_{MDI}(\bar{x}) = 1 - \sum_{i=1}^{N} |x_i - \frac{1}{N}| \tag{25}$$

If we want to maximize the diversification of our portfolio using this risk measure, the task can be linearized: We introduce another variable vector $\bar{y}$ with the same length as $\bar{x}$. We then introduce the constraints

$$\begin{aligned} \forall i \in \{1, .., N\} :& \\ y_i \geq x_i - \frac{1}{N}& \\ y_i \leq \frac{1}{N} - x_i& \\ y_i \geq 0& \end{aligned} \tag{26}$$

and formulate the problem as

$$\max_{\bar{x}, \mathbf{y}} \quad - \sum_{i=1}^{N} y_i \tag{27}$$

Note that this diversification measure is a **linear function** of portfolio weights.

Diversifying dependence

Another approach to diversification is through measuring mutual asset dependencies. If we increase the number of invested assets, we do not necessarily decrease the risk of extreme events. Adding an asset to a portfolio that is highly correlated to an invested asset bears the risk that both assets simultaneously loose value. Therefore instead of just spreading the portfolio's weights $\bar{x}$, one could try to reduce the overall dependencies inside the portfolio.
There are multiple ways to quantify the dependence between two assets. The simplest one is the linear Pearson correlation that is related to the covariance in equation 8. For two random variables X_i, X_j, the correlation is defined as

$$\rho_{i,j} = \rho(X_i, X_j) = \frac{\sigma(X_i, X_j)}{\sqrt{\sigma(X_i)\sigma(X_j)}} \tag{28}$$

[12]. Note that $\rho_{i,j} \in [-1, 1]$.
A diversifaction measure including the correlation of the assets can then be defined by simply replacing the identity matrix in 24 with the correlation matrix:

$$\mathrm{Div}_\rho(\bar{x}) = 1 - \bar{x}^T \boldsymbol{\rho} \bar{x} \tag{29}$$

where $\boldsymbol{\rho} = \left(\rho_{i,j}\right)$. In the case that all assets are completely uncorrelated, equation 29 reduces to equation 24. Similar to the covariance, Div_ρ is **strictly concave modulo translation**.
The linear Pearson correlation faces the same difficulties as the mean covariance risk measure (see section 2.2) which will not be repeated here. A more target-aimed way for dependence diversification is to look at *tail dependencies* between two assets.
Lower tail dependency is a measure to quantify the probability of a simultaneous loss of two assets. If one asset experiences an extreme loss, the tail dependency is the probability that the other asset experiences one as well. Figure 7 shows two random variables with extremly high lower tail dependency.

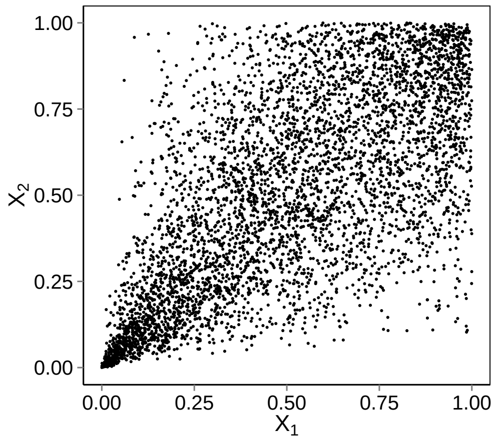

Figure 7: A sample of 10 000 points drawn from a tail-dependent Clayton copula.

Definition 4. *The lower tail dependency λ between two random variables X_i, X_j is defined as the following conditional probability:*

$$\lambda_{i,j} = \lim_{q \nearrow 0} P(X_i \leq F_i^{-1}(q) | X_j \leq F_j^{-1}(q)) \tag{30}$$

where $F_i(x)$ is the cumulative distribution of X_i [12].

Like the correlation, $\lambda_{i,j} \in [0, 1]$.
We can modify the diversification measure 29 by replacing $\boldsymbol{\rho}$ with $\boldsymbol{\lambda} = \left(\lambda_{i,j}\right)$:

$$\mathrm{Div}_\lambda(\bar{x}) = 1 - \bar{x}^T \boldsymbol{\lambda} \bar{x} \tag{31}$$

Note that the estimation of $\lambda_{i,j}$ is more robust than for $\rho_{i,j}$ since it does not rely on the second moment of the asset distributions (similar to the fact that the estimation of the CVaR is more robust than the estimation of the covariance). However the estimation of the tail dependence is not a simple task and relies on the theory of copulae. More information can be found in *Quantitative risk management* by Embrechts et al. [12].
Like the correlation matrix, $\lambda_{i,j}$ is positive semi-definite and thus $\mathrm{Div}_\lambda(\bar{x})$ convex. $\lambda_{i,j}$ is furthermore positive definite and thus $\mathrm{Div}_\lambda(\bar{x})$ is **strictly convex** if

$$\forall i \in \{1, .., N\} : \sum_{k \neq i} \lambda_{i,k} < 1, \tag{32}$$

i.e. if no asset is completely tail dependent on a linear combination of other assets. This condition is fulfilled if the conditions of theorem 1 apply to the given set of assets.

3 Multicriteria Optimization

Multicriteria optimization, also called multi-objective optimization, is a field in optimization sciences that deals with problems having more than one objective function. In other words, a problem having multiple objectives that are not necessarily compatible is a multicriteria problem. Multicriteria optimization is an important topic for investors as the problem of finding a portfolio that simultaneously maximizes returns and minimizes risk is a multicriteria problem. In this chapter we will discuss different ways of approaching the problem.

During this chapter we will work with the definitions of portfolio weights given in chapter 2.1. All definitions will be formulated in such a way that they directly apply to the portfolio framework.

3.1 Definitions

If there are N different assets available for investment and we forbid short selling, then the set of all possible portfolios is given by

$$\mathcal{X} = \{\bar{x} \ : \ \sum_{i}^{N} x_i = 1 \quad , \quad x_i \geq 0\} \tag{33}$$

This set is called the *feasible set*.

To each portfolio certain characteristics such as its risk, its expected return or its diversification can be assigned. A set of k particular characteristics of a portfolio can be calculated as a function of the portfolio's weights $\bar{f}(\bar{x}) = (f_1(\bar{x}), .., f_k(\bar{x})^T$.

If we want this feasible set visualized by quantities like the risk and the expected return, we use the feasible set in objective space:

$$\mathcal{Y} := \bar{f}(\mathcal{X}) \tag{34}$$

where $\bar{f}(\bar{x})$ is the multi-dimensional function to calculate the desired quantities. If we want the feasible set in the risk-return space, $\bar{f}$ would be given by $\bar{f}(\bar{x}) := (\text{Risk}(\bar{x}), \bar{x} \cdot \bar{\mu})^T$. Figure 8 depicts the feasible set in the CVaR-return space both for the Dow Jones index and for the LPP2005.

Normally an investor desires to optimize more than one quality of his portfolio. Mathematically, this is a so called multi-objective problem and can be written as

$$\min_{\bar{x} \in \mathcal{X}} \ \big(f_1(\bar{x}), ..., f_k(\bar{x})\big) \tag{35}$$

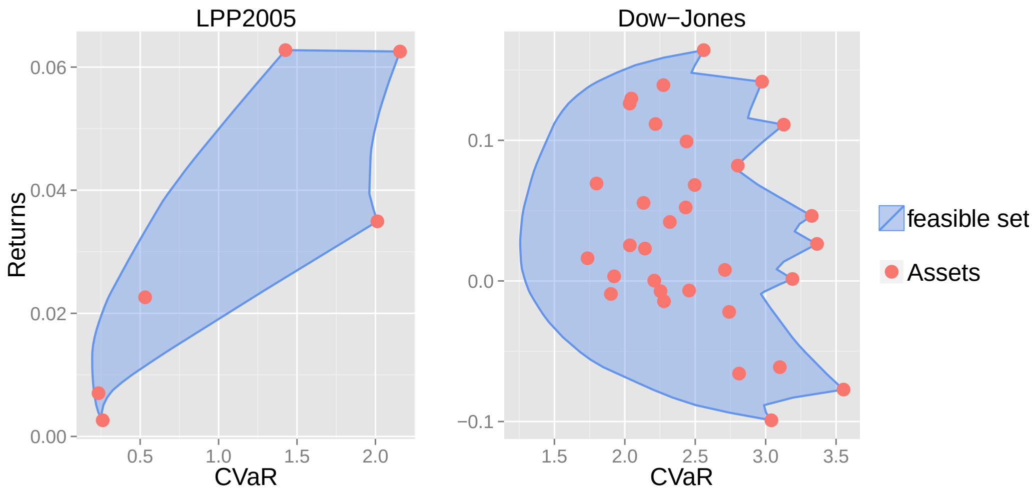

Figure 8: Visualization of the feasible set $\mathcal{Y}$ in the Risk-Return-space and the position of the available assets. The chosen risk measure is the CVaR.

where k is the number of objectives. For example, the objective functions to optimize the portfolio problem is given by

$$\bar{f}(\bar{x}) = \begin{pmatrix} \mathrm{CVaR}_\alpha(\bar{x}) \\ -\bar{x} \cdot \bar{\mu} \end{pmatrix} \tag{36}$$

In general, there exists no portfolio $\bar{x}$ that optimizes all objective functions. This can also be seen when looking at the feasible sets in figure 8: The portfolio that maximizes the return does not simultaneously minimize the portfolio's expected shortfall. Therefore other criteria to define the "*optimal*" portfolio have to be found:

Definition 5. *A portfolio $\bar{x}$ is called Pareto optimal [9] or Pareto efficient for $\bar{f} = (f_1, f_2, ..., f_n)$ if there exists* ***no*** *other portfolio $\bar{x}' \in \mathcal{X}$ with*

$$\begin{aligned} \exists i \in \{1, 2, .., n\}: \quad & f_i(\bar{x}') < f_i(\bar{x}) \\ \forall j \in \{1, .., i-1, i+1, .., n\}: \quad & f_j(\bar{x}') \leq f_j(\bar{x}) \end{aligned} \tag{37}$$

or in other words, no portfolio that can achieve better results than $\bar{x}$ for one quantity while not performing lower for any of the other quantities.
The set of all Pareto efficient portfolios will be denoted with $\mathcal{X}_E$.

Definition 6. *A portfolio $\bar{x}$ is called weakly Pareto optimal [9] or weakly Pareto efficient for $f = (f_1, f_2, ..., f_n)$ if there exists* ***no*** *other portfolio $\bar{x}' \in \mathcal{X}$*

with

$$\forall i \in \{1, 2, .., n\}: \quad f_i(\bar{x}') < f_i(\bar{x}) \tag{38}$$

or in other words, no portfolio that can achieve better results than $\bar{x}$ for every quantity.
The set of all weakly Pareto efficient portfolios will be denoted with $\mathcal{X}_{wE}$.

Naturally, every portfolio that is considered for investment should be Pareto-optimal under the chosen objective function.

3.2 Traditional Portfolio Theory

The traditional goal of portfolio optimization is to find the combination of assets such that an investment minimizes a given risk measure while maintaining a certain expectation of return. In order to achieve this, investors rely on the diversification of the investment and the analysis of posterior asset returns.
We will first to introduce the portfolio optimization framework for general risk measures. Then we refine this framework by inserting the CVaR as a general representative for scenario-based risk measures. Keep in mind that the CVaR can be replaced with any scenario-based risk measure throughout this thesis.
In the following we will assume two things:

1. The assets fulfill the conditions of theorem 1, i.e. the set of assets cannot be perfectly correlated on any non-zero subset of the probability sample space Ω.

2. There are no two assets with exactly the same expected return.

These assumptions will become useful when talking about the efficiency of our portfolios and in the later discussion of optimization approaches: Assumption 1 will guarantee us strictly convex risk measures. Assumption 2 guarantees the uniqueness of the portfolio with the maximum expected return.
These assumptions are not unreasonable: Even though most portfolio theories assume it, information exhange is not perfect in the financial market for all participants [11]. It is therefore unlikely to observe a perfect correlation between multiple assets for an observable set of scenarios.
Assumption 2 can always be justified by including more digits into calculation.

The most simple way of approaching the portfolio problem is to neglect one objective and only minimize the other one. For example, an investor might only be interested in finding the least risky portfolio. This problem could be formulated as following:

$$\begin{aligned} \min_{\bar{x}} \ \text{Risk}(\bar{x}) \\ s.t. \\ \sum_i x_i = 1 \\ x_i \geq 0 \end{aligned} \tag{39}$$

The constraints guarantee that the portfolio obeys our framework introduced in chapter 2.1. The solution of equation 39 is called the global minimum risk portfolio. Naturally the portfolio depends on the choice of the risk measure. We will refer to the global minimum CVaR portfolio as MCVP. Because of assumption 1 and theorem 1, our risk measure is strictly convex. Hence the solution is unique as a strictly convex function has a unique minimum: If there were two solutions $\bar{x}_1$ and $\bar{x}_2$, a portfolio consisting of a linear combination of x_1 and x_2 would have a lower variance than x_1 and x_2 as x_1 and x_2 are not perfectly correlated. Solutions of equation 39 are not necessarily unique for the VaR risk measure as the VaR is not strictly convex.

An investor could also only be interested in maximizing the return of the portfolio. For that the formula for the expected return in equation 7 can be used. A formulation of the problem looks like this:

$$\begin{aligned} \max_{\bar{x}} \ \bar{x}^T \bar{\mu} \\ s.t. \\ \sum_i x_i = 1 \\ x_i \geq 0 \end{aligned} \tag{40}$$

The solution of equation 40 is called the global maximum return portfolio (MRP) and consists just the asset with the highest expected return. As we assumed that there are no two assets with the same expected return, this solution is unique as well.

We showed that the solutions of equation 39 and 40 are unique and minimize one of the two objective functions of the portfolio problem. Consequently they are Pareto optimal.

The above discussed approaches are two extreme cases. Normally investors are interested in a combination of both cases, i.e. a portfolio that has an appropriate expected return while not being too risky. Naturally, these portfolios should be Pareto optimal.
As already mentioned, this is a multi-objective problem. The most common way of solving this problem (which was also introduced my Markowitz) is to use the ϵ-constraint method. In this method, only one objective function f_k is optimized while the other objective functions are constrained:

$$\begin{aligned} \min_{\bar{x}} f_k(\bar{x}) \\ s.t.\ f_i(\bar{x}) \leq \epsilon_i\ ,\quad i \neq k \\ \bar{x} \in \mathcal{X} \end{aligned} \tag{41}$$

Ehrgott [9] proved that every unique solution $\mathbf{x}$ of equation 41 is in $\mathcal{X}_E$.
There are two possible ways of formulating the portfolio problem with this method [22]:
Firstly, one tries to minimize the overall risk of the portfolio while keeping the expected return above a constant value. This value is the so called *target return*. When requiring full investment and forbidding short selling and using equation 7, this problem can be written as

$$\begin{aligned} \min_{\bar{x}}\ \text{Risk}(\bar{x}) \\ s.t. \\ \bar{x}^T\overline{\mu} \geq r \\ \sum_i x_i = 1 \\ x_i \geq 0 \end{aligned} \tag{42}$$

Similarly to equation 39, every solution of equation 42 is unique for the CVaR and the variance risk. If there were two solutions $\bar{x}_1$ and $\bar{x}_2$ of equation 42, every linear combination of $\bar{x}_1$ and $\bar{x}_2$ would have a lower risk than $\bar{x}_1$ and $\bar{x}_1$ and still fulfill the constraint. Consequently, every solution of 42 is unique and is thus after Ehrgott in $\mathcal{X}_E$.
The Pareto optimal portfolios are distributed between two points (the maximum return portfolio and the minimum CVaR portfolio), the $\mathcal{X}_E$ forms a line and is 2-dimensional. This line is also called the *efficient frontier* and can be seen in figure 9 for the CVaR risk measure.
As the efficient frontier spans between the MCVP and the MRP and every portfolio on it is unique, we can find every Pareto optimal portfolio simply by

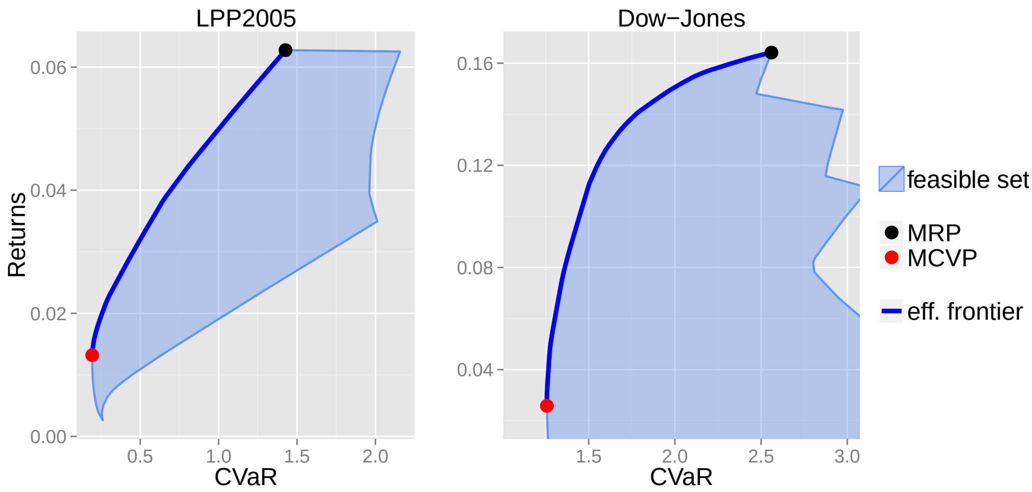

Figure 9: Visualization of the set of risk-return Pareto optimal portfolios and maximum-return portfolio (MRP) as well as the minimum-CVaR portfolio (MCVP.)

chosing any $r \in [\min(\bar{\mu}), \max(\bar{\mu})]$. Note that in general it is not that simple to find every $\bar{x} \in \mathcal{X}_E$. for equation 41 [9].
The second approach is to maximizes the expectation of the return while limiting the risk:

$$\begin{aligned} \max_{\bar{x}} \ \bar{x}^T\overline{\mu} \\ s.t. \\ \text{Risk}(\bar{x}) \leq r \\ \sum_i x_i = 1 \\ x_i \geq 0 \end{aligned} \tag{43}$$

Similar to equation 42, the set of solutions of equation 43 is equal to $\mathcal{X}_E$.

3.3 Mean-CVaR-Portfolio

The implementation of a risk measure into equation 42 and 43 appears to be straight forward. In principle, one could just substitute the Risk($\bar{x}$)-term with an expression for the CVaR and choose the desired quantile. Equation 42 could be reformulated as follows:

$$
\begin{aligned}
\min_{\bar{x}} \ \mathrm{CVaR}_{\alpha}(\bar{x}) \\
s.t. \\
\bar{x}^T \overline{\mu} \geq r \\
\sum_i x_i = 1 \\
x_i \geq 0
\end{aligned}
\tag{44}
$$

However, the expression for $\mathrm{CVaR}_{\alpha}(\bar{x})$ in equation 22 is non-linear and thus it is very hard to solve $\min_{\bar{x}} \ \mathrm{CVaR}_{\alpha}(\bar{x})$ in computational optimization. However there exists a way to linearize the problem of equation 44. In the following the main steps are presented. For a more detailed discussion, the reader is advised to read *Optimization of conditional value-at-risk* by R. Rockafellar and S. Uryasev [24].

Reformulation
We define $\bar{r}$ as the vector of all asset returns r_i. The return function $f(\bar{x}, \bar{r})$ of a portfolio with weights $\bar{x}$ is then given by

$$
f(\bar{x}, r) = \bar{x}^T \bar{r} \tag{45}
$$

If every r_i is a random variable, so is $\bar{r}$. We define $\rho(\bar{r})$ to be the distribution function of $\bar{r}$. With this framework, the expected return of portfolio $\bar{x}$ is then given by

$$
E(R_{\bar{x}}) = \int_{\mathbb{R}} \bar{x}^T \bar{r} \rho(\bar{r}) d\bar{r}
$$

With this formulation we can rewrite equation 22 as

$$
\mathrm{CVaR}_{\alpha}(\bar{x}) = -\frac{1}{\alpha} \int_{\bar{x}^T \bar{r} \leq -\mathrm{VaR}_{\alpha}(\bar{x})} \bar{x}^T \bar{r} \rho(\bar{r}) d\bar{r} \tag{46}
$$

This formulation is also more in the sense of the definition of a scenario-based risk measure (see equation 20).
Equation 46 is dependent on $\mathrm{VaR}_{\alpha}(\bar{x})$ and thus difficult to optimize. Computing the CVaR requires a computation of the VaR first and thus finding the minimum of the CVaR is very difficult for industrial solvers. Rockafeller

and Uryasev [24] showed that 46 can be rewritten as a function that is independent of the VaR-function. For that we define the new function

$$F_\alpha(\bar{x}, \beta) = \beta - \frac{1}{\alpha} \int_{\mathbb{R}} [-\bar{x}^T \bar{r} - \beta]^+ \rho(\bar{r}) d\bar{r} \tag{47}$$

where

$$[t]^+ = \begin{cases} t & \text{when} \quad t \geq 0 \\ 0 & \text{when} \quad t \leq 0 \end{cases} \tag{48}$$

Rockafeller and Uryasev then proved that $\mathrm{CVaR}_\alpha(\bar{x})$ can be calculated as the minimum of $F_\alpha(\bar{x}, \beta)$ in β:

$$\mathrm{CVaR}_\alpha(\bar{x}) = \min_{\beta \in \mathbb{R}} F_\alpha(\bar{x}, \beta) \tag{49}$$

which is now independent of VaR_α.

Discretization

Normally we do not know the exact form of $\rho(\bar{r})$, therefore we have to use a sample estimator. The best unbiased estimator for a an integral in the form

$$\int_{\mathbb{R}} f(\bar{r}) \rho(\bar{r}) d\bar{r}$$

is

$$\frac{1}{S} \sum_{i=1}^{S} f(\bar{r}_i) \tag{50}$$

where $\bar{r}_1, .., \bar{r}_S$ are the available sample points. Therefore we rewrite our function $F_\alpha(\bar{x}, \bar{r})$ into the estimated function

$$\widetilde{F}_\alpha(\bar{x}, \beta) = \beta - \frac{1}{S \cdot \alpha} \sum_{i=1}^{S} [-\bar{x}^T \bar{r}_i - \beta]^+ \tag{51}$$

We now have

$$\mathrm{CVaR}_\alpha(\bar{x}) = \min_{\beta \in \mathbb{R}} \ \beta - \frac{1}{S \cdot \alpha} \sum_{i=1}^{S} [-\bar{x}^T \bar{r}_i - \beta]^+ \tag{52}$$

Linearization
Optimizing equation 52 is still hard to solve numerically as it is non-linear due to the expression $[-\bar{x}^T \mathbf{r}_i - \beta]^+$. Fortunately it is possible to transform this expression into a linear form. For that we introduce the variable $\bar{u} \in \mathbb{R}^S$ which is subject to the constraints

$$\begin{aligned} u_k &\geq 0 \\ u_k &\geq -\bar{x}^T \bar{r}_k - \beta \end{aligned} \tag{53}$$

The constraints ensure that

$$\min_{u_k} \; u_k = [-\bar{x}^T \bar{r}_i - \beta]^+$$

We then rewrite the minimization problem in equation 52 as

$$\mathrm{CVaR}_\alpha(\bar{x}) = \min_{\beta,\mathbf{u}} \beta + \frac{1}{S \cdot \alpha} \sum_{i=1}^{S} u_k \tag{54}$$

We are finally able to write a linear form of equation 44:

$$\begin{aligned} \min_{\bar{x},\beta,\bar{u} \in \mathbb{R}^S} \; \beta + \frac{1}{S \cdot \alpha} \sum_{i=1}^{S} u_k \\ s.t. \\ \bar{x}^T \overline{\mu} \geq r \\ \sum_i x_i = 1 \\ x_i \geq 0 \\ u_k \geq 0 \\ u_k \geq -\bar{x}^T \bar{r}_k - \beta \end{aligned} \tag{55}$$

Krokhmal, Palmquist and Uryasev [19] showed that this framework can also be used for designing CVaR-constrainst. In detail, they showed the following identity:

$$\begin{array}{ccc} \max_{\bar{x}} \; \bar{x}^T \bar{\mu} & = & \max_{\bar{x},\beta,\bar{u}} \; \bar{x}^T \bar{\mu} \\ s.t. & & s.t. \\ \widetilde{F}_\alpha(\bar{x}, \beta) \leq \delta & & CVaR_\alpha(\bar{x}) \leq \delta \end{array} \tag{56}$$

We can thus formulate equation 43 with CVaR-risk measure as well:

$$
\begin{aligned}
\max_{\bar{x},\beta,\bar{u}\in\mathbb{R}^S} \quad & \bar{x}^T\bar{\mu} \\
s.t. & \\
\beta + \frac{1}{S\cdot\alpha}\sum_{i=1}^{S} u_k &\leq r \\
\sum_{i} x_i &= 1 \\
x_i &\geq 0 \\
u_k &\geq 0 \\
u_k &\geq -\bar{x}^T\bar{r}_k - \beta
\end{aligned}
\tag{57}
$$

In the following we will always choose $\alpha = 0.05$ when computing CVaR_α.

4 Weighted Sum Method

As described in the introduction, we will introduce a method to extend the bi-criteria portfolio problem into a tri-criteria one. However the ϵ-constraint method faces problems when optimizing more than two objectives. To be able to achieve this, we will now introduce a different approach to solve multi-criteria optimization problems: The *weighted sum method*.
We will first describe and analyze this method and apply it to the bi-criteria portfolio problem. Afterwards we will use this method to extend the problems into a tri-criteria one and discuss its benefits and flaws.

4.1 Definitions

The weighted sum method is the simplest and best known solution method to multi-objective problems [9]. A multi-objective problem of the form

$$\min_{\bar{x}\in\mathcal{X}} \ \big(f_1(\bar{x}), ..., f_k(\bar{x})\big) \tag{58}$$

can be solved as a single objective problem of the form

$$\min_{\bar{x}} \ \sum_{i=1}^{k} \lambda_i f_i(\bar{x}) \tag{59}$$

with $\lambda_i \in [0,1]$. The value of λ_i represents the emphasis on the i'th objective.

We assume that $\exists i \in \{1, .., k\} : \ \lambda_i > 0$, otherwise problem 59 is not well defined. We write this condition as $\bar{\lambda} \geq 0$. Furthermore, if $\forall i \in \{1, ..., k\} : \ \lambda_i > 0$, we write $\bar{\lambda} > 0$.
Obviously, if $\bar{\lambda}$ is of the form

$$\bar{\lambda} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{60}$$

equation 59 reduces to a problem of the form of equation 39.

It is common practice to constrain the choice of $\bar{\lambda}$ such that

$$\sum_{i=1}^{n} \lambda_i = 1 \tag{61}$$

This makes sense since the set of Pareto optimal solutions is at most a $k-1$ dimensional manifold but $\bar{\lambda}$ is k-dimensional. Consequently there are many values for $\bar{\lambda}$ for which equation 59 is equivalent. For instance, the problem is the same if $\forall i : \lambda_i = 1$ or if $\forall i : \lambda_i = 0.5$. For every $\bar{\lambda}$ and $\hat{\bar{\lambda}}$ that are linearly dependent $\hat{\bar{\lambda}} = \alpha\bar{\lambda}$, the problem is equivalent.
In general if $\sum_{i=1}^{n} \lambda_i \ll 1$ the solutions are not likely to be numerically stable [15]. Furthermore the objective functions should be normalized such that their extrema lie in the same value range.
We are now interested in the efficiency of solutions of equation 59. Ehrgott [9] proved the following theorems:

Theorem 2. *Assume that $\hat{\bar{x}}$ is an optimal solution for a problem of the form of equation 59 with $\lambda_i \in [0,1]$. Then the following is true:*

1. *If $\bar{\lambda} > 0$, then $\hat{\bar{x}} \in \mathcal{X}_E$*
2. *If $\bar{\lambda} \geq 0$, then $\hat{\bar{x}} \in \mathcal{X}_{wE}$*
3. *If $\bar{\lambda} \geq 0$ and $\hat{\bar{x}}$ is a unique optimal solution, then $\hat{\bar{x}} \in \mathcal{X}_E$*

Theorem 3. *Assume that $\mathcal{X}$ is a convex set and $\forall i \in \{1, ..., k\}$, f_i are convex functions. Then the following is true:*

1. *If $\hat{\bar{x}} \in \mathcal{X}_E$, then $\exists \bar{\lambda} \geq 0$ such that $\hat{\bar{x}}$ is a solution of the problem 59 with $\bar{\lambda}$.*
2. *Assume further that there at least one function f_i is **strictly** convex. If $\bar{\lambda} \geq 0$ and $\lambda_i \neq 0$, then there is a unique solution $\hat{\bar{x}} \in \mathcal{X}_E$ for 59[7].*

A direct consequence of theorem 2 and 3 is the following statement:

Assume that $\mathcal{X}$ is a convex set and that $f_i,\ i \in \{1, ..., k\}$ are **strictly** convex functions. Then every solution $\hat{\bar{x}}$ of 59 is unique and the set of solutions of is equal to $\mathcal{X}_E$.

[7]The second statement of theorem 3 was not actually proven by Ehrgott in *Multicriteria optimization*, but is a consequence of the fact that every strictly convex function has a unique minimum and the fact that the sum of a convex function and a strictly convex function is strictly convex.

These theorems will be very important to us in the following.
Please note that if $f_1, .., f_k$ are convex but not strictly convex (i.e. linear), every Pareto optimal $\bar{x}$ is a solution of 59 for some $\bar{\lambda}$, but most $\bar{x} \in \mathcal{X}_E$ are not unique solutions for this $\bar{\lambda}$. This can be seen in the following example:
Let the interval $[0, 1]$ be our feasible set $\mathcal{X}$. Assume we have two objectives and our linear objective functions are given by

$$\begin{aligned} f_1(x) &:= x \\ f_2(x) &:= -(x-1) \end{aligned} \tag{62}$$

The minimum of f_1 is $x = 0$, the minimum of f_2 is $x = 1$. Obviously the whole interval $[0, 1]$ consists of Pareto optimal solutions for our problem. We write $\lambda_1 =\in [0, 1]$, $\lambda_2 = (1 - \lambda_1)$. The sum of equation 59 is then given by

$$g(x, \lambda) = \sum_{i=1}^{2} \lambda_i f_i(x) = x(2\lambda_1 - 1) \tag{63}$$

This function has its global minimum for $\lambda \in [0, 0.5)$ at $x = 1$, for $\lambda \in (0.5, 1]$ at $x = 0$. For $\lambda = 0.5$, $g(x, 0.5) = 0$ and thus every value in the interval is an optimal solution to the problem

$$\min_{x\in[0,1]} g(x, \lambda = 0.5) = [0, 1] \tag{64}$$

For computational solvers, this situation is normally very difficult and the solver will typically pick one of the possible values instead of returning the whole interval. Consequently the whole set of Pareto optimal solutions will not be found completely. The situation is different if we deal with strictly convex functions, for example

$$\begin{aligned} f_1(x) &:= x^2 \\ f_2(x) &:= (x-1)^2 \\ g(x, \lambda) &= x^2 + 2\lambda x - 2x - (\lambda - 1) \end{aligned} \tag{65}$$

where for every λ the solution is unique. A comparison is depicted in figure 10.
A disadvantage of the weighted sum method is the fact that different $\bar{\lambda}$ can yield the same solution, even if the objective functions are strictly convex. Additionally, a uniformly distributed set of $\bar{\lambda}$ values is no guarantee for a uniformly distributed set of solutions. This is also true for strictly convex objective functions. A rule of thumb is that the solutions are distributed better across $\mathcal{X}_E$ if the objective functions are nowhere close to being linear. Also in both cases it helps if the objective functions are normalized such that their images $f_i(\mathcal{X}_E)$ fill the same interval.

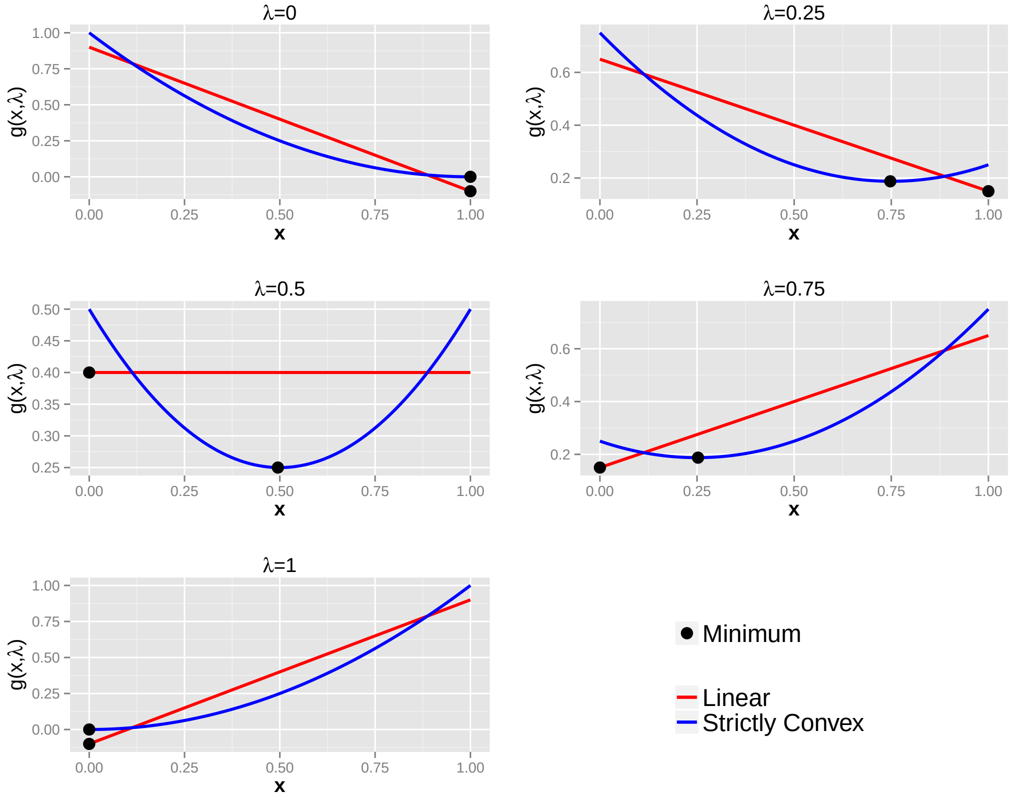

Figure 10: A comparison of the effectiveness of the weighted sum method for linear (red) and strictly convex (blue) objective functions. Note that the linear functions only gives two different minima while the strictly convex functions gives the whole interval.

4.2 Critical Line Algorithm

We will now look at the implementation of the weighted sum method to solve the Portfolio problem.
In 1956 Markowitz introduced an early version of this method. We will look at it a bit closer as it is crucial for the next chapter.
A way to find the portfolio weights on the efficient frontier is to introduce the variable $\lambda \in [0, 1]$ and formulate the optimization problem as

$$\min_{\bar{x} \in \mathcal{X}} \quad -\lambda \text{Return}(\bar{x}) + (1 - \lambda)\text{Risk}(\bar{x})$$

λ represents the emphasis of the investor on the particular criteria: If $\lambda = 0$, the interest lies only on minimizing the risk, if $\lambda = 1$ the investor is only

interested in maximizing the expected return.
This is a version of the weighted sum method. Our objective functions are given by

$$\begin{aligned} f_1 = -\text{Return}(\bar{x}) = -\bar{x}^T \bar{\mu} \\ f_2 = \text{Risk}(\bar{x}) \end{aligned} \tag{66}$$

As mentioned above, the expected return is a linear function of the portfolio weights whereas the risk is for scenario-based risk measures a strictly convex risk measure (as we assumed that there are no assets that are perfectly correlated for a non-zero set of scenarios).
As a consequence of theorem 3, every Pareto optimal portfolio $\bar{x} \in \mathcal{X}_E$ can be found by this method. Furthermore, we can apply the second statement of theorem 3 for every $\lambda < 1$. Thus the portfolio solutions $\bar{x}$ for $\lambda < 1$ are unique and thus in $\mathcal{X}_E$.
If $\lambda = 1$, the portfolio solution is the maximum return portfolio MRP. We already showed that this portfolio is a unique solution of equation 40 and thus also of equation 66.
We have thus shown that every Pareto optimal portfolio $\bar{x}$ is a unique solution for some $\lambda \in [0, 1]$.
The implementation of this algorithm for the CVaR-risk measure looks like this:

$$\max_{\bar{x},\mathbf{e},\beta} \quad \lambda_1 \cdot \bar{x}^T \bar{\mu} - (1 - \lambda) \cdot (\beta - \frac{1}{\alpha s} \sum_{i=0}^{s} e_i) \tag{67}$$

$$s.t. \tag{68}$$

$$\sum_i x_i = 1 \tag{69}$$

$$x_i \geq 0. \tag{70}$$

$$e_i \geq \beta - \sum_{j=0}^{s} x_i r_{i,j} \tag{71}$$

At this moment, the risk contribution of equation 67 is not necessarily in the same order of magnitude as the risk contribution, i.e. they are not normalized to the same order of magnitude. This can lead to undesired behaviour: If the objective function f_i takes values in a much greater range than the other objectives, the sum of the objectives will take its minimum very close to the minimum of f_i. As a consequence, the resulting portfolios will not be distributed uniformly along the efficient frontier. To prevent this, we will

normalize the CVaR and the return term so that both can take values between 0 and 1.
The most extreme points on the efficient frontier for both the CVaR and the return are the MRP and the MCVP. We thus use their coordinates in CVaR-Return plane for the normalization:
We then rewrite the return term as

$$\frac{\bar{x}^T \cdot \bar{\mu} - r_{MCVP}}{r_{MRP} - r_{MCVP}} \tag{72}$$

where

$$r_{MCVP} = \bar{x}^T_{MCVP} \cdot \bar{\mu}$$
$$r_{MRP} = \bar{x}^T_{MRP} \cdot \bar{\mu}$$

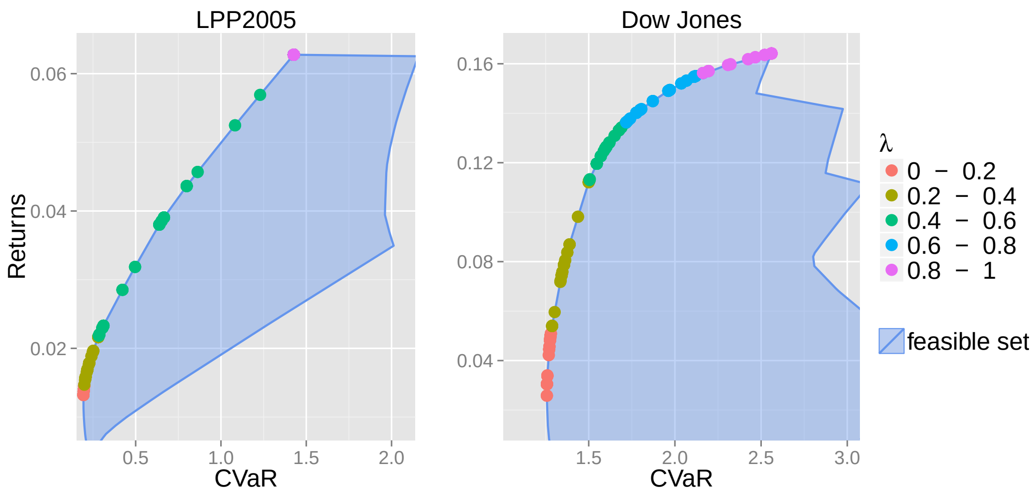

Figure 11: Distribution of the portfolio solutions of equation 74 for different λ values.

We normalize the CVaR-term in the same way:

$$\frac{1}{c_{MRP} - c_{MCVP}}(\mathrm{VaR} - \frac{1}{\alpha s}\sum_{i=0}^{s} e_i - c_{MCVP}) \tag{73}$$

where c_{MCVP} is the CVaR-value for the MCVP and c_{MRP} is the CVaR-value for the MRP. The problem can now be written as

$$
\begin{aligned}
\max_{\bar{x},\mathbf{e},\beta} \quad \lambda_1 \cdot \frac{\bar{x}^T \cdot \bar{\mu} - r_{MCVP}}{r_{MRP} - r_{MCVP}} \\
- \frac{(1-\lambda)}{c_{MRP} - c_{MCVP}} (\text{VaR} - \frac{1}{\alpha s} \sum_{i=0}^{s} e_i - c_{MCVP}) \\
s.t. \\
\sum_i x_i = 1 \\
x_i \geq 0. \\
e_i \geq \text{VaR} - \sum_{j=0}^{s} x_i r_{i,j}
\end{aligned}
\tag{74}
$$

A visualization of the portfolio distribution of the unnormalized critical line algorithm can be found in figure 25.

The critical line algorithm in principle parametrizes the efficient frontier. This parametrization can take very unpleasant forms depending on the shape of the efficient frontier. The behaviour for different λ-values is depicted in figure 11.

As one can see in figure 11, for the **LPP2005** the largest part of the efficient frontier is filled by the small range $\lambda \in (0.4, 0.6)$. Also notice the concentration of portfolios on the MRP and the MCVP: For all $\lambda > 0.576$, the portfolio obtained from the critical line algorithm will be the MRP.

The situation is not as dramatic for the efficient frontier of the **Dow Jones index**: The portfolios are much more spread along the efficient frontier and there is not such a high concentration on the MRP and the MCVP. The MRP is hit for $\lambda > 0.903$.

The reason for this difference lies in the different shapes of the efficient frontier: For the Dow Jones index the frontier is less close to a linear function whereas a large part of the frontier of the LPP2005 is almost linear. As explained above, the weighted sum method only finds every Pareto optimal portfolio if the objective functions are not linear. This is due to the fact that for linear objective functions, almost all portfolio solutions are only solutions to the same $\bar{\lambda}$. If the objective functions are deviated from being linear into being strictly convex, this single $\bar{\lambda}$ for which these portfolios are a solution is then stretched to a set of $\bar{\lambda}$ values. Still if the objective functions are very close to being linear, this set of $\bar{\lambda}$ values is very small. The portfolio solutions are therefore distributed more uniformly if the objective functions are strictly convex and not close to being linear (i.e. rounder).

As already mentioned in the introduction, the LPP2005 is an index illustrating the limitations of the weighted sum method. Figure 12 shows that the hull of the feasible set of the LPP2005 is more stretched and less round than the hulls of the Dow-Jones index or the DAX for comparison (see figure 12 and figure 27). Consequently the efficient frontier is closer to a linear curve. This phenomenon can be explained due to the small number of assets in the LPP2005. Larger indices tend to have a rounder hull and thus a more curvy efficient frontier (see the hull of the Dow-Jones index in figure 13 and the hull of the DAX index in figure 27).

4.3 Three Goal Programming

Diversification has proven to be very efficient at reducing risk from individual assets. As the CVaR is a *a posteriori* risk measure, i.e. it only considers past returns, it might not capture the future risk of an asset fully. An asset might have generated high and stable returns in the past, but due to structural changes this behaviour might change drastically in the future. Diversification can help reduce the dependence on such assets.

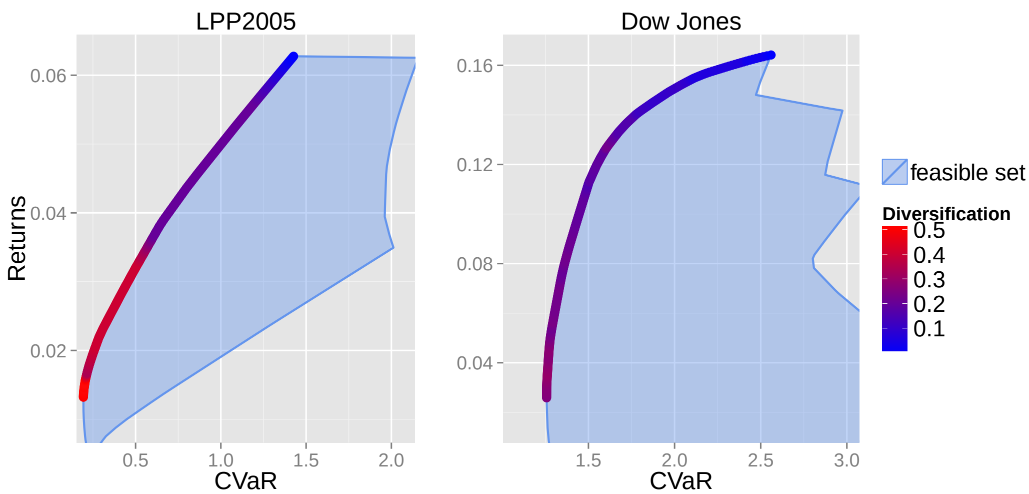

Figure 12: Visualization of the diversification along the efficient frontier

In figure 12 the diversification of the portfolios on the efficient frontier is depicted. As it is visible, the portfolios along the critical line are not very diversified. The MRP is a single asset and the MCVP has a high concentration risk as well. Regarding the above considerations, the investor might not be satisfied by these portfolios, i.e. the Pareto efficient solutions from the

bi-objective CVaR-return problem might not be optimal in reality as they are not diversified enough. To improve that, we turn the bi-objective problem into a tri-objective one. Additionally to the CVaR and the expected return, we add the diversification as an objective.
In this case, the Pareto optimal portfolios are spanned between three points: The MRP, the MCVP and the *equal weights portfolio* (EWP) which is the most diversified portfolio. Consequently the set of all Pareto optimal portfolios is a two dimensional surface with three frontiers.
In a tri-objective problem it is not easy to find all Pareto-optimal portfolios using the ϵ-constraint method: Depending on the choice of $\bar{\epsilon}$ the problem might have multiple solutions or none. If $\bar{\epsilon}$ is chosen too pessimistic, many weakly efficient portfolios will satisfy the constraints additionally to the Pareto efficient one. If $\bar{\epsilon}$ is chosen too optimistic, the constraints might be so restrictive that there is no portfolio satisfying them. It is therefore difficult to find $\bar{\epsilon}$ such that the solution of 41 is Pareto efficient.
If our problem is strictly convex we already know that we will find all Pareto optimal solutions with the weighted sum method. It therefore appears to be a better approach on a tri-objective problem than the ϵ-constraint method:

$$\begin{aligned} \min_{\bar{x}\in\mathcal{X}} \quad -\lambda_1 \cdot \text{Return}(\bar{x}) + \lambda_2 \cdot \text{CVaR}(\bar{x}) - \lambda_3 \cdot \text{Diversification}(\bar{x}) \\ s.t. \\ \lambda_1 + \lambda_2 + \lambda_3 = 1 \end{aligned} \tag{75}$$

The first thing we have to do is choose a diversification measure for the objective function. The mean-deviation index is a linear function of the portfolio weights. As the expected return is a linear function as well, we cannot apply the second statement of theorem 3 if $\lambda_2 = 0$. Therefore the mean-deviation index is not suited as an objective function. However if we choose a strictly convex (concave) diversification measure, we can apply 3 for $\lambda_1 < 1$ even though the expected return is linear.
The Herfindhal index $\text{Div}_H(\bar{x}$ is a strictly concave risk measure and can thus fulfill our requirements on an objective function. The tail dependence diversification $Div_\lambda(\bar{x})$ is also strictly concave under our assumptions. Both risk measures are easy to implement in our framework as they are quadratic and therefore numerically optimizable.
We will demonstrate the method with the Herfindhal index:

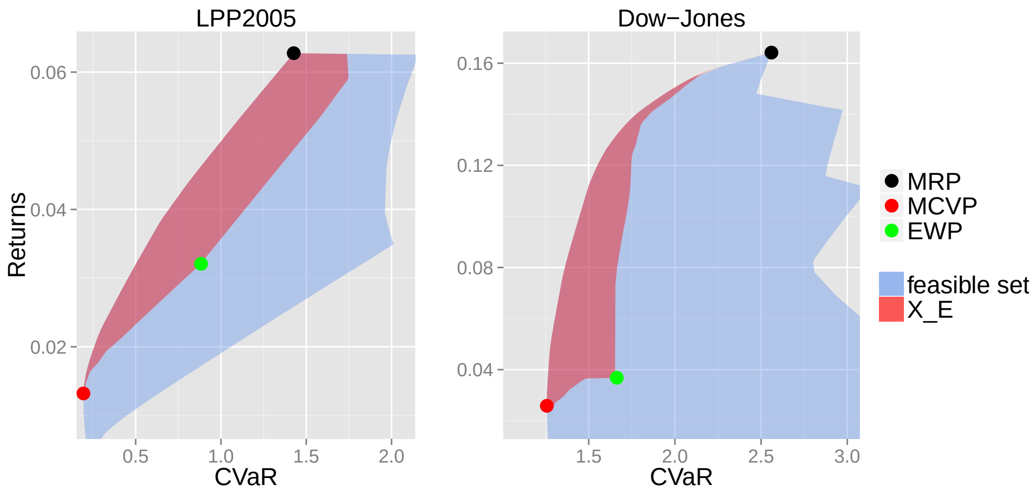

Figure 13: Visualization of $\mathcal{X}_E$ for the problem in equation 77 and the positions maximum-return portfolio (MRP), the minimum-CVaR portfolio (MCVP) and the equal weights portfolio (EWP).

$$
\begin{aligned}
f_1(\bar{x}) &:= -\bar{x}^T \bar{\mu} \\
f_2(\bar{x}) &:= \mathrm{CVaR}(\bar{x}) \\
f_3(\bar{x}) &:= \bar{x}^T \ \mathbf{I} \ \bar{x}
\end{aligned}
\tag{76}
$$

Keep in mind that it is uncomplicated to substitute the Herfindhal index with $Div_\lambda(\bar{x})$ as we just have to substitute the identity matrix $\mathbf{I}$ with $\boldsymbol{\lambda}$. Because of our normalization in equation 61, we can write $\lambda_3 = (1-\lambda_1-\lambda_2)$. The implementation of equation 75 then looks as following:

$$
\begin{aligned}
\min_{\bar{x},\mathbf{e},\beta} \quad -\lambda_1 \cdot \bar{x}^T \bar{\mu} + (1-\lambda) \cdot (\beta - \frac{1}{\alpha s} \sum_{i=0}^{s} e_i) + \bar{x}^T \cdot \bar{x} \\
s.t. \\
\sum_i x_i = 1 \\
x_i \geq 0. \\
e_j \geq \beta - \sum_{i=0}^{N} x_i r_{i,j}
\end{aligned}
\tag{77}
$$

As both f_2 and f_3 are strictly convex and f_1 is convex, we can apply theorem 3 and find that every Pareto optimal portfolio is found by equation 77. As for $\lambda_1 < 1$ every solution is unique due to theorem 3 and we already proved that for $\lambda_1 = 1$ the solution is the unique MRP, we can conclude that for every $\bar{\lambda}$ there exists a unique portfolio solution $\bar{x} \in \mathcal{X}_E$.
Figure 13 shows the surface of Pareto optimal portfolios in the CVaR-Return plane for the LPP2005 and for the Dow Jones index. Note that although not every point in the CVaR-Return plane is a unique portfolio, every point in the Pareto-optimal surface corresponds to exactly one Pareto optimal portfolio.
We will again normalize f_1 and f_2 as in chapter 4.2 to take values in $[0, 1]$. However we have to make adjustments: The minimum return-portfolio and the maximum CVaR-portfolio are not trivial anymore. We will make the following estimations: r_{MCVP} will be replaced with $r_{\min} = \min[r_{\text{MCVP}}, r_{\text{EWP}}]$ and c_{MRP} will be replaced with $c_{\max} = \max[c_{\text{MRP}}, c_{\text{EWP}}$. It is not guaranteed that f_1 and f_2 take values only $[0, 1]$ as there might be efficient portfolios with a smaller return than $r_{\min}$ on the frontier between the MCVP and the EWP or portfolios with a higher CVaR than $c_{\max}$ on the frontier between the MRP and the EWP. However it is unlikely that these values will significantly higher/lower than our estimates and our normalization is still close enough to $[0, 1]$.
As $f_3(\bar{x}) \in [0, 1]$, we do not have to normalize f_3. With normalization, equation 77 looks like this:

$$\begin{aligned} \min_{\bar{x},\mathbf{e},\beta} \quad & -\lambda_1 \cdot \tfrac{\bar{x}^T \cdot \bar{\mu} - r_{\min}}{r_{\text{MRP}} - r_{\min}} \\ & + \tfrac{(1-\lambda)}{c_{\min} - c_{\text{MCVP}}} (\beta - \tfrac{1}{\alpha s} \textstyle\sum_{i=0}^{s} e_i - c_{\text{MCVP}}) \\ & + (1 - \lambda_2 - \lambda_1) \cdot \bar{x}^T \bar{x} \\ & s.t. \\ & \sum_i x_i = 1 \\ & x_i \geq 0. \\ & e_i \geq \beta - \sum_{j=0}^{N} x_i r_{i,j} \end{aligned} \tag{78}$$

In figure 14, a set of portfolios calculated for the LPP2005 and the Dow Jones index is depicted in the CVaR-Return plane. The portfolios were calculated using equation 78 for different λ_1, λ_2-values. These values are drawn from a

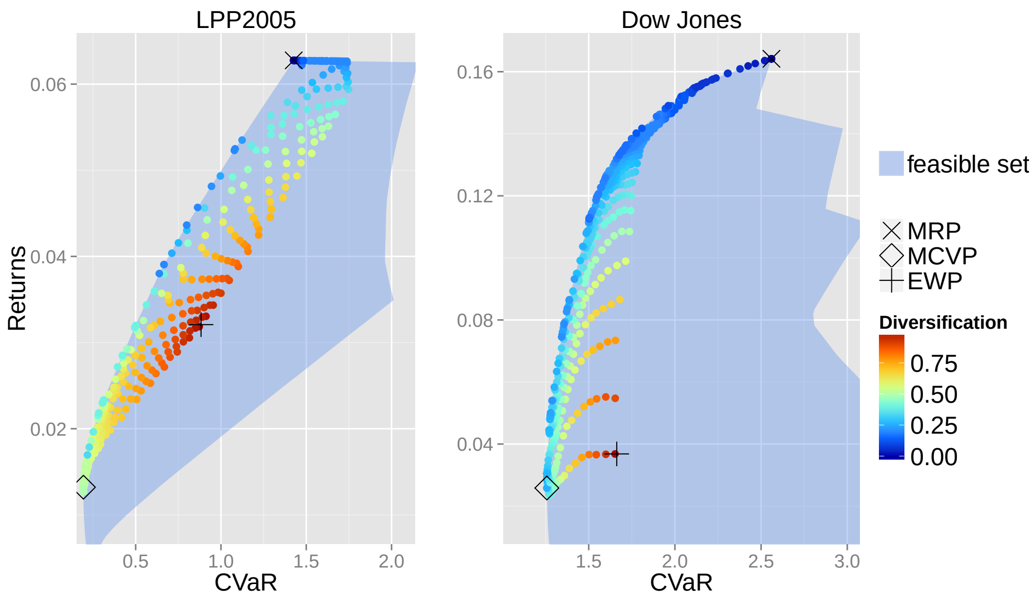

Figure 14: Portfolio positions in the CVaR-Return plane when optimized with goal programming using equation 78. The points are drawn from a grid with distance $\frac{1}{30}$ between the λ-values. The color of a point indicates the diversification of the portfolio.

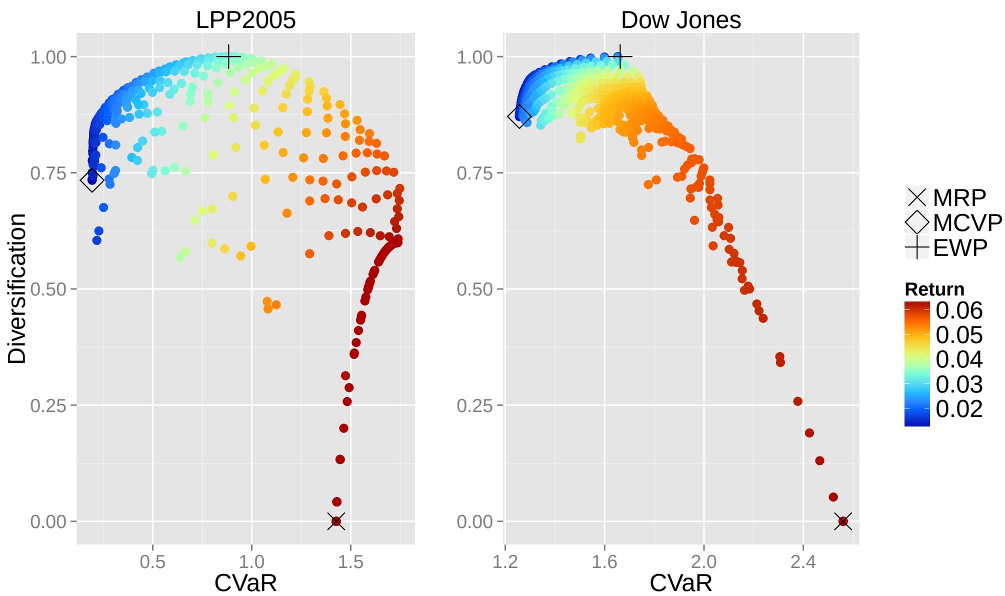

Figure 15: Portfolio positions in the CVaR-Diversification plane when optimized with goal programming using equation 78. The points are drawn from a grid with distance $\frac{1}{30}$ between the λ-values. The color of a point indicates the expected return of the portfolio.

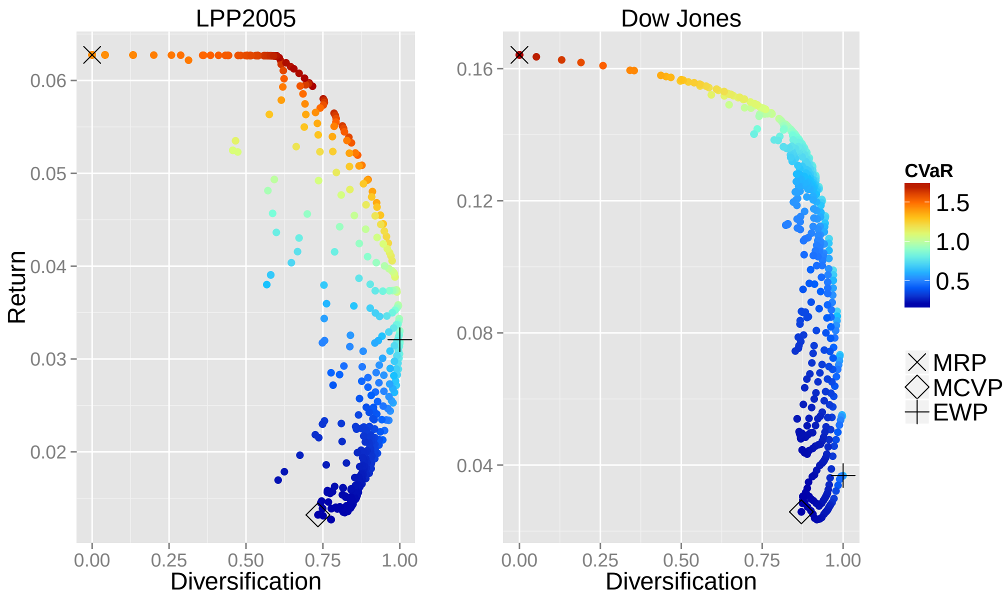

Figure 16: Portfolio positions in the Diversification-Return plane when optimized with goal programming using 78. The points are drawn from a grid with distance $\frac{1}{30}$ between the λ-values. The color of a point indicates the CVaR of the portfolio.

grid of equidistant $\lambda_1 - \lambda_2$-points with $\lambda_1 + \lambda_2 \leq 1$. The distance between the $\bar{\lambda}$-points is $\frac{1}{30}$. This grid can be seen in figure 26.
We can also display the portfolio solutions in other planes than the CVaR-Risk plane: Figure 15 and 16 show the set of calculated portfolios in the CVaR-Diversification plane and in the Diversification-Return plane.
When comparing the different frontiers the effect of the eminently non-linear form of the diversification measure can be perceived. We can see how uniformly the portfolios are distributed across the frontier between the EWP and the MCVP for both the Dow Jones and the LPP2005 in figure 15. Also in figure 16 the frontier between the EWP and in this case the MRP is populated evenly for both plots. This stands in contrast to the frontier between the MCVP and the MRP that we discussed in chapter 4.2. This frontier is populated very sparsely for the LPP2005 and only moderately for the Dow Jones.
The uniform behaviour of the weighted sum method for the frontiers spanning from the EWP can be explained by non-linearity of the Herfindhal index. It is quadratic and thus everywhere far from being linear. Thus for both frontiers at least one objective function and consequently the sum of all objective

functions show non-linear behaviour. For the frontier between the MCVP and the MRP the contribution of the diversification is always zero. Due to the linearity of the expected return and the occurring closeness of the CVaR to a linear function, the weighted sum method does not distribute the portfolios as even as for the other frontiers.

4.4 Ternary maps

A *ternary map* or *ternary plot* is a plot used to show the compositions of systems composed of three species [27]. A ternary map has a triangular barycentric coordinate system in which every position corresponds to a ratio of three variables that sum to a constant. Each of the three corners correspond to a composition of exclusively one species. Ternary maps are very common in metallurgy where they depict characteristics of alloys as a function of their composition. Since λ_1, λ_2 and λ_3 in equation 78 also sum up to a constant, every $\bar{\lambda}$-values can be seen as a position in a ternary map.
In figure 17 to 22, ternary maps for the LPP2005 and the Dow Jones are depicted. Additionally, ternary maps for the DAX can be found in the appendix in figure 28, 29 and 30.
These ternary maps have three axes, one for each λ_i. As $\lambda_3 = (1 - \lambda_1 - \lambda_2)$, all allowed values for $\bar{\lambda}$ can be mapped to a point on the grid. In each of the three corners lies one of the three extreme portfolios, the MRP, the MCVP or the EWP. Contour lines are added to illustrate behaviour of the expected return, the CVaR and the diversification in dependence of $\bar{\lambda}$.
Figure 17, 19 and 21 display the Diversification, the CVaR and the expected return for the LPP2005. The different distribution of portfolios along the three frontiers can be seen very clearly in the ternary plots: The frontier between the MCVP and the MRP can be seen between λ_1 and λ_2. The contour lines lying on this frontier between are all in very close intervals for all three plots.
In the center of the ternary maps as well as on the other two frontiers the portfolios appear to be distributed better. The contour lines are separated more and spread across the map. Again this demonstrates the importance of the strict convexity of the objective functions.
Figure 18, 20 and 22 display the Diversification, the CVaR and the expected return for the Dow-Jones index. The situation is not as dramatic as for the LPP2005: The contour lines are more spread across the different frontier for all three plots and also look reasonable in the center of the maps. This is also the case for DAX as visible in figure 28 to 30
Ideally the method in equation 78 allocates the portfolio coherently as a func-

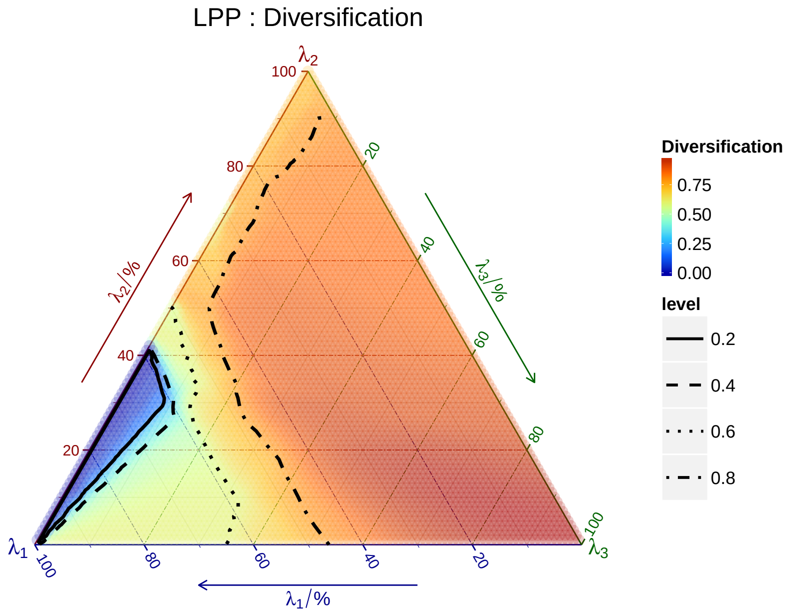

Figure 17: Ternary map showing the levels of diversification for different $\bar{\lambda}$-values for the LPP2005. Here again $\lambda_3 = (1 - \lambda_1 - \lambda_2)$.

tion of $\bar{\lambda}$. The maps should show the same behaviour for different datasets such that portfolios obtained for a distinct $\bar{\lambda}$ value have the same characteristics for different sets of available assets.
As already discussed the LPP2005 is a special case that illustrates the deficiencies of the method in equation 78. However except for the frontier between the MRP and the MCVP, all ternary plots show a similar behaviour for the LPP2005 and the Dow-Jones index. Especially the CVaR and the expected return allocation appear alike for the LPP2005 and the Dow-Jones index.
If we compare the ternary maps of the Dow-Jones index, we find a very similar behaviour of the method in the ternary maps. The contour lines mostly have the same form. From this we conclude that the tri-criteria optimization method allocates portfolios coherently in dependence of $\bar{\lambda}$ for bigger sets of available assets.

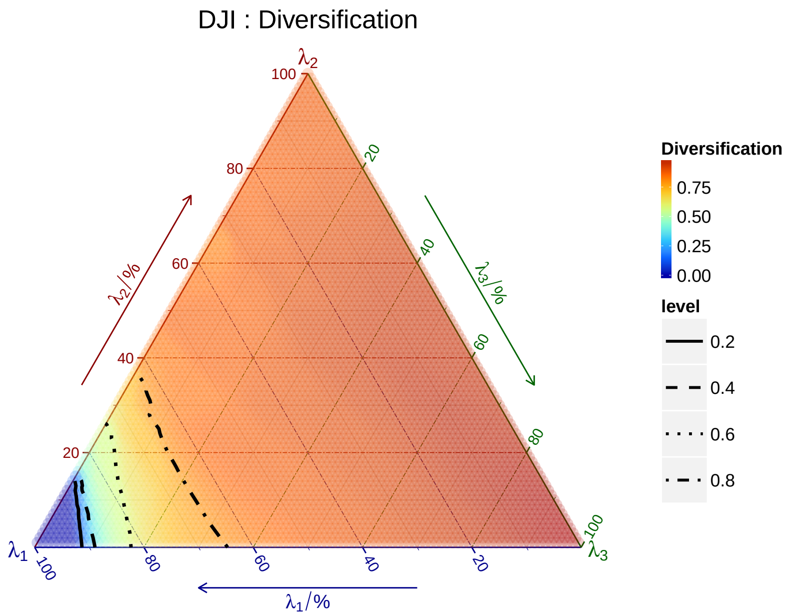

Figure 18: Ternary map showing the levels of diversification for different $\bar{\lambda}$-values for the Dow Jones index.

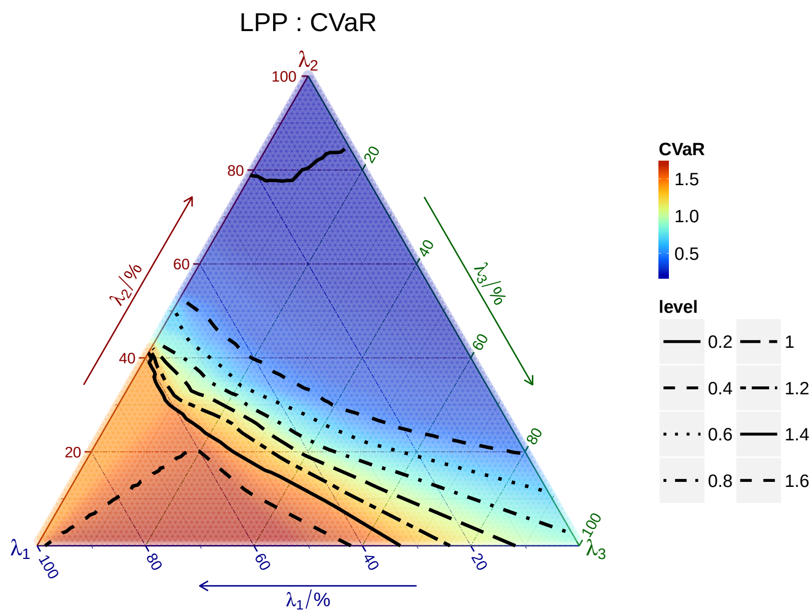

Figure 19: Ternary map showing the CVaR-levels for different $\bar{\lambda}$-values for the LPP2005.

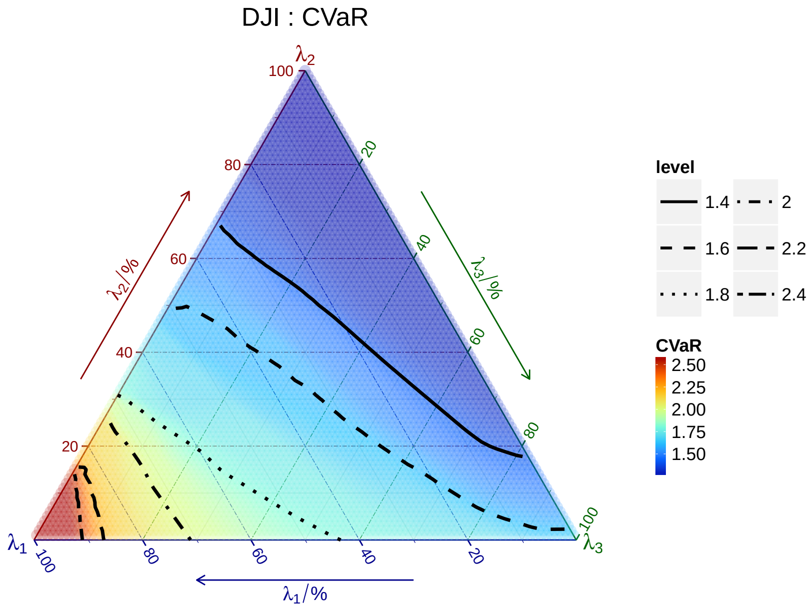

Figure 20: Ternary map showing the CVaR-levels for different $\bar{\lambda}$-values for the Dow Jones index.

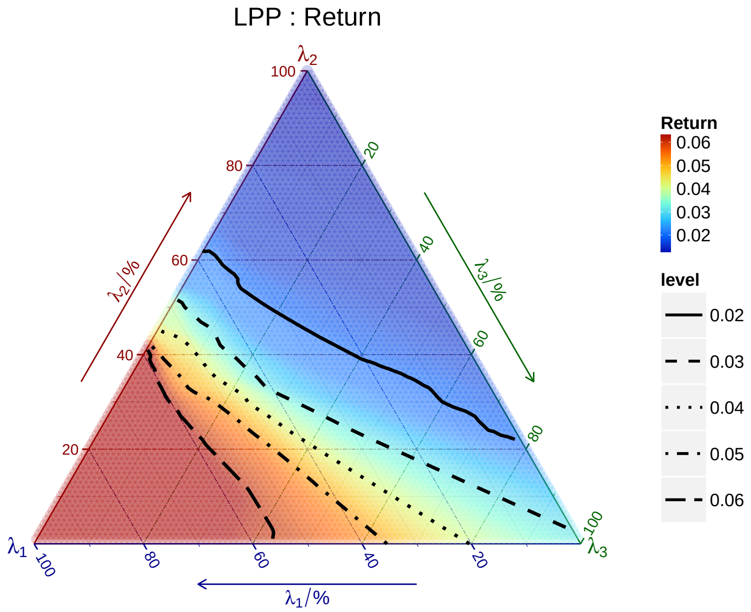

Figure 21: Ternary map showing the expected return levels for different $\bar{\lambda}$-values for the LPP2005.

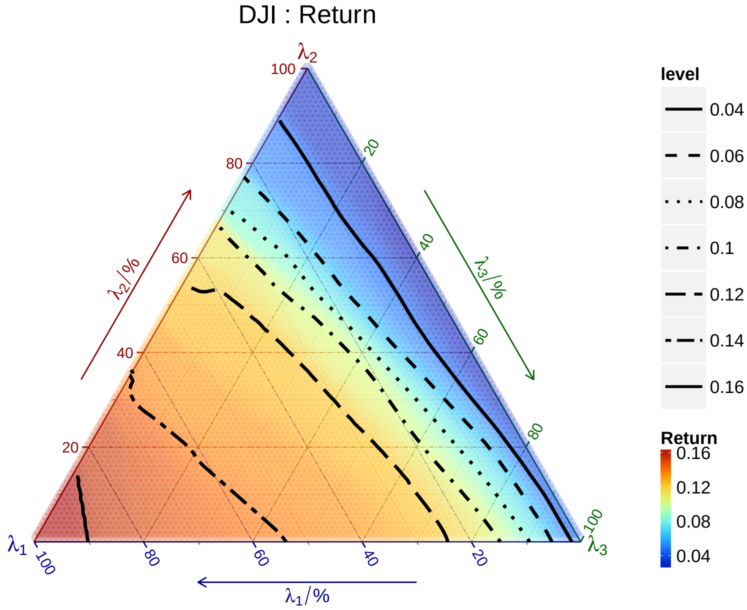

Figure 22: Ternary map showing the expected return levels for different $\bar{\lambda}$-values for the Dow Jones index.

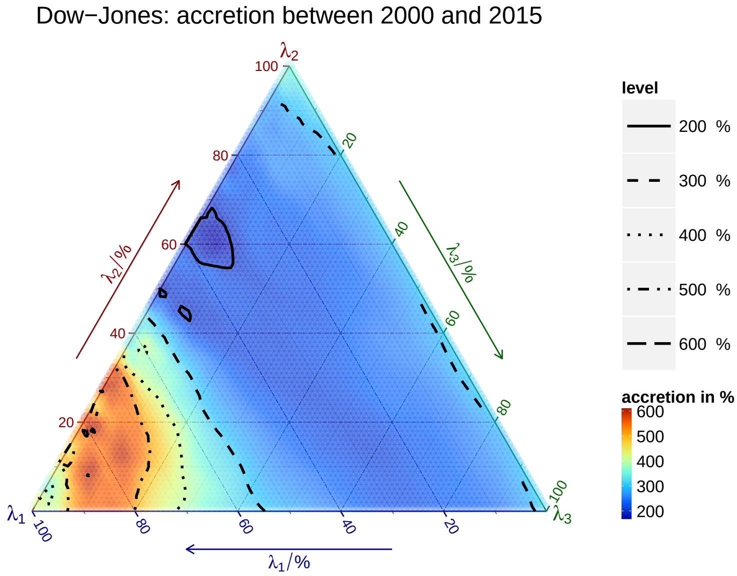

Figure 23: Ternary map showing the accrecation levels for different $\bar{\lambda}$-values for the Dow Jones index for a period between March 20 2008 and June 25 2015.

5 Rolling-window analysis

We tested the method developed in the last chapter with a rolling-window analysis in order to test the methods profitability over time. To test the method we used a dataset of the stock components of the Dow-Jones index between 2000-01-05 and 2015-07-08 and a dataset of the stock components of the DAX index between 2003-01-07 and 2015-07-27.
The analysis was done for three months of estimation time and one month of investment time: The first three consecutive months of daily returns of the dataset are used as sample input for the algorithm in equation 78. The portfolio weights are calculated from equation 78 for a grid of $\bar{\lambda}$-values. Afterwards the accretion of each portfolio for the next consecutive month is calculated and stored. Then procedure is repeated for a time window that is shifted by one month such that the investment time of the newly calculated portfolio is one month later than in the preceding step. This analysis was done for a grid of $\bar{\lambda}$-values
Figure 23 and 24 show the total accretion for each $\bar{\lambda}$-strategy. As visible, both maps show a similar behaviour in the profitability of the different $\bar{\lambda}$-values. In both maps the most profitable $\bar{\lambda}$-strategies have $\lambda_1 \in [0.7, 0.95]$. It appears that the emphasis between λ_2 and λ_3 does not influence the profitability

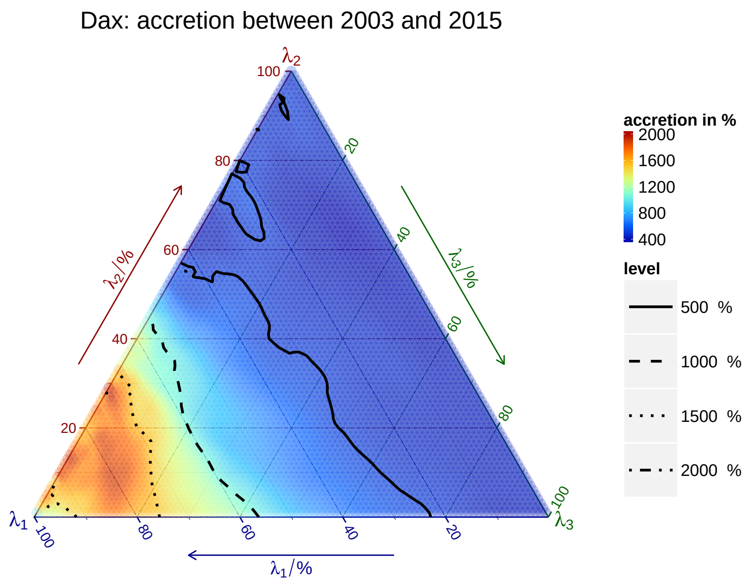

Figure 24: Ternary map showing the accrecation levels for different $\bar{\lambda}$-values for the LPP2005 for a period between 2005 and 2006.

much in this case.

6 Runtime comparisons

In order to test the implementation of equation 78, we relied on the algebraic modeling language *AMPL*[8] that is built for large-scale optimization problems. It supports multiple commercial and open-source solvers from the *COIN-OR* project[9] such as CBC, CPLEX, SNOPT and MINOS. Multiple runtime performance comparisons are depicted below. The implemenation of equation 78 among other implementations is available in the function *CVaRObj* in the R-package *MOPP: Multi-Objective Portfolio Programming*[8].
The calculations were done on a personal computer running *Windows 7*-OS with 8,00 GB memory and a *Intel Core i7-2600*-processor with four 3.4 GHz cores.
We tested the implementation for different numbers of available assets and for different grid-widths of the $\bar{\lambda}$-values. The runtimes can be seen in table 1 and 2.

	Numer of assets	5	25	50	100
Runtime [s]	Minos	11.5	12.2	16.1	19.5
	CPLEX	13.75	47.5	59.9	68.2
	Gurobi	17.3	20.6	46.5	39.5
	SNOPT	9.8	12.5	17.9	18.9

Table 1: Runtime comparisons of AMPL-solvers for multiple numbers of available assets. The λ-grid that was used was of mesh width 1/30.

As visible, the runtime scales weakly with the number of available assets and depends strongly on the solver. As expected, the runtime scales inverse quadratically with the grid-with of the $\bar{\lambda}$-values.

	λ-grid	1/15	1/30	1/60	1/120
Runtime [s]	Minos	4.2	13.9	61.2	218.9

Table 2: Runtime comparisons for different mesh widths. The number of available assets was set to 50.

[8] `www.ampl.com`
[9] `http://www.coin-or.org/download/binary/`

7 Conclusion and Outlook

The weighted sum method is an elegant and naturally comprehensible method to optimize convex multi-criteria problems. For strictly convex problems it finds only Pareto efficient solutions and is able to find the whole set of Pareto efficient solutions.

The discussion of the weighted sum method showed that it is a valid approach for portfolio optimization. The traditional portfolio problem is a bi-criteria problem. It was shown in this thesis that it can be solved with the weighted sum method for the expected return and a scenario-based risk measure as objectives. Under reasonable conditions on the available assets, the scenario-based risk measure is strictly convex and consequently the weighted sum method finds all portfolios on the efficient frontier.

However the quality of the portfolio allocations by this method is not necessarily coherent and depends on the number of available assets as demonstrated with the LPP2005 index. For larger numbers of available assets, the allocation appears to be more coherent.

The weighted sum method allows the optimization of multiple objectives given that only one objective function is not strictly convex. As the portfolios on the efficient frontier are not very diversified, we introduced a method that adds the diversification as an additional objective. This method can be implemented for an arbitrary scenario-based risk measure and a quadratic diversification measure. With equation 78 we presented an implementation of this method for the linearized CVaR and the Herfindahl index. This method is able to find all Pareto efficient solutions under these three objectives.

Although the Herfindhal index adds convexity to the problem, the allocation of portfolios in dependence of the parameter $\bar{\lambda}$ again appears to be more coherent for sets of many available assets than for smaller sets. In these cases the diversification, the CVaR and the expected return are properties that are distributed smoothly for different $\bar{\lambda}$-values.

The rolling-window analysis showed that adding diversification as an objective indeed improves the profitability on a long term basis. The most profitable portfolios are found close to the MRP but have significantly higher accretion than the MRP.

Outlook

In the future this method can be used to further analyze dependencies between profitability of portfolios and their diversification, their CVaR and their expected return as it enables an fast way of finding all efficient portfo-

lios. The implementation of equation 78 is a quadratic problem and can thus be efficiently solved by industrial solvers such as *CPLEX* and *Minos*. It is therefore a fast way to analyze large sets of assets.
As explained, the objective functions can be easily stubstituted by other scenario-based risk measures and quadratic diversification measures such as the tail dependence diversification.
In principle it is possible to add additional objectives to the implemented method. The condition on these objectives is that the objective functions are strictly convex. With more objectives, connections between multiple portfolio characteristics can be examined.

References

[1] Carlo Acerbi, Claudio Nordio, and Carlo Sirtori. Expected shortfall as a tool for financial risk management. *arXiv preprint cond-mat/0102304*, 2001.

[2] Carlo Acerbi and Dirk Tasche. Expected shortfall: a natural coherent alternative to value at risk. *Economic notes*, 31(2):379–388, 2002.

[3] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical finance*, 9(3):203–228, 1999.

[4] Eric Bouyé, Valdo Durrleman, Ashkan Nikeghbali, Gaël Riboulet, and Thierry Roncalli. Copulas for finance-a reading guide and some applications. *Available at SSRN 1032533*, 2000.

[5] Gary Chamberlain. A characterization of the distributions that imply mean-variance utility functions. *Journal of Economic Theory*, 29(1):185–201, 1983.

[6] Patrick Cheridito and Tianhui Li. Dual characterization of properties of risk measures on orlicz hearts. *Mathematics and Financial Economics*, 2(1):29–55, 2008.

[7] Rémy Chicheportiche and Jean-Philippe Bouchaud. The joint distribution of stock returns is not elliptical. *International Journal of Theoretical and Applied Finance*, 15(03), 2012.

[8] Henry Clausen and Tomas Tamfal. *MOPP: Multi-objective Portfolio Programming*, 2015. R package version 1.1 — For new features, see the 'Changelog' file (in the package source).

[9] Matthias Ehrgott. *Multicriteria optimization.* Springer Science & Business Media, 2006.

[10] Noureddine El Karoui et al. High-dimensionality effects in the markowitz problem and other quadratic programs with linear constraints: risk underestimation. *The Annals of Statistics*, 38(6):3487–3566, 2010.

[11] Edwin J Elton, Martin J Gruber, Stephen J Brown, and William N Goetzmann. *Modern portfolio theory and investment analysis.* John Wiley & Sons, 2009.

[12] Paul Embrechts, Rdiger Frey, and Alexander McNeil. Quantitative risk management. *Princeton Series in Finance, Princeton*, 10, 2005.

[13] Frank J Fabozzi, Petter N Kolm, Dessislava Pachamanova, and Sergio M Focardi. *Robust portfolio optimization and management*. John Wiley & Sons, 2007.

[14] Willliam Feller. *An introduction to probability theory and its applications*, volume 2. John Wiley & Sons, 2008.

[15] Oleg Grodzevich and Oleksandr Romanko. Normalization and other topics in multi-objective optimization. 2006.

[16] Philippe Jorion. *Value at risk: the new benchmark for managing financial risk*, volume 3. McGraw-Hill New York, 2007.

[17] Noureddine El Karoui. On the realized risk of high-dimensional markowitz portfolios. *SIAM Journal on Financial Mathematics*, 4(1):737–783, 2013.

[18] Ulrich Kirchner and Caroline Zunckel. Measuring portfolio diversification. *arXiv preprint arXiv:1102.4722*, 2011.

[19] Pavlo Krokhmal, Jonas Palmquist, and Stanislav Uryasev. Portfolio optimization with conditional value-at-risk objective and constraints. *Journal of risk*, 4:43–68, 2002.

[20] Harry Markowitz. Portfolio selection*. *The journal of finance*, 7(1):77–91, 1952.

[21] Helmut Mausser and Dan Rosen. Beyond var: From measuring risk to managing risk. In *Computational Intelligence for Financial Engineering, 1999.(CIFEr) Proceedings of the IEEE/IAFE 1999 Conference on*, pages 163–178. IEEE, 1999.

[22] John Norstad. An introduction to portfolio theory, 1999.

[23] Joel Owen and Ramon Rabinovitch. On the class of elliptical distributions and their applications to the theory of portfolio choice. *The Journal of Finance*, 38(3):745–752, 1983.

[24] R Tyrrell Rockafellar and Stanislav Uryasev. Optimization of conditional value-at-risk. *Journal of risk*, 2:21–42, 2000.

[25] R Tyrrell Rockafellar and Stanislav Uryasev. Conditional value-at-risk for general loss distributions. *Journal of banking & finance*, 26(7):1443–1471, 2002.

[26] Arthur Sullivan. Economics: Principles in action. 2003.

[27] Ronald Frank Tylecote. *A history of metallurgy.* The Metals Society, 1976.

A Appendix

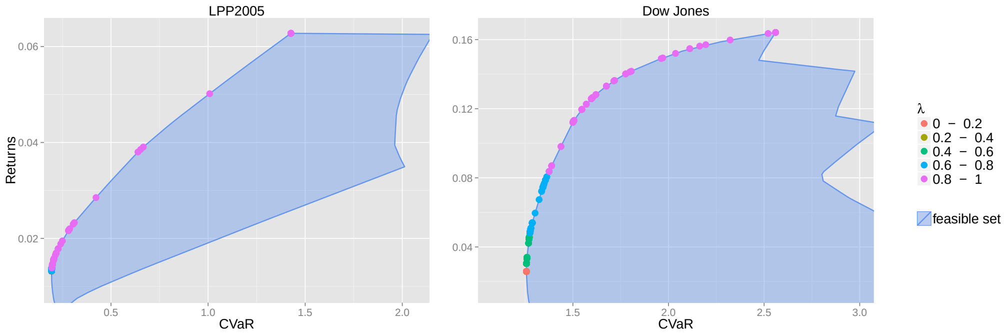

Figure 25: Visualization of the portfolio distribution for different λ values for the unnormalized critical line algorithm.

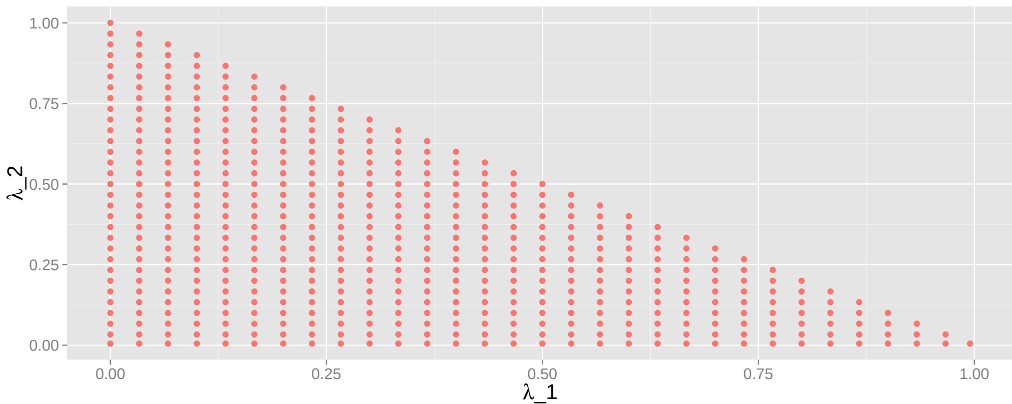

Figure 26: Positions of the different $\lambda_1 - \lambda_2$ -points.

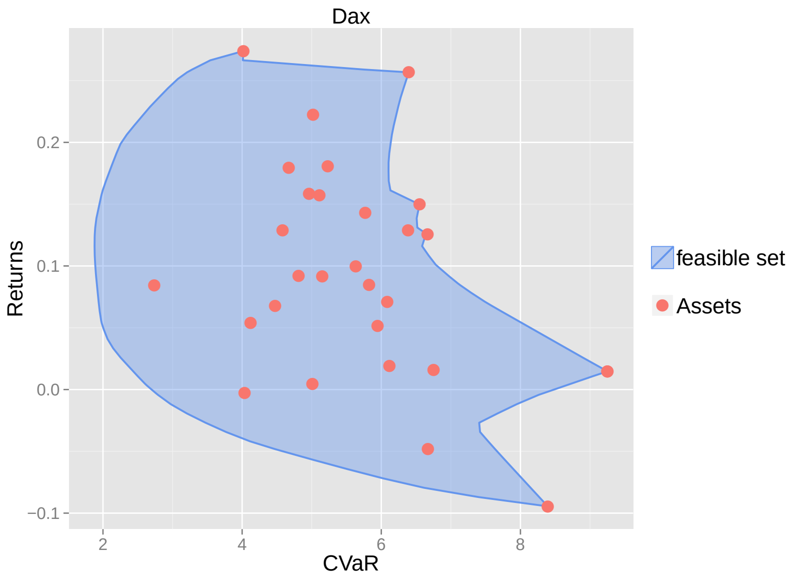

Figure 27: Visualization of the feasible set $\mathcal{Y}$ in the Risk-Return-space and the position of the available assets for the DAX. The chosen risk measure is the CVaR.

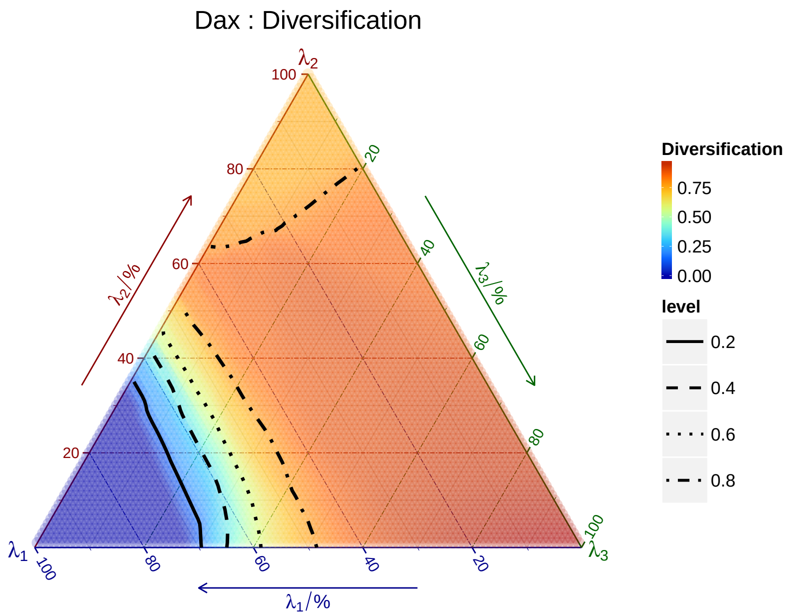

Figure 28: Ternary map showing the diversification levels for different $\bar{\lambda}$-values for the DAX index.

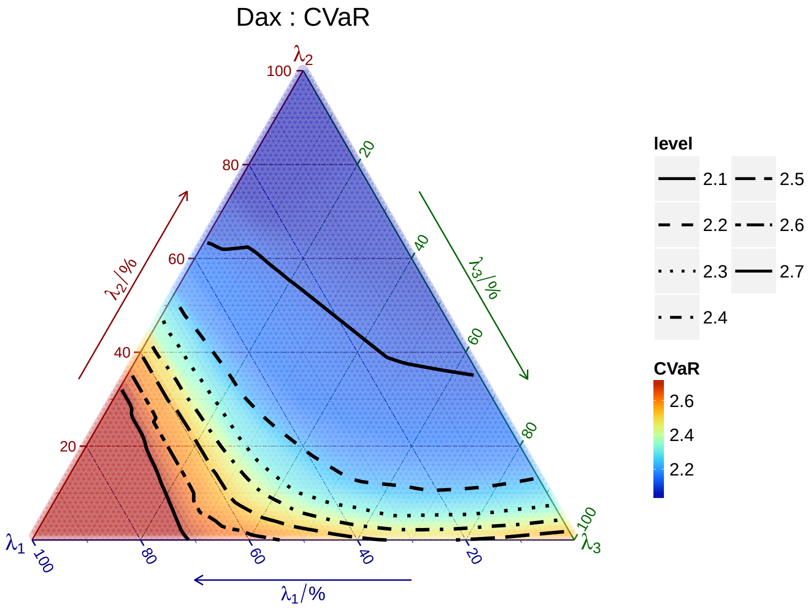

Figure 29: Ternary map showing the CVaR-levels for different $\bar{\lambda}$-values for the DAX index.

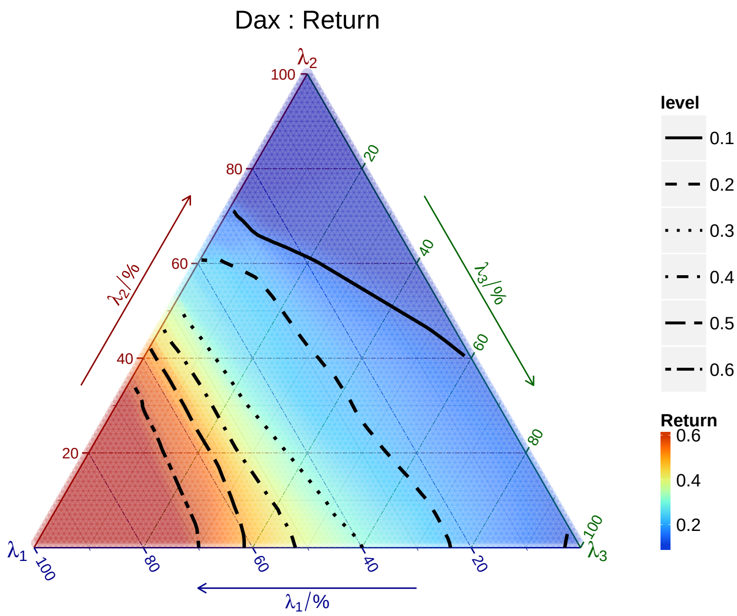

Figure 30: Ternary map showing the expected return levels for different $\bar{\lambda}$-values for the DAX index.