COMISEF Computational Optimization Methods in Statistics, Econometrics and Finance

- Marie Curie Research and Training Network funded by the EU Commission through MRTN-CT-2006-034270 -

COMISEF WORKING PAPERS SERIES

WPS-017 04/11/2009

Worst-Case Value-at-Risk of Non-Linear Portfolios

S. Zymler
D. Kuhn
B. Rustem

Worst-Case Value-at-Risk of Non-Linear Portfolios

Steve Zymler*, Daniel Kuhn and Berç Rustem

Department of Computing
Imperial College of Science, Technology and Medicine
180 Queen's Gate, London SW7 2AZ, UK.

August 18, 2009

Abstract

Portfolio optimization problems involving Value-at-Risk (VaR) are often computationally intractable and require complete information about the return distribution of the portfolio constituents, which is rarely available in practice. These difficulties are further compounded when the portfolio contains derivatives. We develop two tractable conservative approximations for the VaR of a derivative portfolio by evaluating the worst-case VaR over all return distributions of the derivative underliers with given first- and second-order moments. The derivative returns are modelled as convex piecewise linear or—by using a delta-gamma approximation—as (possibly non-convex) quadratic functions of the returns of the derivative underliers. These models lead to new Worst-Case Polyhedral VaR (WCPVaR) and Worst-Case Quadratic VaR (WCQVaR) approximations, respectively. WCPVaR is a suitable VaR approximation for portfolios containing long positions in European options expiring at the end of the investment horizon, whereas WCQVaR is suitable for portfolios containing long and/or short positions in European and/or exotic options expiring beyond the investment horizon. We prove that WCPVaR and WCQVaR optimization can be formulated as tractable second-order cone and semidefinite programs, respectively, and reveal interesting connections to robust portfolio optimization. Numerical experiments demonstrate the benefits of incorporating non-linear relationships between the asset returns into a worst-case VaR model.

Key words. Value-at-Risk, Derivatives, Robust Optimization, Second-Order Cone Programming, Semidefinite Programming

1 Introduction

Investors face the challenging problem of how to distribute their current wealth over a set of available assets with the goal to earn the highest possible future wealth. One of the first mathematical models for this problem was formulated by Markowitz [17], who observed that a prudent investor does not aim solely at maximizing the expected return of an investment, but

*Corresponding author: sz02@doc.ic.ac.uk

also at minimizing its risk. In the Markowitz model, the risk of a portfolio is measured by the variance of the portfolio return.

Although mean-variance optimization is appropriate when the asset returns are symmetrically distributed, it is known to result in counter intuitive asset allocations when the portfolio return is skewed. This shortcoming triggered extensive research on downside risk measures. Due to its intuitive appeal and since its use is enforced by financial regulators, Value-at-Risk (VaR) remains the most popular downside risk measure [14]. The VaR at level ϵ is defined as the $(1 - \epsilon)$-quantile of the portfolio loss distribution.

Despite its popularity, VaR lacks some desirable theoretical properties. Firstly, VaR is known to be a non-convex risk measure. As a result, VaR optimization problems usually are computationally intractable. In fact, they belong to the class of chance-constrained stochastic programs, which are notoriously difficult to solve. Secondly, VaR fails to satisfy the subadditivity property of coherent risk measures [3]. Thus, the VaR of a portfolio can exceed the weighted sum of the VaRs of its constituents. In other words, VaR may penalize diversification. Thirdly, the computation of VaR requires precise knowledge of the joint probability distribution of the asset returns, which is rarely available in practice.

A typical investor may know the first- and second-order moments of the asset returns but is unlikely to have complete information about their distribution. Therefore, El Ghaoui *et al.* [11] propose to maximize the VaR of a given portfolio over all asset return distributions consistent with the known moments. The resulting Worst-Case VaR (WCVaR) represents a conservative (that is, pessimistic) approximation for the true (unknown) portfolio VaR. In contrast to VaR, WCVaR represents a convex function of the portfolio weights and can be optimized efficiently by solving a tractable second-order cone program. El Ghaoui *et al.* [11] also disclose an interesting connection to robust optimization [5, 6, 22]: WCVaR coincides with the worst-case portfolio loss when the asset returns are confined to an *ellipsoidal uncertainty set* determined through the known means and covariances.

In this paper we study portfolios containing derivatives, the most prominent examples of which are European call and put options. Sophisticated investors frequently enrich their portfolios with derivative products, be it for hedging and risk management or speculative purposes. In the presence of derivatives, WCVaR still constitutes a tractable conservative approximation for the true portfolio VaR. However, it tends to be over-pessimistic and thus may result in undesirable portfolio allocations. The main reasons for the inadequacy of WCVaR are the following.

- The calculation of WCVaR requires the first- and second-order moments of the derivative returns as an input. These moments are difficult or (in the case of exotic options) almost impossible to estimate due to scarcity of time series data.

- WCVaR disregards perfect dependencies between the derivative returns and the underlying asset returns. These (typically non-linear) dependencies are known in practice as they can be inferred from contractual specifications (payoff functions) or option pricing models. Note that the covariance matrix of the asset returns, which is supplied to the WCVaR model, fails to capture non-linear dependencies among the asset returns, and therefore WCVaR tends to severely *overestimate* the true VaR of a portfolio containing derivatives.

Recall that WCVaR can be calculated as the optimal value of a robust optimization problem with an ellipsoidal uncertainty set, which is highly symmetric. This symmetry hints at the inadequacy of WCVaR from a geometrical viewpoint. An intuitively appealing uncertainty set should be asymmetric to reflect the skewness of the derivative returns. Recently, Natarajan *et al.* [19] included asymmetric distributional information into the WCVaR optimization in order to obtain a tighter approximation of VaR. However, their model requires forward- and backward-deviation measures as an input, which are difficult to estimate for derivatives. In contrast, reliable information about the functional relationships between the returns of the derivatives and their underlying assets is readily available.

In this paper we develop novel Worst-Case VaR models which explicitly account for perfect non-linear dependencies between the asset returns. We first introduce the *Worst-Case Polyhedral VaR* (WCPVaR), which provides a tight conservative approximation for the VaR of a portfolio containing European-style options expiring at the end of the investment horizon. In this situation, the option returns constitute convex piecewise-linear functions of the underlying asset returns. WCPVaR evaluates the worst-case VaR over all asset return distributions consistent with the given first- and second-order moments of the option underliers and the piecewise linear relation between the asset returns. Under a no short-sales restriction on the options, we are able to formulate WCPVaR optimization as a convex second-order cone program, which can be solved efficiently [2]. We also establish the equivalence of the WCPVaR model to a robust optimization model described in [27].

Next, we introduce the *Worst-Case Quadratic VaR* (WCQVaR) which approximates the VaR of a portfolio containing long and/or short positions in plain vanilla and/or exotic options with arbitrary maturity dates. In contrast to WCPVaR, WCQVaR assumes that the derivative returns are representable as (possibly non-convex) quadratic functions of the underlying asset returns. This can always be enforced by invoking a *delta-gamma approximation*, that is, a second-order Taylor approximation of the portfolio return. The delta-gamma approximation is popular in many branches of finance and is accurate for short investment periods. Moreover, it has been used extensively for VaR estimation, see, e.g., the surveys by Jaschke [13] and Mina and Ulmer [18]. However, to the best of our knowledge, the delta-gamma approximation has never been used in a VaR optimization model. We define WCQVaR as the worst-case VaR over all asset return distributions consistent with the known first- and second-order moments of the option underliers and the given quadratic relation between the asset returns. WCQVaR provides a tight conservative approximation for the true portfolio VaR if the delta-gamma approximation is accurate. We show that WCQVaR optimization can be formulated as a convex semidefinite program, which can be solved efficiently [26], and we establish a connection to a novel robust optimization problem. The main contributions of this paper can be summarized as follows:

(1) We generalize the WCVaR model [11] to explicitly account for the non-linear relationships between the derivative returns and the underlying asset returns. To this end, we develop the WCPVaR and WCQVaR models as described above. We show that in the absence of derivatives both models reduce to the WCVaR model. Moreover, we formulate WCPVaR optimization as a second-order cone program and WCQVaR optimization as a semidefinite program. Both models are polynomial time solvable.

(2) We show that both the WCPVaR and the WCQVaR models have equivalent reformulations as robust optimization problems. We explicitly construct the associated uncertainty sets which are, unlike conventional ellipsoidal uncertainty sets, asymmetrically oriented around the mean values of the asset returns. This asymmetry is caused by the non-linear dependence of the derivative returns on their underlying asset returns. Simple examples illustrate that the new models may approximate the true portfolio VaR significantly better than WCVaR in the presence of derivatives.

(3) The robust WCQVaR model is of relevance beyond the financial domain because it constitutes a tractable approximation of a chance-constrained stochastic program that is affine in the decision variables but (possibly non-convex) *quadratic* in the uncertainties. Although tractable approximations for chance constrained programs with affine perturbations have been researched extensively (see, e.g., [20]), the case of quadratic data dependence has remained largely unexplored (with the exception of [4, §1.4]).

(4) We evaluate the WCQVaR model in the context of an index tracking application. We show that when investment in options is allowed, the optimal portfolios exhibit vastly improved out-of-sample performance compared to the optimal portfolios based on stocks only.

The remainder of the paper is organized as follows. In Section 2 we review the mathematical definitions of VaR and WCVaR. Moreover, we recall the relationship between WCVaR optimization and robust optimization. In Section 3 we highlight the shortcomings of WCVaR in the presence of derivatives. In Section 4 we develop the WCPVaR model in which the option returns are modelled as convex piecewise-linear functions of the underlying asset returns. We prove that it can be reformulated as a second-order cone program and construct the uncertainty set which generates the equivalent robust portfolio optimization model. In Section 5 we describe the WCQVaR model, which approximates the portfolio return by a quadratic function of the underlying asset returns. We show that it can be reformulated as a semidefinite program and prove its equivalence to an augmented robust optimization problem whose uncertainty set is embedded into the space of positive semidefinite matrices. Section 6 evaluates the out-of-sample performance of the WCQVaR model in the context of an index tracking application. Conclusions are drawn in Section 7.

Notation. We use lower-case bold face letters to denote vectors and upper-case bold face letters to denote matrices. The space of symmetric matrices of dimension n is denoted by $\mathbb{S}^n$. For any two matrices $\mathbf{X}, \mathbf{Y} \in \mathbb{S}^n$, we let $\langle \mathbf{X}, \mathbf{Y} \rangle = \mathrm{Tr}(\mathbf{XY})$ be the trace scalar product, while the relation $\mathbf{X} \succcurlyeq \mathbf{Y}$ ($\mathbf{X} \succ \mathbf{Y}$) implies that $\mathbf{X} - \mathbf{Y}$ is positive semidefinite (positive definite). Random variables are always represented by symbols with tildes, while their realizations are denoted by the same symbols without tildes. Unless stated otherwise, equations involving random variables are assumed to hold almost surely. In the case of distributional ambiguity, the equations hold almost surely with respect to each distribution under consideration.

2 Worst-Case Value-at-Risk Optimization

Consider a market consisting of m assets such as equities, bonds, and currencies. We denote the present as time $t = 0$ and the end of the investment horizon as $t = T$. A portfolio is characterized by a vector of asset weights $\boldsymbol{w} \in \mathbb{R}^m$, whose elements add up to 1. The component w_i denotes the percentage of total wealth which is invested in the ith asset at time $t = 0$. Furthermore, $\tilde{\boldsymbol{r}}$ denotes the $\mathbb{R}^m$-valued random vector of relative assets returns over the investment horizon. By definition, an investor will receive $1 + \tilde{r}_i$ dollars at time T for every dollar invested in asset i at time 0. The return of a given portfolio $\boldsymbol{w}$ over the investment period is thus given by the random variable

$$\tilde{r}_p = \boldsymbol{w}^T \tilde{\boldsymbol{r}}. \tag{1}$$

Loosely speaking, we aim at finding an allocation vector $\boldsymbol{w}$ which entails a high portfolio return, whilst keeping the associated risk at an acceptable level. Depending on how risk is defined, we end up with different portfolio optimization models.

Arguably one of the most popular measures of risk is the *Value-at-Risk* (VaR). The VaR at level ϵ is defined as the $(1-\epsilon)$-percentile of the portfolio loss distribution, where ϵ is typically chosen as 1% or 5%. Put differently, $\mathrm{VaR}_\epsilon(\boldsymbol{w})$ is defined as the smallest real number γ with the property that $-\boldsymbol{w}^T\tilde{\boldsymbol{r}}$ exceeds γ with a probability not larger than ϵ, that is,

$$\mathrm{VaR}_\epsilon(\boldsymbol{w}) = \min\left\{\gamma \ : \ \mathbb{P}\{\gamma \leq -\boldsymbol{w}^T\tilde{\boldsymbol{r}}\} \leq \epsilon\right\}, \tag{2}$$

where $\mathbb{P}$ denotes the distribution of the asset returns $\tilde{\boldsymbol{r}}$.

In this paper we investigate portfolio optimization problems of the type

$$\begin{aligned} \underset{\boldsymbol{w} \in \mathbb{R}^m}{\text{minimize}} \quad & \mathrm{VaR}_\epsilon(\boldsymbol{w}) \\ \text{subject to} \quad & \boldsymbol{w} \in \mathcal{W}, \end{aligned} \tag{3}$$

where $\mathcal{W} \subseteq \mathbb{R}^m$ denotes the set of admissible portfolios. The inclusion $\boldsymbol{w} \in \mathcal{W}$ usually implies the budget constraint $\boldsymbol{w}^T\boldsymbol{e} = 1$ (where $\boldsymbol{e}$ denotes the vector of 1s). Optionally, the set $\mathcal{W}$ may account for bounds on the allocation vector $\boldsymbol{w}$ and/or a constraint enforcing a minimum expected portfolio return. In this paper we only require that $\mathcal{W}$ must be a convex polyhedron.

By using (2), the VaR optimization model (3) can be reformulated as

$$
\begin{aligned}
\underset{\boldsymbol{w}\in\mathbb{R}^m,\gamma\in\mathbb{R}}{\text{minimize}} \quad & \gamma \\
\text{subject to} \quad & \mathbb{P}\{\gamma + \boldsymbol{w}^T\tilde{\boldsymbol{r}} \geq 0\} \geq 1-\epsilon \\
& \boldsymbol{w} \in \mathcal{W},
\end{aligned} \tag{4}
$$

which constitutes a chance-constrained stochastic program. Optimization problems of this kind are usually difficult to solve since they tend to have non-convex or even disconnected feasible sets. Furthermore, the evaluation of the chance constraint requires precise knowledge of the probability distribution of the asset returns, which is rarely available in practice.

2.1 Two Analytical Approximations of Value-at-Risk

In order to overcome the computational difficulties and to account for the lack of knowledge about the distribution of the asset returns, the objective function in (3) must usually be approximated. Most existing approximation techniques fall into one of two main categories: *non-parametric approaches* which approximate the asset return distribution by a discrete (sampled or empirical) distribution and *parametric approaches* which approximate the asset return distribution by the best fitting member of a parametric family of continuous distributions. We now give a brief overview of two analytical VaR approximation schemes that are of particular relevance for our purposes.

Both in the financial industry as well as in the academic literature, it is frequently assumed that the asset returns $\tilde{\boldsymbol{r}}$ are governed by a Gaussian distribution with given mean vector $\boldsymbol{\mu_r} \in \mathbb{R}^m$ and covariance matrix $\boldsymbol{\Sigma_r} \in \mathbb{S}^m$. This assumption has the advantage that the VaR can be calculated analytically as

$$
\mathrm{VaR}_\epsilon(\boldsymbol{w}) = -\boldsymbol{\mu}_{\boldsymbol{r}}^T\boldsymbol{w} - \Phi^{-1}(\epsilon)\sqrt{\boldsymbol{w}^T\boldsymbol{\Sigma_r}\boldsymbol{w}}, \tag{5}
$$

where Φ is the standard normal distribution function. This model is sometimes referred to as *Normal VaR* (see, e.g., [19]). In practice, the distribution of the asset returns often fails to be Gaussian. In these cases, (5) can still be used as an approximation. However, it may lead to gross *underestimation* of the actual portfolio VaR when the true portfolio return distribution is leptokurtic or heavily skewed, as is the case for portfolios containing options.

To avoid unduly optimistic risk assessments, El Ghaoui *et al.* [11] suggest a conservative (that is, pessimistic) approximation for VaR under the assumption that only the mean values and covariance matrix of the asset returns are known. Let $\mathcal{P}_{\boldsymbol{r}}$ be the set of all probability distributions on $\mathbb{R}^m$ with mean value $\boldsymbol{\mu}_{\boldsymbol{r}}$ and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{r}}$. We emphasize that $\mathcal{P}_{\boldsymbol{r}}$ contains also distributions which exhibit considerable skewness, so long as they match the given mean vector and covariance matrix. The *Worst-Case Value-at-Risk* for portfolio $\boldsymbol{w}$ is now defined as

$$\mathrm{WCVaR}_\epsilon(\boldsymbol{w}) = \min\left\{\gamma \;:\; \sup_{\mathbb{P}\in\mathcal{P}_{\boldsymbol{r}}} \mathbb{P}\{\gamma \leq -\boldsymbol{w}^T\tilde{\boldsymbol{r}}\} \leq \epsilon\right\}. \tag{6}$$

El Ghaoui *et al.* demonstrate that WCVaR has the closed form expression

$$\mathrm{WCVaR}_\epsilon(\boldsymbol{w}) = -\boldsymbol{\mu}_{\boldsymbol{r}}^T\boldsymbol{w} + \kappa(\epsilon)\sqrt{\boldsymbol{w}^T\boldsymbol{\Sigma}_{\boldsymbol{r}}\boldsymbol{w}}, \tag{7}$$

where $\kappa(\epsilon) = \sqrt{(1-\epsilon)/\epsilon}$. WCVaR represents a tight approximation for VaR in the sense that there exists a worst-case distribution $\mathbb{P}^* \in \mathcal{P}_{\boldsymbol{r}}$ such that VaR with respect to $\mathbb{P}^*$ is equal to WCVaR.

When using WCVaR instead of VaR as a risk measure, we end up with the portfolio optimization problem

$$\begin{aligned} &\underset{\boldsymbol{w}\in\mathbb{R}^m}{\text{minimize}} && -\boldsymbol{\mu}_{\boldsymbol{r}}^T\boldsymbol{w} + \kappa(\epsilon)\left\|\boldsymbol{\Sigma}_{\boldsymbol{r}}^{1/2}\boldsymbol{w}\right\|_2 \\ &\text{subject to} && \boldsymbol{w}\in\mathcal{W}, \end{aligned} \tag{8}$$

which represents a second-order cone program that is amenable to efficient numerical solution procedures.

2.2 Robust Optimization Perspective on Worst-Case VaR

Consider the following *uncertain* linear program.

$$\begin{aligned} &\underset{\boldsymbol{w}\in\mathbb{R}^m,\gamma\in\mathbb{R}}{\text{minimize}} && \gamma \\ &\text{subject to} && \gamma + \boldsymbol{w}^T\tilde{\boldsymbol{r}} \geq 0 \\ & && \boldsymbol{w}\in\mathcal{W} \end{aligned} \tag{9}$$

Since the asset return vector is uncertain, this model essentially represents a whole family of optimization problems, one for each possible realization of $\tilde{\boldsymbol{r}}$. Therefore, (9) fails to provide a unique implementable investment decision. One way to disambiguate this model is to require that the explicit inequality constraint in (9) is satisfied with a given probability. By using this approach, we recover the chance-constrained stochastic program (4). Robust optimization [5, 6] pursues a different approach to disambiguate the model. The idea is to select a decision which is optimal with respect to the worst-case realization of $\tilde{\boldsymbol{r}}$ within a prescribed *uncertainty set* $\mathcal{U}$. This set may cover only a subset of all possible realizations of $\tilde{\boldsymbol{r}}$ and is chosen by the modeller. The *robust counterpart* of problem (9) is then defined as

$$
\begin{aligned}
\underset{\boldsymbol{w}\in\mathbb{R}^m,\gamma\in\mathbb{R}}{\text{minimize}} \quad & \gamma \\
\text{subject to} \quad & \gamma + \boldsymbol{w}^T\boldsymbol{r} \geq 0 \quad \forall \boldsymbol{r} \in \mathcal{U} \\
& \boldsymbol{w} \in \mathcal{W}.
\end{aligned}
\tag{10}
$$

The shape of the uncertainty set $\mathcal{U}$ should reflect the modeller's knowledge about the asset return distribution, e.g., full or partial information about the support and certain moments of the random vector $\tilde{\boldsymbol{r}}$. Moreover, the size of $\mathcal{U}$ determines the degree to which the user wants to safeguard feasibility of the corresponding explicit inequality constraint. The semi-infinite constraint in the robust counterpart (10) is therefore closely related to the chance constraint in the stochastic program (4). For a large class of convex uncertainty sets, the semi-infinite constraint in the robust counterpart can be reformulated in terms of a small number of tractable (i.e., linear, second-order conic, or semidefinite) constraints [5, 6].

An uncertainty set that enjoys wide popularity in the robust optimization literature is the *ellipsoidal set*,

$$
\mathcal{U} = \{\boldsymbol{r} \in \mathbb{R}^m \ : \ (\boldsymbol{r} - \boldsymbol{\mu}_{\boldsymbol{r}})^T \boldsymbol{\Sigma}_{\boldsymbol{r}}^{-1} (\boldsymbol{r} - \boldsymbol{\mu}_{\boldsymbol{r}}) \leq \delta^2\},
$$

which is defined in terms of the mean vector $\boldsymbol{\mu}_{\boldsymbol{r}}$ and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{r}}$ of the asset returns as well as a size parameter δ. By conic duality it can be shown that the following equivalence holds for any fixed $(\boldsymbol{w}, \gamma) \in \mathcal{W} \times \mathbb{R}$.

$$
\gamma + \boldsymbol{w}^T\boldsymbol{r} \geq 0 \quad \forall \boldsymbol{r} \in \mathcal{U} \iff -\boldsymbol{\mu}_{\boldsymbol{r}}^T\boldsymbol{w} + \delta \left\| \boldsymbol{\Sigma}_{\boldsymbol{r}}^{1/2}\boldsymbol{w} \right\|_2 \leq \gamma
\tag{11}
$$

Problem (10) can therefore be reformulated as the following second-order cone program.

$$\begin{aligned} \underset{\boldsymbol{w}\in\mathbb{R}^m}{\text{minimize}} \quad & -\boldsymbol{\mu}_{\boldsymbol{r}}^T\boldsymbol{w} + \delta\left\|\boldsymbol{\Sigma}_{\boldsymbol{r}}^{1/2}\boldsymbol{w}\right\|_2 \\ \text{subject to} \quad & \boldsymbol{w}\in\mathcal{W} \end{aligned} \tag{12}$$

By comparing (8) and (12), El Ghaoui *et al.* [11] noticed that optimizing WCVaR at level ϵ is equivalent to solving the robust optimization problem (10) under an ellipsoidal uncertainty set with size parameter $\delta = \kappa(\epsilon)$, see also Natarajan *et al.* [19]. This uncertainty set will henceforth be denoted by $\mathcal{U}_\epsilon$.

In this paper we extend the WCVaR model (7) and the equivalent robust optimization model (10) to situations in which there are non-linear relationships between the asset returns, as is the case in the presence of derivatives.

3 Worst-Case VaR for Derivative Portfolios

From now on assume that our market consists of $n \leq m$ *basic assets* and $m-n$ *derivatives.* We partition the asset return vector as $\tilde{\boldsymbol{r}} = (\tilde{\boldsymbol{\xi}}, \tilde{\boldsymbol{\eta}})$, where the $\mathbb{R}^n$-valued random vector $\tilde{\boldsymbol{\xi}}$ and $\mathbb{R}^{m-n}$-valued random vector $\tilde{\boldsymbol{\eta}}$ denote the basic asset returns and derivative returns, respectively.

To approximate the VaR of some portfolio $\boldsymbol{w}\in\mathcal{W}$ containing derivatives, one can principally still use the WCVaR model (7), which has the advantage of computational tractability and accounts for the absence of distributional information beyond first- and second-order moments. However, WCVaR is not a suitable approximation for VaR in the presence of derivatives due to the following reasons.

The first- and second-order moments of the derivative returns, which must be supplied to the WCVaR model, are difficult to estimate reliably from historical data, see, e.g., [9]. Note that the moments of the basic assets returns (i.e., stocks and bonds etc.) can usually be estimated more accurately due to the availability of longer historical time series. However, even if the means and covariances of the derivative returns were precisely known, WCVaR would still provide a poor approximation of the actual portfolio VaR because it disregards known perfect dependencies between the derivative returns and their underlying asset returns. In fact, the returns of the derivatives are uniquely determined by the returns of the underlying assets, that is, there exists a (typically non-linear) measurable function $f : \mathbb{R}^n \to \mathbb{R}^m$ such that $\tilde{\boldsymbol{r}} = f(\tilde{\boldsymbol{\xi}})$.[1] Put differently, the derivatives introduce no new uncertainties in the market; their returns are uncertain only because the underlying asset returns are uncertain. The function f can usually be inferred reliably from contractual specifications (payoff functions) or pricing models of the derivatives.

In summary, WCVaR provides a conservative approximation to the actual VaR. However, it relies on first- and second-order moments of the derivative returns, which are difficult to obtain in practice, but disregards the perfect dependencies captured by the function f, which is typically known.

When f is non-linear, WCVaR tends to severely *overestimate* the actual VaR since the covariance matrix $\boldsymbol{\Sigma_r}$ accounts only for *linear* dependencies. The robust optimization perspective on WCVaR manifests this drawback geometrically. Recall that the ellipsoidal uncertainty set $\mathcal{U}_\epsilon$ introduced in Section 2.2 is symmetrically oriented around the mean vector $\boldsymbol{\mu_r}$. If the underlying assets of the derivatives have approximately symmetrically distributed returns, then the derivative returns are heavily skewed. An ellipsoidal uncertainty set fails to capture this asymmetry. This geometric argument supports our conjecture that WCVaR provides a poor (over-pessimistic) VaR estimate when the portfolio contains derivatives.

In the remainder of the paper we assume to know the first- and second-order moments of the basic asset returns as well as the function f, which captures the non-linear dependencies between the basic asset and derivative returns. In contrast, we assume that the moments of the derivative returns are unknown.

In the next sections we derive generic Worst-Case Value-at-Risk models that explicitly account for non-linear (piecewise linear or quadratic) relationships between the asset returns. These new models provide tighter approximations for the actual VaR of portfolios containing derivatives than the WCVaR model, which relies solely on moment information.

Below, we will always denote the mean vector and the covariance matrix of the basic asset returns by $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, respectively. Without loss of generality we assume that $\boldsymbol{\Sigma}$ is strictly positive definite.

[1]For ease of exposition, we assume that the returns of the derivative underliers are the only risk factors determining the option returns.

4 Worst-Case Polyhedral VaR Optimization

In this section we describe a Worst-Case VaR model that explicitly accounts for piecewise linear relationships between option returns and their underlying asset returns. We show that this model can be cast as a tractable second-order cone program and establish its equivalence to a robust optimization model that admits an intuitive interpretation.

4.1 Piecewise Linear Portfolio Model

We now assume that the $m - n$ derivatives in our market are *European-style call and/or put options* derived from the basic assets. All these options are assumed to mature at the end of the investment horizon, that is, at time T.

For ease of exposition, we partition the allocation vector as $\boldsymbol{w} = (\boldsymbol{w}^{\boldsymbol{\xi}}, \boldsymbol{w}^{\boldsymbol{\eta}})$, where $\boldsymbol{w}^{\boldsymbol{\xi}} \in \mathbb{R}^n$ and $\boldsymbol{w}^{\boldsymbol{\eta}} \in \mathbb{R}^{m-n}$ denote the percentage allocations in the basic assets and options, respectively. In this section we forbid short-sales of options, that is, we assume that the inclusion $\boldsymbol{w} \in \mathcal{W}$ implies $\boldsymbol{w}^{\boldsymbol{\eta}} \geq \mathbf{0}$. Recall that the set $\mathcal{W}$ of admissible portfolios was assumed to be a convex polyhedron.

We now derive an explicit representation for f by using the known payoff functions of the basic assets as well as the European call and put options. Since the first n components of $\tilde{\boldsymbol{r}}$ represent the basic asset returns $\tilde{\boldsymbol{\xi}}$, we have $f_j(\tilde{\boldsymbol{\xi}}) = \tilde{\xi}_j$ for $j = 1, \ldots, n$. Next, we investigate the option returns $\tilde{r}_j$ for $j = n+1, \ldots, m$. Let asset j be a call option with strike price k_j on the basic asset i, and denote the return and the initial price of the option by $\tilde{r}_j$ and c_j, respectively. If s_i denotes the initial price of asset i, then its end-of-period price amounts to $s_i(1 + \tilde{\xi}_i)$. We can now explicitly express the return $\tilde{r}_j$ as a convex piecewise linear function of $\tilde{\xi}_i$,

$$\begin{aligned} f_j(\tilde{\boldsymbol{\xi}}) &= \frac{1}{c_j} \max\left\{0, s_i(1 + \tilde{\xi}_i) - k_j\right\} - 1 \\ &= \max\left\{-1, a_j + b_j \tilde{\xi}_i - 1\right\}, \quad \text{where} \quad a_j = \frac{s_i - k_j}{c_j} \quad \text{and} \quad b_j = \frac{s_i}{c_j}. \end{aligned} \tag{13a}$$

Similarly, if asset j is a put option with price p_j and strike price k_j on the basic asset i, then its return $\tilde{r}_j$ is representable as a different convex piecewise linear function,

$$f_j(\tilde{\boldsymbol{\xi}}) = \max\left\{-1, a_j + b_j \tilde{\xi}_i - 1\right\}, \quad \text{where} \quad a_j = \frac{k_j - s_i}{p_j} \quad \text{and} \quad b_j = -\frac{s_i}{p_j}. \tag{13b}$$

Using the above notation, we can write the vector of asset returns $\tilde{\boldsymbol{r}}$ compactly as

$$\tilde{\boldsymbol{r}} = f(\tilde{\boldsymbol{\xi}}) = \begin{pmatrix} \tilde{\boldsymbol{\xi}} \\ \max\left\{-\boldsymbol{e}, \boldsymbol{a} + \mathbf{B}\tilde{\boldsymbol{\xi}} - \boldsymbol{e}\right\} \end{pmatrix}, \tag{14}$$

where $\boldsymbol{a} \in \mathbb{R}^{m-n}$, $\mathbf{B} \in \mathbb{R}^{(m-n)\times n}$ are known constants determined through (13a) and (13b), $\boldsymbol{e} \in \mathbb{R}^{m-n}$ is the vector of 1s, and 'max' denotes the component-wise maximization operator. Thus, the return $\tilde{r}_p$ of some portfolio $\boldsymbol{w} \in \mathcal{W}$ can be expressed as

$$\begin{aligned}\tilde{r}_p &= \boldsymbol{w}^T \tilde{\boldsymbol{r}} = (\boldsymbol{w}^{\boldsymbol{\xi}})^T \tilde{\boldsymbol{\xi}} + (\boldsymbol{w}^{\boldsymbol{\eta}})^T \tilde{\boldsymbol{\eta}} \\ &= \boldsymbol{w}^T f(\tilde{\boldsymbol{\xi}}) = (\boldsymbol{w}^{\boldsymbol{\xi}})^T \tilde{\boldsymbol{\xi}} + (\boldsymbol{w}^{\boldsymbol{\eta}})^T \max\left\{-\boldsymbol{e}, \boldsymbol{a} + \mathbf{B}\tilde{\boldsymbol{\xi}} - \boldsymbol{e}\right\}. \end{aligned} \tag{15}$$

4.2 Worst-Case Polyhedral VaR Model

For any portfolio $\boldsymbol{w} \in \mathcal{W}$, we define the *Worst-Case Polyhedral VaR* (WCPVaR) as

$$\begin{aligned}\mathrm{WCPVaR}_\epsilon(\boldsymbol{w}) &= \min\left\{\gamma \ : \ \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\left\{\gamma \le -\boldsymbol{w}^T f(\tilde{\boldsymbol{\xi}})\right\} \le \epsilon\right\} \\ &= \min\left\{\gamma \ : \ \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\left\{\gamma \le -(\boldsymbol{w}^{\boldsymbol{\xi}})^T \tilde{\boldsymbol{\xi}} - (\boldsymbol{w}^{\boldsymbol{\eta}})^T \max\left\{-\boldsymbol{e}, \boldsymbol{a} + \mathbf{B}\tilde{\boldsymbol{\xi}} - \boldsymbol{e}\right\}\right\} \le \epsilon\right\}, \end{aligned} \tag{16}$$

where $\mathcal{P}$ denotes the set of all probability distributions of the *basic* asset returns $\tilde{\boldsymbol{\xi}}$ with a given mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. WCPVaR provides a tight conservative approximation for the VaR of a portfolio whose return constitutes a convex piecewise linear (i.e., polyhedral) function of the basic asset returns.

In the remainder of this section we derive a manifestly tractable representation for WCPVaR. As a first step to achieve this goal, we simplify the maximization problem

$$\sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\left\{\gamma \le -(\boldsymbol{w}^{\boldsymbol{\xi}})^T \tilde{\boldsymbol{\xi}} - (\boldsymbol{w}^{\boldsymbol{\eta}})^T \max\left\{-\boldsymbol{e}, \boldsymbol{a} + \mathbf{B}\tilde{\boldsymbol{\xi}} - \boldsymbol{e}\right\}\right\}, \tag{17}$$

which can be identified as the subordinate optimization problem in (16).

For some fixed portfolio $\boldsymbol{w} \in \mathcal{W}$ and $\gamma \in \mathbb{R}$, we define the set $\mathcal{S}_\gamma \subseteq \mathbb{R}^n$ as

$$\mathcal{S}_\gamma = \{\boldsymbol{\xi} \in \mathbb{R}^n \ : \ \gamma + (\boldsymbol{w}^{\boldsymbol{\xi}})^T \boldsymbol{\xi} + (\boldsymbol{w}^{\boldsymbol{\eta}})^T \max\{-\boldsymbol{e}, \boldsymbol{a} + \mathbf{B}\boldsymbol{\xi} - \boldsymbol{e}\} \le 0\}.$$

For any $\boldsymbol{\xi} \in \mathbb{R}^n$ and nonnegative $\boldsymbol{w}^{\boldsymbol{\eta}} \in \mathbb{R}^{m-n}$ we have

$$\begin{aligned}(\boldsymbol{w}^{\boldsymbol{\eta}})^T \max\{-\boldsymbol{e}, \boldsymbol{a} + \mathbf{B}\boldsymbol{\xi} - \boldsymbol{e}\} &= \min_{\boldsymbol{g} \in \mathbb{R}^{m-n}} \left\{\boldsymbol{g}^T \boldsymbol{w}^{\boldsymbol{\eta}} \ : \ \boldsymbol{g} \geq -\boldsymbol{e}, \ \boldsymbol{g} \geq \boldsymbol{a} + \mathbf{B}\boldsymbol{\xi} - \boldsymbol{e}\right\} \\ &= \max_{\boldsymbol{y} \in \mathbb{R}^{m-n}} \left\{\boldsymbol{y}^T(\boldsymbol{a} + \mathbf{B}\boldsymbol{\xi}) - \boldsymbol{e}^T \boldsymbol{w}^{\boldsymbol{\eta}} \ : \ \mathbf{0} \leq \boldsymbol{y} \leq \boldsymbol{w}^{\boldsymbol{\eta}}\right\},\end{aligned}$$

where the second equality follows from strong linear programming duality. Thus, the set $\mathcal{S}_\gamma$ can be written as

$$\mathcal{S}_\gamma = \left\{\boldsymbol{\xi} \in \mathbb{R}^n \ : \ \max_{\mathbf{0} \leq \boldsymbol{y} \leq \boldsymbol{w}^{\boldsymbol{\eta}}} \left\{\gamma + (\boldsymbol{w}^{\boldsymbol{\xi}})^T \boldsymbol{\xi} + \boldsymbol{y}^T(\boldsymbol{a} + \mathbf{B}\boldsymbol{\xi}) - \boldsymbol{e}^T \boldsymbol{w}^{\boldsymbol{\eta}}\right\} \leq 0\right\}. \tag{18}$$

The optimal value of problem (17) can be obtained by solving the worst-case probability problem

$$\pi_{\mathrm{wc}} = \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{P}\{\tilde{\boldsymbol{\xi}} \in \mathcal{S}_\gamma\}. \tag{19}$$

The next lemma reviews a general result about worst-case probability problems and will play a key role in many of the following derivations. The proof is due to Calafiore *et al.* [8] but is repeated in Appendix A.1 to keep this paper self-contained.

Lemma 4.1 *Let $\mathcal{S} \subseteq \mathbb{R}^n$ be any Borel measurable set (which is not necessarily convex), and define the worst-case probability π_{wc} as*

$$\pi_{\mathrm{wc}} = \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{P}\{\tilde{\boldsymbol{\xi}} \in \mathcal{S}\}, \tag{20}$$

where $\mathcal{P}$ is the set of all probability distributions of $\tilde{\boldsymbol{\xi}}$ with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma} \succ \mathbf{0}$. Then,

$$\pi_{\mathrm{wc}} = \inf_{\mathbf{M} \in \mathbb{S}^{n+1}} \left\{\langle \boldsymbol{\Omega}, \mathbf{M}\rangle \ : \ \mathbf{M} \succcurlyeq \mathbf{0}, \quad \left[\boldsymbol{\xi}^T \ 1\right] \mathbf{M} \left[\boldsymbol{\xi}^T \ 1\right]^T \geq 1 \quad \forall \boldsymbol{\xi} \in \mathcal{S}\right\}, \tag{21}$$

where

$$\boldsymbol{\Omega} = \begin{bmatrix} \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T & \boldsymbol{\mu} \\ \boldsymbol{\mu}^T & 1 \end{bmatrix} \tag{22}$$

is the second-order moment matrix of $\tilde{\boldsymbol{\xi}}$.

Lemma 4.1 enables us to reformulate the worst-case probability problem (19) as

$$\begin{aligned} \pi_{\rm wc} = \inf_{\mathbf{M}\in\mathbb{S}^{n+1}} \quad & \langle \boldsymbol{\Omega}, \mathbf{M}\rangle \\ \text{s.t.} \quad & \left[\boldsymbol{\xi}^T \ 1\right] \mathbf{M} \left[\boldsymbol{\xi}^T \ 1\right]^T \geq 1 \ \forall \boldsymbol{\xi} : \max_{\mathbf{0}\leq \boldsymbol{y} \leq \boldsymbol{w}^{\boldsymbol{\eta}}} \{\gamma + (\boldsymbol{w}^{\boldsymbol{\xi}})^T\boldsymbol{\xi} + \boldsymbol{y}^T(\boldsymbol{a} + \mathbf{B}\boldsymbol{\xi}) - \boldsymbol{e}^T\boldsymbol{w}^{\boldsymbol{\eta}}\} \leq 0 \\ & \mathbf{M} \succcurlyeq \mathbf{0}. \end{aligned} \tag{23}$$

We now recall the non-linear Farkas Lemma, which is a fundamental theorem of alternatives in convex analysis and will enable us to simplify the optimization problem (23), see, e.g., [21, Theorem 2.1] and the references therein.

Lemma 4.2 (Farkas Lemma) *Let $f_0, \ldots, f_p : \mathbb{R}^n \to \mathbb{R}$ be convex functions, and assume that there exists a strictly feasible point $\bar{\boldsymbol{\xi}}$ with $f_i(\bar{\boldsymbol{\xi}}) < 0$, $i = 1, \ldots, p$. Then, $f_0(\boldsymbol{\xi}) \geq 0$ for all $\boldsymbol{\xi}$ with $f_i(\bar{\boldsymbol{\xi}}) \leq 0$, $i = 1, \ldots, p$, if and only if there exist constants $\tau_i \geq 0$ such that*

$$f_0(\boldsymbol{\xi}) + \sum_{i=1}^{p} \tau_i f_i(\boldsymbol{\xi}) \geq 0 \quad \forall \boldsymbol{\xi} \in \mathbb{R}^n.$$

We will now argue that problem (23) can be reformulated as follows.

$$\begin{aligned} \inf \quad & \langle \boldsymbol{\Omega}, \mathbf{M}\rangle \\ \text{s.t.} \quad & \mathbf{M} \in \mathbb{S}^{n+1}, \quad \tau \in \mathbb{R}, \quad \mathbf{M} \succcurlyeq \mathbf{0}, \quad \tau \geq 0 \\ & \left[\boldsymbol{\xi}^T \ 1\right] \mathbf{M} \left[\boldsymbol{\xi}^T \ 1\right]^T - 1 + 2\tau \left(\max_{\mathbf{0}\leq \boldsymbol{y} \leq \boldsymbol{w}^{\boldsymbol{\eta}}} \{\gamma + (\boldsymbol{w}^{\boldsymbol{\xi}})^T\boldsymbol{\xi} + \boldsymbol{y}^T(\boldsymbol{a} + \mathbf{B}\boldsymbol{\xi}) - \boldsymbol{e}^T\boldsymbol{w}^{\boldsymbol{\eta}}\} \right) \geq 0 \ \ \forall \boldsymbol{\xi} \in \mathbb{R}^n \end{aligned} \tag{24}$$

For ease of exposition, we first first define

$$h = \min_{\boldsymbol{\xi}\in\mathbb{R}^n} \max_{\mathbf{0}\leq \boldsymbol{y} \leq \boldsymbol{w}^{\boldsymbol{\eta}}} \{\gamma + (\boldsymbol{w}^{\boldsymbol{\xi}})^T\boldsymbol{\xi} + \boldsymbol{y}^T(\boldsymbol{a} + \mathbf{B}\boldsymbol{\xi}) - \boldsymbol{e}^T\boldsymbol{w}^{\boldsymbol{\eta}}\}.$$

The equivalence of (23) and (24) is proved case by case. Assume first that $h < 0$. Then, the equivalence follows from the Farkas Lemma. Assume next that $h > 0$. Then, the semi-infinite constraint in (23) becomes redundant and, since $\boldsymbol{\Omega} \succ \mathbf{0}$, the optimal solution of (23) is given by $\mathbf{M} = \mathbf{0}$ with a corresponding optimal value of 0. The optimal value of problem (24) is also equal to 0. Indeed, by choosing $\tau = 1/h$, the semi-infinite constraint in (24) is satisfied independently of $\mathbf{M}$. Finally, assume that $h = 0$. In this degenerate case the equivalence follows from a standard continuity argument. Details are omitted for brevity of exposition.

It can be seen that since $\tau \geq 0$, the semi-infinite constraint in (24) is equivalent to the assertion that there exists some $\mathbf{0} \leq \boldsymbol{y} \leq \boldsymbol{w}^{\boldsymbol{\eta}}$ with

$$[\boldsymbol{\xi}^T\ 1]\,\mathbf{M}\,[\boldsymbol{\xi}^T\ 1]^T - 1 + 2\tau\left(\gamma + (\boldsymbol{w}^{\boldsymbol{\xi}})^T\boldsymbol{\xi} + \boldsymbol{y}^T(\boldsymbol{a} + \mathbf{B}\boldsymbol{\xi}) - \boldsymbol{e}^T\boldsymbol{w}^{\boldsymbol{\eta}}\right) \geq 0 \qquad \forall \boldsymbol{\xi} \in \mathbb{R}^n.$$

This semi-infinite constraint can be written as

$$\begin{aligned} &\begin{bmatrix} \boldsymbol{\xi} \\ 1 \end{bmatrix}^T \left(\mathbf{M} + \begin{bmatrix} \mathbf{0} & \tau(\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T\boldsymbol{y}) \\ \tau(\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T\boldsymbol{y})^T & -1 + 2\tau(\gamma + \boldsymbol{y}^T\boldsymbol{a} - \boldsymbol{e}^T\boldsymbol{w}^{\boldsymbol{\eta}}) \end{bmatrix}\right) \begin{bmatrix} \boldsymbol{\xi} \\ 1 \end{bmatrix} \geq \mathbf{0} \quad \forall \boldsymbol{\xi} \in \mathbb{R}^n \\ \iff \quad &\mathbf{M} + \begin{bmatrix} \mathbf{0} & \tau(\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T\boldsymbol{y}) \\ \tau(\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T\boldsymbol{y})^T & -1 + 2\tau(\gamma + \boldsymbol{y}^T\boldsymbol{a} - \boldsymbol{e}^T\boldsymbol{w}^{\boldsymbol{\eta}}) \end{bmatrix} \succcurlyeq \mathbf{0}. \end{aligned}$$

Thus, the worst-case probability problem (23) can equivalently be formulated as

$$\begin{aligned} \pi_{\mathrm{wc}} = \inf \quad & \langle \boldsymbol{\Omega}, \mathbf{M} \rangle \\ \text{s.t.} \quad & \mathbf{M} \in \mathbb{S}^{n+1}, \quad \boldsymbol{y} \in \mathbb{R}^{m-n}, \quad \tau \in \mathbb{R} \\ & \mathbf{M} \succcurlyeq \mathbf{0}, \quad \tau \geq 0, \quad \mathbf{0} \leq \boldsymbol{y} \leq \boldsymbol{w}^{\boldsymbol{\eta}} \\ & \mathbf{M} + \begin{bmatrix} \mathbf{0} & \tau(\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T\boldsymbol{y}) \\ \tau(\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T\boldsymbol{y})^T & -1 + 2\tau(\gamma + \boldsymbol{y}^T\boldsymbol{a} - \boldsymbol{e}^T\boldsymbol{w}^{\boldsymbol{\eta}}) \end{bmatrix} \succcurlyeq \mathbf{0}. \end{aligned} \tag{25}$$

By using (25) we can express WCPVaR in (16) as the optimal value of the following minimization problem.

$$\begin{aligned} \mathrm{WCPVaR}_\epsilon(\boldsymbol{w}) = \inf \quad & \gamma \\ \text{s.t.} \quad & \mathbf{M} \in \mathbb{S}^{n+1}, \quad \boldsymbol{y} \in \mathbb{R}^{m-n}, \quad \tau \in \mathbb{R}, \quad \gamma \in \mathbb{R} \\ & \langle \boldsymbol{\Omega}, \mathbf{M} \rangle \leq \epsilon, \quad \mathbf{M} \succcurlyeq \mathbf{0}, \quad \tau \geq 0, \quad \mathbf{0} \leq \boldsymbol{y} \leq \boldsymbol{w}^{\boldsymbol{\eta}} \\ & \mathbf{M} + \begin{bmatrix} \mathbf{0} & \tau(\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T\boldsymbol{y}) \\ \tau(\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T\boldsymbol{y})^T & -1 + 2\tau(\gamma + \boldsymbol{y}^T\boldsymbol{a} - \boldsymbol{e}^T\boldsymbol{w}^{\boldsymbol{\eta}}) \end{bmatrix} \succcurlyeq \mathbf{0} \end{aligned} \tag{26}$$

Problem (26) is non-convex due to the bilinear terms in the matrix inequality constraint. It can easily be shown that $\langle \boldsymbol{\Omega}, \mathbf{M} \rangle \geq 1$ for any feasible point with vanishing τ-component. However, since $\epsilon < 1$, this is in conflict with the constraint $\langle \mathbf{\Omega}, \mathbf{M} \rangle \leq \epsilon$. We thus conclude that no feasible point can have a vanishing τ-component. This allows us to divide the matrix inequality in problem (26) by τ. Subsequently we perform variable substitutions in which we replace $1/\tau$ by τ and $\mathbf{M}/\tau$ by $\mathbf{M}$. This yields the following reformulation of problem (26).

$$
\begin{aligned}
\mathrm{WCPVaR}_\epsilon(\boldsymbol{w}) = \inf \quad & \gamma \\
\text{s.t.} \quad & \mathbf{M} \in \mathbb{S}^{n+1}, \quad \boldsymbol{y} \in \mathbb{R}^{m-n}, \quad \tau \in \mathbb{R}, \quad \gamma \in \mathbb{R} \\
& \langle \mathbf{\Omega}, \mathbf{M} \rangle \leq \tau\epsilon, \quad \mathbf{M} \succcurlyeq \mathbf{0}, \quad \tau \geq 0, \quad \mathbf{0} \leq \boldsymbol{y} \leq \boldsymbol{w}^{\boldsymbol{\eta}} \\
& \mathbf{M} + \begin{bmatrix} \mathbf{0} & \boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T \boldsymbol{y} \\ (\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T \boldsymbol{y})^T & -\tau + 2(\gamma + \boldsymbol{y}^T \boldsymbol{a} - \boldsymbol{e}^T \boldsymbol{w}^{\boldsymbol{\eta}}) \end{bmatrix} \succcurlyeq \mathbf{0}
\end{aligned} \tag{27}
$$

Observe that (27) constitutes a semidefinite program (SDP) that can be used to efficiently compute the WCPVaR of a given portfolio $\boldsymbol{w} \in \mathcal{W}$. However, it would be desirable to obtain an equivalent second-order cone program (SOCP) because SOCPs exhibit better scalability properties than SDPs [2]. Theorem 4.1 shows that such a reformulation exists.

Theorem 4.1 *Problem* (27) *can be reformulated as*

$$
\mathrm{WCPVaR}_\epsilon(\boldsymbol{w}) = \min_{\mathbf{0} \leq \boldsymbol{g} \leq \boldsymbol{w}^{\boldsymbol{\eta}}} \; -\boldsymbol{\mu}^T(\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T \boldsymbol{g}) + \kappa(\epsilon) \left\| \mathbf{\Sigma}^{1/2} (\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T \boldsymbol{g}) \right\|_2 - \boldsymbol{a}^T \boldsymbol{g} + \boldsymbol{e}^T \boldsymbol{w}^{\boldsymbol{\eta}}, \tag{28}
$$

which constitutes a tractable SOCP.

Proof: The proof follows a similar reasoning as in [11, Theorem 1] and is therefore relegated to Appendix A.2.

Remark 4.1 *In the absence of derivatives, that is, when the market only contains basic assets, then $m = n$ and $\boldsymbol{w} = \boldsymbol{w}^{\boldsymbol{\xi}}$. In this special case we obtain*

$$
\mathrm{WCPVaR}_\epsilon(\boldsymbol{w}) = -\boldsymbol{\mu}^T \boldsymbol{w} + \kappa(\epsilon) \left\| \mathbf{\Sigma}^{1/2} \boldsymbol{w} \right\|_2 = \mathrm{WCVaR}_\epsilon(\boldsymbol{w}).
$$

Thus, the WCPVaR *model encapsulates the* WCVaR *model* (7) *as a special case.*

The problem of minimizing the WCVaR of a portfolio containing European options can now be conservatively approximated by

$$\begin{array}{ll}\underset{\boldsymbol{w}\in\mathbb{R}^m}{\text{minimize}} & \text{WCPVaR}_\epsilon(\boldsymbol{w}) \\ \text{subject to} & \boldsymbol{w}\in\mathcal{W},\end{array}$$

which is equivalent to the tractable SOCP

$$\begin{array}{ll}\text{minimize} & \gamma \\ \text{subject to} & \boldsymbol{w}^{\boldsymbol{\xi}}\in\mathbb{R}^n, \quad \boldsymbol{w}^{\boldsymbol{\eta}}\in\mathbb{R}^{m-n}, \quad \boldsymbol{g}\in\mathbb{R}^{m-n}, \quad \gamma\in\mathbb{R} \\ & -\boldsymbol{\mu}^T(\boldsymbol{w}^{\boldsymbol{\xi}}+\mathbf{B}^T\boldsymbol{g})+\kappa(\epsilon)\left\|\boldsymbol{\Sigma}^{1/2}(\boldsymbol{w}^{\boldsymbol{\xi}}+\mathbf{B}^T\boldsymbol{g})\right\|_2-\boldsymbol{a}^T\boldsymbol{g}+\boldsymbol{e}^T\boldsymbol{w}^{\boldsymbol{\eta}}\leq\gamma \\ & \mathbf{0}\leq\boldsymbol{g}\leq\boldsymbol{w}^{\boldsymbol{\eta}}, \quad \boldsymbol{w}=(\boldsymbol{w},\boldsymbol{w}^{\boldsymbol{\eta}}), \quad \boldsymbol{w}\in\mathcal{W}.\end{array} \tag{29}$$

Recall that the set of admissible portfolios $\mathcal{W}$ precludes short positions in options, that is, $\boldsymbol{w}\in\mathcal{W}$ implies $\boldsymbol{w}^{\boldsymbol{\eta}}\geq\mathbf{0}$.

4.3 Robust Optimization Perspective on WCPVaR

In Section 2 we highlighted a known relationship between WCVaR optimization and robust optimization. Moreover, in Section 3 we argued that the ellipsoidal uncertainty set related to the WCVaR model is symmetric and as such fails to capture the asymmetric dependencies between options and their underlying assets. In the next theorem we establish that the WCPVaR minimization problem (29) can also be cast as a robust optimization problem of the type (10). However, the uncertainty set which generates WCPVaR is no longer symmetric.

Theorem 4.2 *The* WCPVaR *minimization problem* (29) *is equivalent to the robust optimization problem*

$$\begin{array}{lll}\underset{\boldsymbol{w}\in\mathbb{R}^m,\gamma\in\mathbb{R}}{\text{minimize}} & \gamma & \\ \text{subject to} & -\boldsymbol{w}^T\boldsymbol{r}\leq\gamma & \forall\boldsymbol{r}\in\mathcal{U}^p_\epsilon \\ & \boldsymbol{w}\in\mathcal{W}, & \end{array} \tag{30}$$

where the uncertainty set $\mathcal{U}^p_\epsilon\subseteq\mathbb{R}^m$ *is defined as*

$$\mathcal{U}^p_\epsilon=\left\{\boldsymbol{r}\in\mathbb{R}^m\ :\ \exists\boldsymbol{\xi}\in\mathbb{R}^n,\ (\boldsymbol{\xi}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{\xi}-\boldsymbol{\mu})\leq\kappa(\epsilon)^2,\ \boldsymbol{r}=f(\boldsymbol{\xi})\right\}. \tag{31}$$

Proof: The result is based on conic duality. We refer to [27, Theorem 3.1] for a complete exposition of the proof. ∎

Example 4.1 *Consider a Black-Scholes economy consisting of stocks A and B, a European call option on stock A, and a European put option on stock B. Furthermore, let* $\mathbf{w}$ *be an equally weighted portfolio of these* $m = 4$ *assets, that is, set* $w_i = 1/m$ *for* $i = 1, \dots, m$*.*

We assume that the prices of stocks A and B are governed by a bivariate geometric Brownian motion with drift coefficients of 12% and 8%, and volatilities of 30% and 20% per annum, respectively. The correlation between the instantaneous stock returns amounts to 20%. The initial prices of the stocks are $100. The options mature in 21 days and have strike prices of $100. We assume that the risk-free rate is 3% per annum and that there are 252 trading days per year. By using the Black-Scholes formulas [7], we obtain call and put option prices of $3.58 and $2.18, respectively.

We want to compute the VaR at confidence level ϵ *for portfolio* $\mathbf{w}$ *and a 21-day time horizon. To this end, we randomly generate L=5,000,000 end-of-period stock prices and corresponding option payoffs. These are used to obtain L asset and portfolio return samples. Figure 1 (left) displays the sampled portfolio loss distribution, which exhibits considerable skewness due to the options. The* Monte-Carlo VaR *is obtained by computing the* $(1 - \epsilon)$*-quantile of the sampled portfolio loss distribution. We also compute the sample means and sample covariance matrix of the asset returns, which are used for the calculation of* WCVaR (7) *and* WCPVaR (28)*.*

Figure 1 (right) displays the VaR estimates at different levels of $\epsilon \in [0.01, 0.2]$*. We observe that for all values of* ϵ*, the* WCVaR *and* WCPVaR *values exceed the* Monte-Carlo VaR *estimate. This is not surprising since these models are distributionally robust and as such provide a conservative estimate of VaR. Note that the* Monte-Carlo VaR *can only be calculated accurately if many return samples are available (e.g., if the return distribution is precisely known). However,* WCVaR *vastly overestimates* WCPVaR*. This effect is amplified for lower values of* ϵ*, where the accuracy of the VaR estimate matters most. Indeed, for* $\epsilon = 1\%$*, the* WCVaR *reports an unrealistically high value of* 497%*, which is 7 times larger than the corresponding* WCPVaR *value.*

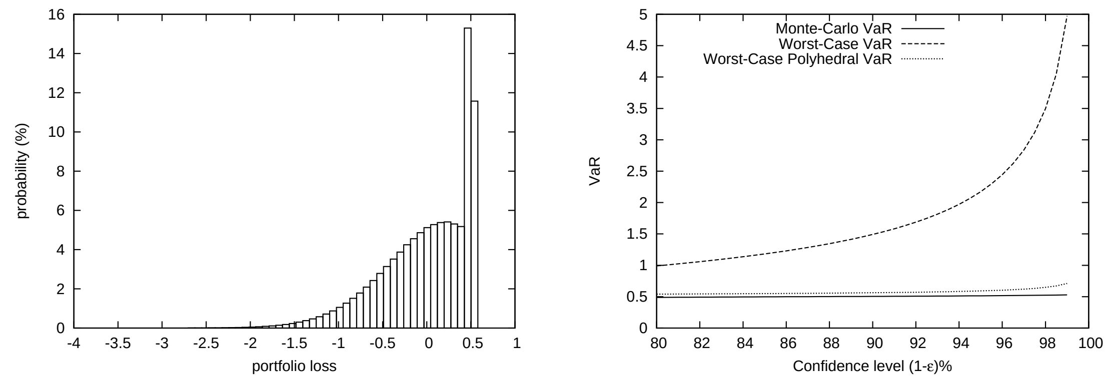

Figure 1: Left: The portfolio loss distribution obtained via Monte-Carlo simulation. Note that negative values represent gains. Right: The VaR estimates at different confidence levels obtained via Monte-Carlo sampling, WCVaR, and WCPVaR.

5 Worst-Case Quadratic VaR Optimization

The WCPVaR model suffers from a number of weaknesses which may make it unattractive for certain investors.

Firstly, in order to obtain a tractable problem reformulation we had to prohibit short-sales of options. Although this is not restrictive for investors who merely want to enrich their portfolios with options in order to obtain insurance benefits (see [27]), it severely constrains the complete set of option strategies that larger institutions might want to include in their portfolios.

Furthermore, we can only calculate and optimize the risk of portfolios comprising options that mature at the end of the investment horizon. As a result, investors cannot use the model, for example, to optimize portfolios including longer term options that mature far beyond the investment horizon.

Finally, the model is only suitable for portfolios containing plain vanilla European options and can not be used when exotic options are included in the portfolio.

In this section we propose an alternative Worst-Case VaR model which mitigates the weaknesses of the WCPVaR model. It is important to note that WCPVaR does not make any assumptions about the pricing model of the options. Only observable market prices and the known payoff functions of the options are used to calculate the option returns. In contrast, the new model proposed in this section requires the availability of a pricing model for the options. Moreover, it approximates the portfolio return using a second-order Taylor expansion which is only accurate for short investment horizons.

5.1 Delta-Gamma Portfolio Model

As in Section 4, we assume that there are $n \leq m$ *basic assets* and $m-n$ *derivatives* whose values are uniquely determined by the values of the basic assets. However, in contrast to Section 4, we do not only focus on European style options but also allow for exotic derivatives. Furthermore, we no longer require that the options mature at the end of the investment horizon.

We let $\tilde{\boldsymbol{s}}(t)$ denote the n-dimensional vector of basic asset prices at time $t \geq 0$ and assume that the prices at time $t = 0$ are known (i.e., deterministic). Moreover, we assume that the value of any (basic or non-basic) asset $i = 1, \ldots, m$ is representable as $v_i(\tilde{\boldsymbol{s}}(t), t)$, where $v_i : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ is twice continuously differentiable.

For a sufficiently short horizon time T, a second-order Taylor expansion accurately approximates the asset values at the end of the investment horizon. For $i = 1, \ldots, m$ we have

$$v_i(\tilde{\boldsymbol{s}}(T), T) - v_i(\boldsymbol{s}(0), 0) \;\approx\; \bar{\theta}_i T + \bar{\boldsymbol{\Delta}}_i^T (\tilde{\boldsymbol{s}}(T) - \boldsymbol{s}(0)) + \frac{1}{2}(\tilde{\boldsymbol{s}}(T) - \boldsymbol{s}(0))^T \bar{\boldsymbol{\Gamma}}_i (\tilde{\boldsymbol{s}}(T) - \boldsymbol{s}(0)),$$

where

$$\bar{\theta}_i = \partial_t v_i(\boldsymbol{s}(0), 0) \in \mathbb{R}, \quad \bar{\boldsymbol{\Delta}}_i = \nabla_{\boldsymbol{s}} v_i(\boldsymbol{s}(0), 0) \in \mathbb{R}^n, \quad \text{and} \quad \bar{\boldsymbol{\Gamma}}_i = \nabla^2_{\boldsymbol{s}} v_i(\boldsymbol{s}(0), 0) \in \mathbb{S}^n. \tag{32}$$

The values computed in (32) are referred to as the 'greeks' of the assets. We emphasize that the computation of the greeks relies on the availability of a pricing model, that is, the value functions v_i must be known. Note that the values of the functions v_i at $(\boldsymbol{s}(0), 0)$ can be observed in the market. However, the values of v_i in a neighborhood of $(\boldsymbol{s}(0), 0)$ are not observable. The proposed second-order Taylor approximation is very popular in finance and is often referred to as the *delta-gamma approximation*, see [13].

By using the *relative greeks*

$$\theta_i = \frac{T}{v_i(\boldsymbol{s}(0), 0)} \bar{\theta}_i, \quad \boldsymbol{\Delta}_i = \frac{1}{v_i(\boldsymbol{s}(0), 0)} \operatorname{diag}(\boldsymbol{s}(0)) \bar{\boldsymbol{\Delta}}_i, \quad \boldsymbol{\Gamma}_i = \frac{1}{v_i(\boldsymbol{s}(0), 0)} \operatorname{diag}(\boldsymbol{s}(0))^T \bar{\boldsymbol{\Gamma}}_i \operatorname{diag}(\boldsymbol{s}(0)),$$

the delta-gamma approximation can be reformulated in terms of relative returns

$$\tilde{r}_i \;\approx\; f_i(\tilde{\boldsymbol{\xi}}) = \theta_i + \boldsymbol{\Delta}_i^T \tilde{\boldsymbol{\xi}} + \frac{1}{2} \tilde{\boldsymbol{\xi}}^T \boldsymbol{\Gamma}_i \tilde{\boldsymbol{\xi}} \quad \forall i = 1, \ldots, m. \tag{33}$$

Here we use the (possibly non-convex) quadratic functions f_i to map the basic asset returns $\tilde{\boldsymbol{\xi}}$ to the asset returns $\tilde{\boldsymbol{r}}$.

The return of a portfolio $\boldsymbol{w} \in \mathcal{W}$ can therefore be approximated by

$$\boldsymbol{w}^T \tilde{\boldsymbol{r}} \approx \theta(\boldsymbol{w}) + \boldsymbol{\Delta}(\boldsymbol{w})^T \tilde{\boldsymbol{\xi}} + \frac{1}{2} \tilde{\boldsymbol{\xi}}^T \boldsymbol{\Gamma}(\boldsymbol{w}) \tilde{\boldsymbol{\xi}}, \tag{34}$$

where we use the auxiliary functions

$$\theta(\boldsymbol{w}) = \sum_{i=1}^{m} w_i \theta_i, \quad \boldsymbol{\Delta}(\boldsymbol{w}) = \sum_{i=1}^{m} w_i \boldsymbol{\Delta}_i, \quad \text{and} \quad \boldsymbol{\Gamma}(\boldsymbol{w}) = \sum_{i=1}^{m} w_i \boldsymbol{\Gamma}_i,$$

which are all linear in $\boldsymbol{w}$. We emphasize that, in contrast to Section 4, we now allow for short-sales of derivatives.

In the remainder of this section we derive a Worst-Case VaR optimization model based on the quadratic approximation (34).

5.2 Worst-Case Quadratic VaR Model

We define the *Worst-Case Quadratic VaR* (WCQVaR) of a fixed portfolio $\boldsymbol{w} \in \mathcal{W}$ in terms of the Taylor expansion (34).

$$\begin{aligned} \text{WCQVaR}_\epsilon(\boldsymbol{w}) &= \min \left\{ \gamma \ : \ \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{P} \left\{ \gamma \leq -\boldsymbol{w}^T f(\tilde{\boldsymbol{\xi}}) \right\} \leq \epsilon \right\} \\ &= \min \left\{ \gamma \ : \ \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{P} \left\{ \gamma \leq -\theta(\boldsymbol{w}) - \boldsymbol{\Delta}(\boldsymbol{w})^T \tilde{\boldsymbol{\xi}} - \frac{1}{2} \tilde{\boldsymbol{\xi}}^T \boldsymbol{\Gamma}(\boldsymbol{w}) \tilde{\boldsymbol{\xi}} \right\} \leq \epsilon \right\} \end{aligned} \tag{35}$$

Note that the WCQVaR approximates the portfolio return $\boldsymbol{w}^T \tilde{\boldsymbol{r}}$ by a (possibly non-convex) quadratic function of the basic asset returns $\tilde{\boldsymbol{\xi}}$.

Theorem 5.1 below shows how the WCQVaR of a portfolio $\boldsymbol{w}$ can be computed by solving a tractable SDP. We first recall the $\mathcal{S}$-lemma, which is a crucial ingredient for the proof of Theorem 5.1. We refer to Pólik and Terlaky [21] for a derivation and an in-depth survey of its manifold uses.

Lemma 5.1 ($\mathcal{S}$-lemma) *Let $f_i(\boldsymbol{\xi}) = \boldsymbol{\xi}^T \mathbf{A}_i \boldsymbol{\xi}$, $i = 0, \ldots, p$ be quadratic functions of $\boldsymbol{\xi} \in \mathbb{R}^n$.*

Then, $f_0(\boldsymbol{\xi}) \geq 0$ *for all* $\boldsymbol{\xi}$ *with* $f_i(\boldsymbol{\xi}) \leq 0$, $i = 1, \ldots, p$, *if there exist constants* $\tau_i \geq 0$ *such that*

$$\mathbf{A}_0 + \sum_{i=1}^{p} \tau_i \mathbf{A}_i \succcurlyeq \mathbf{0}.$$

For $p = 1$, *the converse implication holds if there exists a strictly feasible point* $\bar{\boldsymbol{\xi}}$ *with* $f_1(\bar{\boldsymbol{\xi}}) < 0$.

Theorem 5.1 *The* WCQVaR *of a fixed portfolio* $\boldsymbol{w} \in \mathcal{W}$ *can be computed by solving the following tractable SDP.*

$$\begin{aligned} \text{WCQVaR}_\epsilon(\boldsymbol{w}) = \inf \quad & \gamma \\ \text{s.t.} \quad & \mathbf{M} \in \mathbb{S}^{n+1}, \quad \tau \in \mathbb{R}, \quad \gamma \in \mathbb{R} \\ & \langle \boldsymbol{\Omega}, \mathbf{M} \rangle \leq \tau\epsilon, \quad \mathbf{M} \succcurlyeq \mathbf{0}, \quad \tau \geq 0, \\ & \mathbf{M} + \begin{bmatrix} \boldsymbol{\Gamma}(\boldsymbol{w}) & \boldsymbol{\Delta}(\boldsymbol{w}) \\ \boldsymbol{\Delta}(\boldsymbol{w})^T & -\tau + 2(\gamma + \theta(\boldsymbol{w})) \end{bmatrix} \succcurlyeq \mathbf{0} \end{aligned} \tag{36}$$

Proof: For the given portfolio $\boldsymbol{w} \in \mathcal{W}$ and for any fixed $\gamma \in \mathbb{R}$, we introduce the set $\mathcal{Q}_\gamma \subseteq \mathbb{R}^n$, defined through

$$\mathcal{Q}_\gamma = \left\{ \boldsymbol{\xi} \in \mathbb{R}^n \ : \ \gamma \leq -\theta(\boldsymbol{w}) - \boldsymbol{\Delta}(\boldsymbol{w})^T \boldsymbol{\xi} - \frac{1}{2} \boldsymbol{\xi}^T \boldsymbol{\Gamma}(\boldsymbol{w}) \boldsymbol{\xi} \right\}. \tag{37}$$

As in Section 4, the first step towards a tractable reformulation of WCQVaR is to solve the worst-case probability problem

$$\pi_{\text{wc}} = \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{P}\{\tilde{\boldsymbol{\xi}} \in \mathcal{Q}_\gamma\}, \tag{38}$$

which can be identified as the subordinate maximization problem in (35). Lemma 4.1 implies that (38) can equivalently be formulated as

$$\pi_{\text{wc}} = \inf_{\mathbf{M} \in \mathbb{S}^{n+1}} \left\{ \langle \boldsymbol{\Gamma}, \mathbf{M} \rangle \ : \ \mathbf{M} \succcurlyeq \mathbf{0}, \quad \begin{bmatrix} \boldsymbol{\xi}^T & 1 \end{bmatrix} \mathbf{M} \begin{bmatrix} \boldsymbol{\xi}^T & 1 \end{bmatrix}^T \geq 1 \quad \forall \boldsymbol{\xi} \in \mathcal{Q}_\gamma \right\}. \tag{39}$$

By the definition of $\mathcal{Q}$, the semi-infinite constraint in problem (39) is equivalent to

$$\begin{bmatrix} \boldsymbol{\xi}^T & 1 \end{bmatrix} (\mathbf{M} - \text{diag}(\mathbf{0}, 1)) \begin{bmatrix} \boldsymbol{\xi}^T & 1 \end{bmatrix}^T \geq 0 \quad \forall \boldsymbol{\xi} : \begin{bmatrix} \boldsymbol{\xi}^T & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2}\boldsymbol{\Gamma}(\boldsymbol{w}) & \frac{1}{2}\boldsymbol{\Delta}(\boldsymbol{w}) \\ \frac{1}{2}\boldsymbol{\Delta}(\boldsymbol{w})^T & \gamma + \theta(\boldsymbol{w}) \end{bmatrix} \begin{bmatrix} \boldsymbol{\xi}^T & 1 \end{bmatrix}^T \leq 0.$$

By using the $\mathcal{S}$-lemma and by analogous reasoning as in Section 4.2, we can replace the semi-infinite constraint in problem (39) by

$$\exists \tau \geq 0 \ : \ \mathbf{M} + \begin{bmatrix} \tau \mathbf{\Gamma}(\boldsymbol{w}) & \tau \mathbf{\Delta}(\boldsymbol{w}) \\ \tau \mathbf{\Delta}(\boldsymbol{w})^T & -1 + 2\tau(\gamma + \theta(\boldsymbol{w})) \end{bmatrix} \succcurlyeq \mathbf{0}$$

without changing the optimal value of the problem. Thus, the worst-case probability problem (38) can be rewritten as

$$\begin{aligned} \pi_{\text{wc}} = \inf \quad & \langle \mathbf{\Omega}, \mathbf{M} \rangle \\ \text{s.t.} \quad & \mathbf{M} \in \mathbb{S}^{n+1}, \quad \tau \in \mathbb{R}, \quad \mathbf{M} \succcurlyeq \mathbf{0}, \quad \tau \geq 0 \\ & \mathbf{M} + \begin{bmatrix} \tau \mathbf{\Gamma}(\boldsymbol{w}) & \tau \mathbf{\Delta}(\boldsymbol{w}) \\ \tau \mathbf{\Delta}(\boldsymbol{w})^T & -1 + 2\tau(\gamma + \theta(\boldsymbol{w})) \end{bmatrix} \succcurlyeq \mathbf{0}. \end{aligned} \tag{40}$$

The WCQVaR of the portfolio $\boldsymbol{w}$ can therefore be obtained by solving the following non-convex optimization problem.

$$\begin{aligned} \text{WCQVaR}_{\epsilon}(\boldsymbol{w}) = \inf \quad & \gamma \\ \text{s.t.} \quad & \mathbf{M} \in \mathbb{S}^{n+1}, \quad \tau \in \mathbb{R}, \quad \gamma \in \mathbb{R} \\ & \langle \mathbf{\Omega}, \mathbf{M} \rangle \leq \epsilon, \quad \mathbf{M} \succcurlyeq \mathbf{0}, \quad \tau \geq 0 \\ & \mathbf{M} + \begin{bmatrix} \tau \mathbf{\Gamma}(\boldsymbol{w}) & \tau \mathbf{\Delta}(\boldsymbol{w}) \\ \tau \mathbf{\Delta}(\boldsymbol{w})^T & -1 + 2\tau(\gamma + \theta(\boldsymbol{w})) \end{bmatrix} \succcurlyeq \mathbf{0} \end{aligned} \tag{41}$$

By analogous reasoning as in Section 4.2, it can be shown that any feasible solution of problem (41) has a strictly positive τ-component. Thus we may divide the matrix inequality in (41) by τ. After the variable transformation $\tau \rightarrow 1/\tau$ and $\mathbf{M} \rightarrow \mathbf{M}/\tau$, we obtain the postulated SDP (36). ∎

Remark 5.1 *In the absence of derivatives, that is, if the market only contains basic assets, then* $m = n$, *and the coefficient functions in the delta-gamma approximation* (34) *reduce to* $\theta(\boldsymbol{w}) = 0$, $\mathbf{\Delta}(\boldsymbol{w}) = \boldsymbol{w}$, *and* $\mathbf{\Gamma}(\boldsymbol{w}) = \mathbf{0}$. *In this special case, the* WCQVaR *is computed by solving the* *following SDP.*

$$\begin{aligned}
\mathrm{WCQVaR}_\epsilon(\boldsymbol{w}) = \inf \quad & \gamma \\
\text{s.t.} \quad & \mathbf{M} \in \mathbb{S}^{n+1}, \quad \tau \in \mathbb{R}, \quad \gamma \in \mathbb{R} \\
& \langle \boldsymbol{\Omega}, \mathbf{M} \rangle \leq \tau\epsilon, \quad \mathbf{M} \succcurlyeq \mathbf{0}, \quad \tau \geq 0 \\
& \mathbf{M} + \begin{bmatrix} \mathbf{0} & \boldsymbol{w} \\ \boldsymbol{w}^T & -\tau + 2\gamma \end{bmatrix} \succcurlyeq \mathbf{0}
\end{aligned}$$

El Ghaoui et al. *[11] have shown (using similar arguments as in Theorem 4.1) that this SDP has the closed form solution*

$$\mathrm{WCVaR}(\boldsymbol{w}) = -\boldsymbol{\mu}^T \boldsymbol{w} + \kappa(\epsilon)\sqrt{\boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w}}, \quad \textit{where} \quad \kappa(\epsilon) = \sqrt{\frac{1-\epsilon}{\epsilon}}.$$

Thus, the WCQVaR *model is a direct extension of the* WCVaR *model* (7)*.*

Problem (36) constitutes a convex SDP that facilitates the efficient computation of the WCQVaR for any fixed portfolio $\boldsymbol{w} \in \mathcal{W}$. Since the matrix inequality in (36) is linear in $(\mathbf{M}, \boldsymbol{\tau}, \gamma)$ and $\boldsymbol{w}$, one can reinterpret $\boldsymbol{w}$ as a decision variable without impairing the problem's convexity. This observation reveals that we can efficiently minimize the WCQVaR over all portfolios $\boldsymbol{w} \in \mathcal{W}$ by solving the following SDP.

$$\begin{aligned}
\inf \quad & \gamma \\
\text{s.t.} \quad & \mathbf{M} \in \mathbb{S}^{n+1}, \quad \tau \in \mathbb{R}, \quad \gamma \in \mathbb{R}, \quad \boldsymbol{w} \in \mathcal{W} \\
& \langle \boldsymbol{\Omega}, \mathbf{M} \rangle \leq \tau\epsilon, \quad \mathbf{M} \succcurlyeq \mathbf{0}, \quad \tau \geq 0 \\
& \mathbf{M} + \begin{bmatrix} \boldsymbol{\Gamma}(\boldsymbol{w}) & \boldsymbol{\Delta}(\boldsymbol{w}) \\ \boldsymbol{\Delta}(\boldsymbol{w})^T & -\tau + 2(\gamma + \theta(\boldsymbol{w})) \end{bmatrix} \succcurlyeq \mathbf{0}
\end{aligned} \tag{42}$$

Remark 5.2 *Unlike in Section 4, there seems to be no equivalent SOCP formulation for the SDP* (42)*. In particular, there is no simple way to adapt the arguments in the proof of Theorem 4.1 to the current setting. The reason for this is a fundamental difference between the corresponding SDP problems* (27) *and* (42)*. In fact, the top left principal submatrix in the last LMI constraint is independent of* $\boldsymbol{w}$ *in* (27) *but not in* (42)*.*

5.3 Robust Optimization Perspective on WCQVaR

We now highlight the close connection between robust optimization and WCQVaR minimization. In the next theorem we elaborate an equivalence between the WCQVaR minimization problem and a robust optimization problem whose uncertainty set is embedded into a space of positive semidefinite matrices.

Theorem 5.2 *The* WCQVaR *minimization problem* (42) *is equivalent to the robust optimization problem*

$$\begin{aligned} \underset{\boldsymbol{w}\in\mathbb{R}^m,\gamma\in\mathbb{R}}{\text{minimize}} \quad & \gamma \\ \text{subject to} \quad & -\langle \mathbf{Q}(\boldsymbol{w}), \mathbf{Z}\rangle \le \gamma \qquad \forall \mathbf{Z} \in \mathcal{U}_\epsilon^q \\ & \boldsymbol{w} \in \mathcal{W}, \end{aligned} \tag{43}$$

where

$$\mathbf{Q}(\boldsymbol{w}) = \begin{bmatrix} \frac{1}{2}\boldsymbol{\Gamma}(\boldsymbol{w}) & \frac{1}{2}\boldsymbol{\Delta}(\boldsymbol{w}) \\ \frac{1}{2}\boldsymbol{\Delta}(\boldsymbol{w})^T & \theta(\boldsymbol{w}) \end{bmatrix},$$

and the uncertainty set $\mathcal{U}_\epsilon^q \subseteq \mathbb{S}^{n+1}$ *is defined as*

$$\mathcal{U}_\epsilon^q = \left\{ \mathbf{Z} = \begin{bmatrix} \mathbf{X} & \boldsymbol{\xi} \\ \boldsymbol{\xi}^T & 1 \end{bmatrix} \in \mathbb{S}^{n+1} \ : \ \boldsymbol{\Omega} - \epsilon\mathbf{Z} \succcurlyeq \mathbf{0}, \ \mathbf{Z} \succcurlyeq \mathbf{0} \right\}. \tag{44}$$

Proof: For some fixed portfolio $\boldsymbol{w} \in \mathcal{W}$, the WCQVaR can be computed by solving problem (36), which involves the LMI constraint

$$\mathbf{M} + \begin{bmatrix} \boldsymbol{\Gamma}(\boldsymbol{w}) & \boldsymbol{\Delta}(\boldsymbol{w}) \\ \boldsymbol{\Delta}(\boldsymbol{w})^T & -\tau + 2(\gamma + \theta(\boldsymbol{w})) \end{bmatrix} \succcurlyeq \mathbf{0}. \tag{45}$$

Without loss of generality, we can rewrite the matrix $\mathbf{M}$ as

$$\mathbf{M} = \begin{bmatrix} \mathbf{V} & \boldsymbol{v} \\ \boldsymbol{v}^T & u \end{bmatrix}.$$

With this new notation, the LMI constraint (45) is representable as

$$
\begin{aligned}
& [\boldsymbol{\xi}^T\ 1] \begin{bmatrix} \mathbf{V} + \boldsymbol{\Gamma}(\boldsymbol{w}) & \boldsymbol{v} + \boldsymbol{\Delta}(\boldsymbol{w}) \\ (\boldsymbol{v} + \boldsymbol{\Delta}(\boldsymbol{w}))^T & u - \tau + 2(\gamma + \theta(\boldsymbol{w})) \end{bmatrix} [\boldsymbol{\xi}^T\ 1]^T \geq 0 && \forall \boldsymbol{\xi} \in \mathbb{R}^n \\
\Longleftrightarrow \quad & \boldsymbol{\xi}^T(\mathbf{V} + \boldsymbol{\Gamma}(\boldsymbol{w}))\boldsymbol{\xi} + 2\boldsymbol{\xi}^T(\boldsymbol{v} + \boldsymbol{\Delta}(\boldsymbol{w})) + u - \tau + 2(\gamma + \theta(\boldsymbol{w})) \geq 0 && \forall \boldsymbol{\xi} \in \mathbb{R}^n \\
\Longleftrightarrow \quad & \gamma \geq -\frac{1}{2}\boldsymbol{\xi}^T(\mathbf{V} + \boldsymbol{\Gamma}(\boldsymbol{w}))\boldsymbol{\xi} - \boldsymbol{\xi}^T(\boldsymbol{v} + \boldsymbol{\Delta}(\boldsymbol{w})) - \theta(\boldsymbol{w}) - \frac{1}{2}(u - \tau) && \forall \boldsymbol{\xi} \in \mathbb{R}^n \\
\Longleftrightarrow \quad & \gamma \geq \sup_{\boldsymbol{\xi} \in \mathbb{R}^n} \left\{ -\frac{1}{2}\boldsymbol{\xi}^T(\mathbf{V} + \boldsymbol{\Gamma}(\boldsymbol{w}))\boldsymbol{\xi} - \boldsymbol{\xi}^T(\boldsymbol{v} + \boldsymbol{\Delta}(\boldsymbol{w})) - \theta(\boldsymbol{w}) - \frac{1}{2}(u - \tau) \right\}.
\end{aligned}
$$

Thus, the WCQVaR problem (36) can be rewritten as

$$
\begin{aligned}
\inf \quad & \sup_{\boldsymbol{\xi} \in \mathbb{R}^n} \ -\frac{1}{2}\boldsymbol{\xi}^T(\mathbf{V} + \boldsymbol{\Gamma}(\boldsymbol{w}))\boldsymbol{\xi} - \boldsymbol{\xi}^T(\boldsymbol{v} + \boldsymbol{\Delta}(\boldsymbol{w})) - \theta(\boldsymbol{w}) - \frac{1}{2}(u - \tau) \\
\text{s.t.} \quad & \mathbf{V} \in \mathbb{S}^n, \quad \boldsymbol{v} \in \mathbb{R}^n, \quad \tau \in \mathbb{R}, \quad u \in \mathbb{R} \\
& \begin{bmatrix} \mathbf{V} & \boldsymbol{v} \\ \boldsymbol{v}^T & u \end{bmatrix} \succcurlyeq \mathbf{0}, \quad \tau \geq 0, \quad \langle \mathbf{V}, \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T \rangle + 2\boldsymbol{v}^T\boldsymbol{\mu} + u \leq \tau\epsilon.
\end{aligned} \tag{46}
$$

Note that if $\mathbf{V} + \boldsymbol{\Gamma}(\boldsymbol{w})$ is not positive semidefinite, the inner maximization problem in (46) is unbounded. However, this implies that any $\mathbf{V} \in \mathbb{S}^n$ is infeasible in the outer minimization problem unless $\mathbf{V} + \boldsymbol{\Gamma}(\boldsymbol{w}) \succcurlyeq \mathbf{0}$. Therefore, we can add the constraint $\mathbf{V} + \boldsymbol{\Gamma}(\boldsymbol{w}) \succcurlyeq \mathbf{0}$ to the minimization problem in (46) without changing its feasible region. With this constraint appended, the min-max problem (46) becomes a saddlepoint problem because its objective is concave in $\boldsymbol{\xi}$ for any fixed $(\mathbf{V}, \boldsymbol{v}, u, \tau)$ and convex in $(\mathbf{V}, \boldsymbol{v}, u, \tau)$ for any fixed $\boldsymbol{\xi}$. Moreover, the feasible sets of the outer and inner problems are convex and independent of each other. Thus, we may interchange the 'inf' and 'sup' operators to obtain the following equivalent problem, see, e.g., [10, Theorem 5.1].

$$
\begin{aligned}
\max_{\boldsymbol{\xi} \in \mathbb{R}^n} \ \min \quad & -\frac{1}{2}\boldsymbol{\xi}^T(\mathbf{V} + \boldsymbol{\Gamma}(\boldsymbol{w}))\boldsymbol{\xi} - \boldsymbol{\xi}^T(\boldsymbol{v} + \boldsymbol{\Delta}(\boldsymbol{w})) - \theta(\boldsymbol{w}) - \frac{1}{2}(u - \tau) \\
\text{s.t.} \quad & \mathbf{V} \in \mathbb{S}^n, \quad \boldsymbol{v} \in \mathbb{R}^n, \quad \tau \in \mathbb{R}, \quad u \in \mathbb{R} \\
& \begin{bmatrix} \mathbf{V} & \boldsymbol{v} \\ \boldsymbol{v}^T & u \end{bmatrix} \succcurlyeq \mathbf{0}, \quad \tau \geq 0, \quad \langle \mathbf{V}, \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T \rangle + 2\boldsymbol{v}^T\boldsymbol{\mu} + u \leq \tau\epsilon.
\end{aligned} \tag{47}
$$

We proceed by dualizing the inner minimization problem in (47). After a few elementary simplification steps, this dual problem reduces to

$$\begin{aligned}
\max \quad & -\frac{1}{2}\langle \mathbf{\Gamma}(\boldsymbol{w}), \boldsymbol{\xi}\boldsymbol{\xi}^T + \mathbf{Y}\rangle - \boldsymbol{\xi}^T \mathbf{\Delta}(\boldsymbol{w}) - \theta(\boldsymbol{w}) \\
\text{s.t.} \quad & \mathbf{Y} \in \mathbb{S}^n, \quad \alpha \in \mathbb{R}, \quad \mathbf{Y} \succcurlyeq \mathbf{0}, \quad 1 \le \alpha \le \frac{1}{\epsilon} \\
& \begin{bmatrix} \alpha(\mathbf{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T) - (\boldsymbol{\xi}\boldsymbol{\xi}^T + \mathbf{Y}) & \alpha\boldsymbol{\mu} - \boldsymbol{\xi} \\ (\alpha\boldsymbol{\mu} - \boldsymbol{\xi})^T & \alpha - 1 \end{bmatrix} \succcurlyeq \mathbf{0}.
\end{aligned} \tag{48}$$

Note that strong duality holds because the inner problem in (47) is strictly feasible for any $\epsilon > 0$, see [26]. This allows us to replace the inner minimization problem in (47) by the maximization problem (48), which yields the following equivalent formulation for the WCQVaR problem (36).

$$\begin{aligned}
\max \quad & -\frac{1}{2}\langle \mathbf{\Gamma}(\boldsymbol{w}), \boldsymbol{\xi}\boldsymbol{\xi}^T + \mathbf{Y}\rangle - \boldsymbol{\xi}^T \mathbf{\Delta}(\boldsymbol{w}) - \theta(\boldsymbol{w}) \\
\text{s.t.} \quad & \mathbf{Y} \in \mathbb{S}^n, \quad \boldsymbol{\xi} \in \mathbb{R}^n, \quad \alpha \in \mathbb{R}, \quad \mathbf{Y} \succcurlyeq \mathbf{0}, \quad 1 \le \alpha \le \frac{1}{\epsilon} \\
& \begin{bmatrix} \alpha(\mathbf{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T) - (\boldsymbol{\xi}\boldsymbol{\xi}^T + \mathbf{Y}) & \alpha\boldsymbol{\mu} - \boldsymbol{\xi} \\ (\alpha\boldsymbol{\mu} - \boldsymbol{\xi})^T & \alpha - 1 \end{bmatrix} \succcurlyeq \mathbf{0}
\end{aligned}$$

We now introduce a new decision variable $\mathbf{X} = \boldsymbol{\xi}\boldsymbol{\xi}^T + \mathbf{Y}$, which allows us to reformulate the above problem as

$$\begin{aligned}
\max \quad & -\frac{1}{2}\langle \mathbf{\Gamma}(\boldsymbol{w}), \mathbf{X}\rangle - \boldsymbol{\xi}^T \mathbf{\Delta}(\boldsymbol{w}) - \theta(\boldsymbol{w}) \\
\text{s.t.} \quad & \mathbf{X} \in \mathbb{S}^n, \quad \boldsymbol{\xi} \in \mathbb{R}^n, \quad \alpha \in \mathbb{R}, \quad 1 \le \alpha \le \frac{1}{\epsilon} \\
& \begin{bmatrix} \alpha(\mathbf{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T) - \mathbf{X} & \alpha\boldsymbol{\mu} - \boldsymbol{\xi} \\ (\alpha\boldsymbol{\mu} - \boldsymbol{\xi})^T & \alpha - 1 \end{bmatrix} \succcurlyeq \mathbf{0}, \quad \mathbf{X} - \boldsymbol{\xi}\boldsymbol{\xi}^T \succcurlyeq \mathbf{0}.
\end{aligned}$$

By definition of $\mathbf{\Omega}$ as the second-order moment matrix of the basic asset returns, see (22), the first LMI constraint in the above problem can be rewritten as

$$\alpha\mathbf{\Omega} - \begin{bmatrix} \mathbf{X} & \boldsymbol{\xi} \\ \boldsymbol{\xi}^T & 1 \end{bmatrix} \succcurlyeq \mathbf{0}.$$

Furthermore, by using Schur complements, the following equivalence holds.

$$\mathbf{X} - \boldsymbol{\xi}\boldsymbol{\xi}^T \succcurlyeq \mathbf{0} \iff \begin{bmatrix} \mathbf{X} & \boldsymbol{\xi} \\ \boldsymbol{\xi}^T & 1 \end{bmatrix} \succcurlyeq \mathbf{0}$$

Therefore, problem (48) can be reformulated as

$$\begin{aligned} \max \quad & -\left\langle \begin{bmatrix} \frac{1}{2}\boldsymbol{\Gamma}(\boldsymbol{w}) & \frac{1}{2}\boldsymbol{\Delta}(\boldsymbol{w}) \\ \frac{1}{2}\boldsymbol{\Delta}(\boldsymbol{w})^T & \theta(\boldsymbol{w}) \end{bmatrix}, \begin{bmatrix} \mathbf{X} & \boldsymbol{\xi} \\ \boldsymbol{\xi}^T & 1 \end{bmatrix} \right\rangle \\ \text{s.t.} \quad & \mathbf{X} \in \mathbb{S}^n, \quad \boldsymbol{\xi} \in \mathbb{R}^n, \quad \alpha \in \mathbb{R}, \quad 1 \le \alpha \le \frac{1}{\epsilon} \\ & \alpha\boldsymbol{\Omega} - \begin{bmatrix} \mathbf{X} & \boldsymbol{\xi} \\ \boldsymbol{\xi}^T & 1 \end{bmatrix} \succcurlyeq \mathbf{0}, \quad \begin{bmatrix} \mathbf{X} & \boldsymbol{\xi} \\ \boldsymbol{\xi}^T & 1 \end{bmatrix} \succcurlyeq \mathbf{0}. \end{aligned}$$

Since the objective function is independent of α and $\boldsymbol{\Omega} \succ \mathbf{0}$, the optimal choice for α is $1/\epsilon$; in fact, this choice of α generates the largest feasible set. We conclude that the WCQVaR for a fixed portfolio $\boldsymbol{w}$ can be computed by solving the following problem.

$$\begin{aligned} \max \quad & -\left\langle \begin{bmatrix} \frac{1}{2}\boldsymbol{\Gamma}(\boldsymbol{w}) & \frac{1}{2}\boldsymbol{\Delta}(\boldsymbol{w}) \\ \frac{1}{2}\boldsymbol{\Delta}(\boldsymbol{w})^T & \theta(\boldsymbol{w}) \end{bmatrix}, \begin{bmatrix} \mathbf{X} & \boldsymbol{\xi} \\ \boldsymbol{\xi}^T & 1 \end{bmatrix} \right\rangle \\ \text{s.t.} \quad & \mathbf{X} \in \mathbb{S}^n, \quad \boldsymbol{\xi} \in \mathbb{R}^n, \quad \boldsymbol{\Omega} - \epsilon \begin{bmatrix} \mathbf{X} & \boldsymbol{\xi} \\ \boldsymbol{\xi}^T & 1 \end{bmatrix} \succcurlyeq \mathbf{0}, \quad \begin{bmatrix} \mathbf{X} & \boldsymbol{\xi} \\ \boldsymbol{\xi}^T & 1 \end{bmatrix} \succcurlyeq \mathbf{0} \end{aligned}$$

The WCQVaR minimization problem (42) can therefore be expressed as the min-max problem

$$\min_{\boldsymbol{w} \in \mathcal{W}} \max_{\mathbf{Z} \in \mathcal{U}_\epsilon^q} -\langle \mathbf{Q}(\boldsymbol{w}), \mathbf{Z} \rangle, \tag{49}$$

which is manifestly equivalent to the postulated semi-infinite program (43). ∎

It may not be evident how the uncertainty set $\mathcal{U}_\epsilon^q$ (defined in (44)) associated with the WCQVaR formulation is related to the ellipsoidal uncertainty set $\mathcal{U}_\epsilon$ defined in Section 2.2. We now demonstrate that there exists a strong connection between these two uncertainty sets, even though they are embedded in spaces of different dimensions.

Corollary 5.1 *If the constraint $\boldsymbol{\Gamma}(\boldsymbol{w}) \succcurlyeq \mathbf{0}$ is appended to the definition of the set $\mathcal{W}$ of admis-*

sible portfolios, then the robust optimization problem (43) *reduces to*

$$\begin{aligned} \underset{\boldsymbol{w}\in\mathbb{R}^m,\gamma\in\mathbb{R}}{\text{minimize}} \quad & \gamma \\ \text{subject to} \quad & -\theta(\boldsymbol{w}) - \boldsymbol{\Delta}(\boldsymbol{w})^T\boldsymbol{\xi} - \frac{1}{2}\boldsymbol{\xi}^T\boldsymbol{\Gamma}(\boldsymbol{w})\boldsymbol{\xi} \leq \gamma \qquad \forall \boldsymbol{\xi} \in \mathcal{U}_\epsilon \\ & \boldsymbol{w} \in \mathcal{W}, \end{aligned} \tag{50}$$

where $\mathcal{U}_\epsilon$ *is the ellipsoidal uncertainty set defined in Section 2.2.*

Proof: The inner maximization problem in (49) can be written as

$$\begin{aligned} \max \quad & -\theta(\boldsymbol{w}) - \boldsymbol{\Delta}(\boldsymbol{w})^T\boldsymbol{\xi} - \frac{1}{2}\langle \boldsymbol{\Gamma}(\boldsymbol{w}), \mathbf{X}\rangle \\ \text{s.t.} \quad & \mathbf{X} \in \mathbb{S}^n, \quad \boldsymbol{\xi} \in \mathbb{R}^n, \quad \mathbf{X} - \boldsymbol{\xi}\boldsymbol{\xi}^T \succcurlyeq \mathbf{0} \\ & \begin{bmatrix} (\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T) - \epsilon\mathbf{X} & \boldsymbol{\mu} - \epsilon\boldsymbol{\xi} \\ (\boldsymbol{\mu} - \epsilon\boldsymbol{\xi})^T & 1 - \epsilon \end{bmatrix} \succcurlyeq \mathbf{0}. \end{aligned}$$

By introducing the decision variable $\mathbf{Y} = \mathbf{X} - \boldsymbol{\xi}\boldsymbol{\xi}^T$ as in the proof of Theorem 5.2, the above problem can be reformulated as

$$\begin{aligned} \max \quad & -\theta(\boldsymbol{w}) - \boldsymbol{\Delta}(\boldsymbol{w})^T\boldsymbol{\xi} - \frac{1}{2}\boldsymbol{\xi}^T\boldsymbol{\Gamma}(\boldsymbol{w})\boldsymbol{\xi} - \frac{1}{2}\langle \boldsymbol{\Gamma}(\boldsymbol{w}), \mathbf{Y}\rangle \\ \text{s.t.} \quad & \mathbf{Y} \in \mathbb{S}^n, \quad \boldsymbol{\xi} \in \mathbb{R}^n, \quad \mathbf{Y} \succcurlyeq \mathbf{0} \\ & \begin{bmatrix} (\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T) - \epsilon(\mathbf{Y} + \boldsymbol{\xi}\boldsymbol{\xi}^T) & \boldsymbol{\mu} - \epsilon\boldsymbol{\xi} \\ (\boldsymbol{\mu} - \epsilon\boldsymbol{\xi})^T & 1 - \epsilon \end{bmatrix} \succcurlyeq \mathbf{0}. \end{aligned} \tag{51}$$

We will now argue that $\mathbf{Y} = \mathbf{0}$ at optimality. This holds due to the following two facts: (i) for $\mathbf{Y} = \mathbf{0}$ we obtain the largest feasible set, and (ii) we have $\langle \boldsymbol{\Gamma}(\boldsymbol{w}), \mathbf{Y}\rangle \geq 0$ for all $\mathbf{Y} \succcurlyeq \mathbf{0}$ because $\boldsymbol{\Gamma}(\boldsymbol{w}) \succcurlyeq \mathbf{0}$ by assumption. Thus problem (51) reduces to

$$\begin{aligned} \max_{\boldsymbol{\xi}\in\mathbb{R}^n} \quad & -\theta(\boldsymbol{w}) - \boldsymbol{\Delta}(\boldsymbol{w})^T\boldsymbol{\xi} - \frac{1}{2}\boldsymbol{\xi}^T\boldsymbol{\Gamma}(\boldsymbol{w})\boldsymbol{\xi} \\ \text{s.t.} \quad & \begin{bmatrix} (\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T) - \epsilon\boldsymbol{\xi}\boldsymbol{\xi}^T & \boldsymbol{\mu} - \epsilon\boldsymbol{\xi} \\ (\boldsymbol{\mu} - \epsilon\boldsymbol{\xi})^T & 1 - \epsilon \end{bmatrix} \succcurlyeq \mathbf{0}. \end{aligned}$$

Using similar arguments as in Theorem 4.1 (in particular, see (A.2)), we can show that the semidefinite constraint in the above problem is equivalent to

$$\begin{bmatrix} \boldsymbol{\Sigma} & \boldsymbol{\xi} - \boldsymbol{\mu} \\ (\boldsymbol{\xi} - \boldsymbol{\mu})^T & \kappa(\epsilon)^2 \end{bmatrix} \succcurlyeq \mathbf{0} \iff (\boldsymbol{\xi} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\xi} - \boldsymbol{\mu}) \leq \kappa(\epsilon)^2.$$

Thus the original min-max formulation (49) can be reexpressed as

$$\min_{\boldsymbol{w} \in \mathcal{W}} \max_{\boldsymbol{\xi} \in \mathcal{U}_\epsilon} -\theta(\boldsymbol{w}) - \boldsymbol{\Delta}(\boldsymbol{w})^T \boldsymbol{\xi} - \frac{1}{2} \boldsymbol{\xi}^T \boldsymbol{\Gamma}(\boldsymbol{w}) \boldsymbol{\xi},$$

which is equivalent to the postulated robust optimization problem. ∎

Remark 5.3 *Note that the robust optimization problem* (50) *can be reformulated as*

$$\begin{aligned} \underset{\boldsymbol{w} \in \mathbb{R}^m, \gamma \in \mathbb{R}}{\text{minimize}} \quad & \gamma \\ \text{subject to} \quad & -\boldsymbol{w}^T \boldsymbol{r} \leq \gamma \qquad \forall \boldsymbol{r} \in \mathcal{U}_\epsilon^{q2} \\ & \boldsymbol{w} \in \mathcal{W}, \end{aligned} \tag{52}$$

where the uncertainty set $\mathcal{U}_\epsilon^{q2}$ is defined as

$$\mathcal{U}_\epsilon^{q2} = \left\{ \boldsymbol{r} \in \mathbb{R}^m : \begin{array}{c} \exists \boldsymbol{\xi} \in \mathbb{R}^n \quad \textit{such that} \\ (\boldsymbol{\xi} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\xi} - \boldsymbol{\mu}) \leq \kappa(\epsilon)^2 \quad \textit{and} \\ r_i = \theta_i + \boldsymbol{\xi}^T \boldsymbol{\Delta}_i + \frac{1}{2} \boldsymbol{\xi}^T \boldsymbol{\Gamma}_i \boldsymbol{\xi} \quad \forall i = 1, \ldots, m \end{array} \right\}$$

In contrast to the simple ellipsoidal set $\mathcal{U}_\epsilon$, the set $\mathcal{U}_\epsilon^{q2}$ is asymmetrically oriented around $\boldsymbol{\mu}$. This asymmetry is caused by the quadratic functions that map the basic asset returns $\boldsymbol{\xi}$ to the asset returns $\boldsymbol{r}$. As a result, the WCQVaR *model may provide a tighter approximation of the actual VaR of a portfolio containing derivatives than the* WCVaR *model.*

It seems that a min-max formulation (52) with an uncertainty set embedded into $\mathbb{R}^m$ is only available if $\boldsymbol{\Gamma}(\boldsymbol{w}) \succcurlyeq \mathbf{0}$, that is, if the portfolio return is a convex quadratic function of the basic assets returns. In general, however, one needs to resort to the more general formulation (43), in which the uncertainty set is embedded into $\mathbb{S}^{n+1}$; the dimension increase can compensate for the non-convexity of the portfolio return function.

Example 5.1 *We repeat the same experiment as in Example 4.1 but estimate the portfolio VaR*

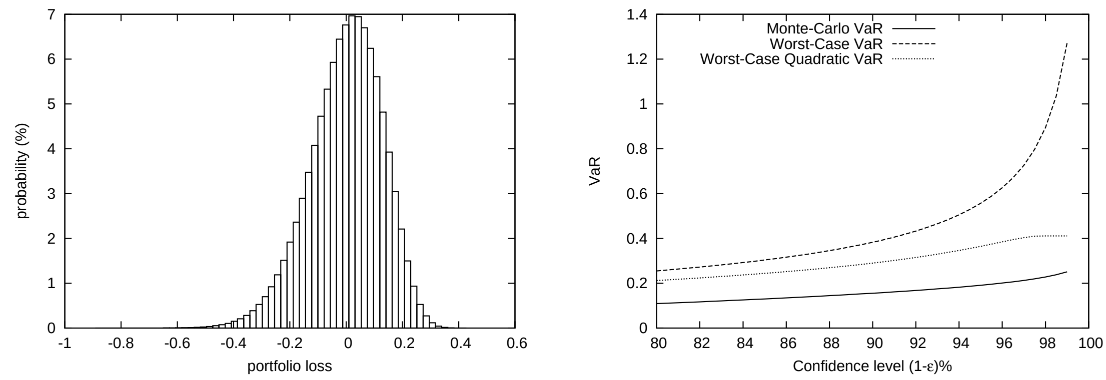

Figure 2: Left: The portfolio loss distribution obtained via Monte-Carlo simulation. Note that negative values represent gains. Right: The VaR estimates at different confidence levels obtained via Monte-Carlo sampling, WCVaR, and WCQVaR.

after 2 days instead of 21 days. Since the VaR is no longer evaluated at the maturity time of the options, we use the WCQVaR *model instead of the* WCPVaR *model. The coefficients of the quadratic approximation function* (34) *are calculated using the standard Black-Scholes greek formulas (see, e.g., [16]). We use an analogous Monte-Carlo simulation as in Example 4.1 to generate the stock and option returns over a 2-day investment period as well as the corresponding sample means and covariances. Figure 2 (left) displays the sampled portfolio loss distribution, which is still skewed, although considerably less than in Example 4.1. In Figure 2 (right) we compare* Monte-Carlo VaR*,* WCVaR*, and* WCQVaR *for different confidence levels. Even for the short horizon time under consideration, the* WCVaR *model still fails to give a realistic VaR estimate. At* $\epsilon = 1\%$*,* WCVaR *is more than 3 times as large as the corresponding* WCQVaR *value. This example demonstrates that the* WCQVaR *can offer significantly better VaR estimates than* WCVaR *when the portfolio contains options.*

6 Computational Results

In Section 6.1 we compare the out-of-sample performance of the WCQVaR in the context of an index tracking application and analyze the benefits of including options in the investment strategy. We refer to [27] for an in-depth analysis of the in- and out-of-sample performance of the robust optimization problem (30), whose equivalence to our novel WCPVaR model was established in Theorem 4.2. All computations are performed within Matlab 2008b and by using the YALMIP interface [15] of the SDPT3 optimization toolkit [25]. We use a 2.0 GHz Core 2 Duo machine running Linux Ubuntu 9.04.

6.1 Index Tracking using Worst-Case VaR

Index tracking is a common and important problem in portfolio management. The aim is to *replicate* the behavior of a given stock market index, sometimes referred to as the *benchmark*, with a given set of other assets not containing the index itself.

We let $\tilde{r}_1$ denote the random return of the benchmark over the investment interval $[0, T]$. In order to replicate this benchmark, we are given $m-1$ assets, whose vector of returns is denoted by $\tilde{\boldsymbol{r}}_{-1}$. This set of assets includes $n-1$ basic assets as well as $m-n$ options derived from the basic assets. We denote by $\boldsymbol{w}_{-1} \in \mathbb{R}^{m-1}$ the asset weights in the replicating portfolio.

Typically, the level of discrepancy between the benchmark and the portfolio is quantified by the *tracking-error* $\mathbb{E}(|\boldsymbol{w}_{-1}^T \tilde{\boldsymbol{r}}_{-1} - \tilde{r}_1|)$. Note that minimizing the tracking-error penalizes both under- and over-performance of the portfolio relative to the benchmark.

In this paper we adopt a slightly different approach. Instead of minimizing the tracking-error, we are only concerned about the portfolio falling short of the benchmark. The *excess-return* of a portfolio $\boldsymbol{w}_{-1}$ relative to the benchmark is computed as $\boldsymbol{w}^T \tilde{\boldsymbol{r}}$ where $\boldsymbol{w} = [-1 \quad \boldsymbol{w}_{-1}^T]^T$ and $\tilde{\boldsymbol{r}} = [\tilde{r}_1 \quad \tilde{\boldsymbol{r}}_{-1}^T]^T$. In order to measure the risk of the replicating portfolio falling below the benchmark, we can use the VaR at confidence $\epsilon = 5\%$.[2] The optimal replicating portfolio is found by minimizing $\text{VaR}_\epsilon(\boldsymbol{w})$ over all admissible portfolios $\boldsymbol{w} \in \mathcal{W}$ with

$$\mathcal{W} = \left\{ \boldsymbol{w} \in \mathbb{R}^m \ : \ \boldsymbol{w}^+ - \boldsymbol{w}^- = \boldsymbol{w}, \ \ \boldsymbol{e}^T \boldsymbol{w}^- \leq \alpha + 1, \ \ \boldsymbol{w}^+ \geq \boldsymbol{0}, \ \ \boldsymbol{w}^- \geq \boldsymbol{0}, \ \ \boldsymbol{e}^T \boldsymbol{w} = 0 \right\}. \tag{53}$$

The inclusion $\boldsymbol{w} \in \mathcal{W}$ implies that the portfolio weights $\boldsymbol{w}_{-1}$ sum up to 1 and that the total amount of shortsales in the replicating portfolio is limited to $\alpha = 4\%$.

Since we include options in the replicating portfolio, we use $\text{WCQVaR}_\epsilon(\boldsymbol{w})$ to approximate the VaR objective. The optimal portfolios are found by solving problem (42).

We now compare the out-of-sample performance of the optimal portfolios containing options with those where investment in options is prohibited. Recall that in the absence of options WCQVaR reduces to WCVaR, see Remark 5.1.

We assess the out-of-sample behavior of the WCQVaR model using a rolling-horizon backtest procedure. The aim is to minimize the under-performance of the replicating portfolio relative

[2] We ran the backtests in this section with different values of ϵ. Although we only report results for $\epsilon = 5\%$, the general conclusions are independent of the choice of ϵ.

to the S&P 500 index, which is often taken as a proxy for the market portfolio. The replicating portfolio is based on the 30 stock constituents of the Dow Jones Industrial Average, as well as some options written on these. We only include options that expire between 30 and 60 days after the investment dates. This ensures that the option payoffs are differentiable and accurately representable by the delta-gamma approximation. Moreover, longer term options tend to be more illiquid and are therefore not included.

Daily stock and option data are obtained from the Optionmetrics IvyDB database, which is one of the most complete sources of historical option data available. We consider a historical data range from January 2nd, 2004 to October 10th, 2008, containing a total of 1181 trading days. We use the following rolling-horizon backtest procedure. At every investment date we estimate the mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ of the stock returns using the daily returns of the previous 600 trading days. Thus, our backtest starts on the 600th trading day in the historical data set. We compute the out-of-sample returns of the optimal replicating portfolios using the stock and option prices on the next available trading day. This process is repeated for all but the first 600 trading days in our data set.

For simplicity, we use the mid-prices of the assets to calculate the returns. Furthermore, the WCQVaR model requires information about the options' delta and gamma sensitivities. These are obtained from the implied volatilities reported in the Optionmetrics database and are calculated using the Black-Scholes formula.[3] We disregard transaction costs and income taxes on option returns, which are beyond the scope of this paper.

The same rolling-horizon procedure is used to obtain the out-of-sample returns of the optimal replicating portfolios with and without options. On average the optimal stock-only portfolios are found in 2.1 seconds, whereas the portfolios with options are found in 7.4 seconds. In total we obtain two sequences of $L = 581$ out-of-sample portfolio returns, corresponding to the strategies with and without options, which are denote by $\{r_l^o\}_{l=1}^L$ and $\{r_l^s\}_{l=1}^L$, respectively. The returns of the benchmark are denoted by $\{r_{1,l}\}_{l=1}^L$.

Since the portfolios minimize the under-performance with respect to the benchmark, it is of interest to analyze how much wealth the robust strategies generate relative to the benchmark. By assuming an initial capital of 1 dollar, we calculate the relative wealth ω_l^k at the end of

[3] In order to avoid the use of erroneous option data, we only selected those options for which the implied volatility was supplied in the database and which had a bid and ask price greater than 0. We found that this procedure allowed us to filter out incorrect entries.

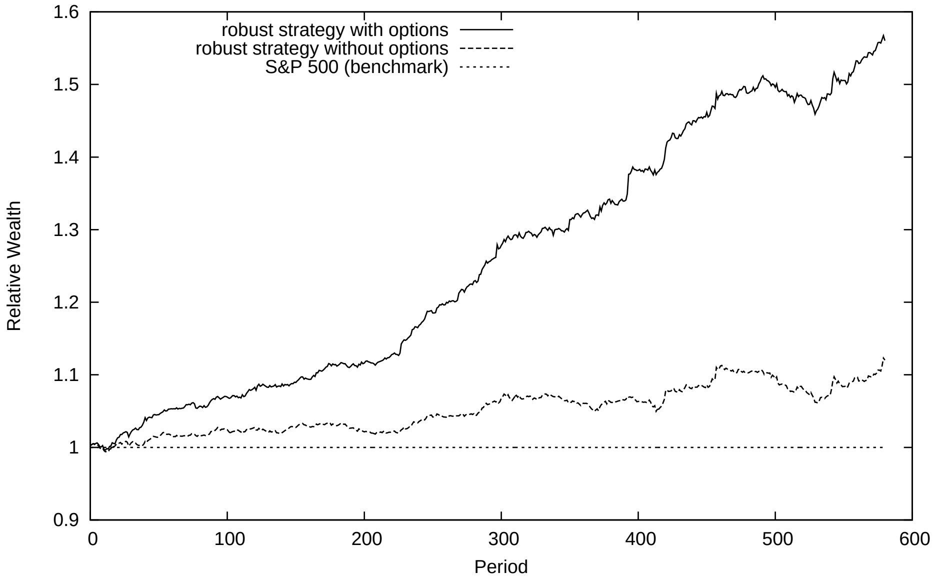

Figure 3: Cumulative relative wealth over time of the robust strategies using daily rebalancing between 22/05/2006 and 10/10/2008.

period l for portfolio strategy $k = o, s$ as

$$\omega_l^k = \frac{\prod_{m=1}^{l}(1 + r_m^k)}{\prod_{m=1}^{l}(1 + r_{1,m})}.$$

Figure 3 displays the relative wealth generated over time by the robust strategies. Both strategies outperform the benchmark over the entire test period. However, the inclusion of options improves the performance considerably. Over the test period, the strategy with options outperforms the benchmark by 56%, whereas the stock-only strategy only outperforms the benchmark by 12%. The annualized average excess-return of the stock-only strategy is 4.9% and that of the option strategy amounts to 19%.

The Sharpe ratio [24] is frequently used to assess the performance of an investment strategy. It is calculated as $(\hat{\mu} - r_f)/\hat{\sigma}$, where $\hat{\mu}$ and $\hat{\sigma}$ represent the annualized estimated mean and standard deviation of the out-of-sample returns, respectively, and $r_f = 3\%$ is the risk-free rate per annum. The stock-only strategy has a Sharpe ratio of 0.13, while the option strategy achieves a value of 0.97. These results clearly demonstrate the benefits of including options in the replicating portfolio.

We observe that all optimal portfolios $\boldsymbol{w}$ satisfy $\boldsymbol{\Gamma}(\boldsymbol{w}) \succcurlyeq \mathbf{0}$, although this was not imposed as a constraint. This implies that the delta-gamma approximation (34) of the optimal portfolio return is convex in the returns of the underlying assets. Alexander has observed this phenomenon in a simulation experiment and argues that it is a natural consequence of the risk minimization process. In fact, a portfolio with a convex payoff loses less from downward price moves and benefits disproportionately from upward price moves of the underlying assets [1].

We further observe that the optimal portfolios hold both long and short positions in options on the same underlying asset. It is known that short-sales of options can generate high expected returns (see, e.g., [9]) but they also carry considerable risk. Thus, optimal portfolios always cover the short-sale of an option by a long position in another option on the same underlying asset. On average the optimal portfolios allocate 11% of wealth in options and 89% in stocks. This implies that the high expected returns generated by the option strategy are not due to risky positions in options, but rather result from a balanced investment in a mixture of both stocks and options.

Next, we assess the *realized* VaRs of the stock-only and option strategies. These are obtained by first computing the ϵ-quantiles of all out-of-sample excess-returns of both strategies and then multiplying these values by -1 (recall that VaR measures the degree of under-performance). For $\epsilon = 5\%$ the realized VaR of the stock-only strategy amounts to 0.29%, while that of the option strategy is 0.33%. For $\epsilon = 1\%$, the realized VaR values are 0.49% and 0.54%, respectively. These results indicate that the option strategy has a slightly higher out-of-sample VaR than the stock-only strategy. However, since the option strategy achieves much higher excess-returns on average, the differences in VaR are negligible. Interestingly, the worst-case daily under-performance of the stock-only strategy is 0.78%, whilst that of the option strategy is 0.61%. Thus, the option strategy performs better in terms of worst-case under-performance relative to the benchmark.

The WCQVaR model described in Section 5 assumes the underlying asset returns to be the only sources of uncertainty in the market. It is known, however, that implied volatilities constitute important risk factors for portfolios containing options. In particular, long dated options are highly sensitive to fluctuations in the volatilities of the underlying assets. The sensitivity of the portfolio return with respect to the volatilities is commonly referred to as *vega risk*. The WCQVaR model can easily be modified to include implied volatilities as additional risk factors. The arising *delta-gamma-vega-approximation* of the portfolio return is still a quadratic function of the risk factors. Thus, the theoretical derivations in Section 5 remain valid in this generalized setting. However, estimating first- and second-order moments of the implied volatilities requires the modeling and calibration of the implied volatility surface over time, which is beyond the scope of this paper. We conjecture that extending the WCQVaR model to account for vega risk can further improve the realized VaR of the option strategy.

7 Conclusions

Derivatives depend non-linearly on their underlying assets. In this paper we generalize the WCVaR model by explicitly incorporating this non-linear relationship into the problem formulation. To this end, we developed two new models.

The WCPVaR model is suited for portfolios containing European options maturing at the investment horizon. WCPVaR expresses the option returns as convex-piecewise linear functions of the underlying assets. A benefit of this model is that it does not require knowledge of the pricing models of the options in the portfolio. However, in order to be tractably solvable, the WCPVaR model precludes short-sales of options.

The WCQVaR model can handle portfolios containing general option types and does not rely on short-sales restrictions. It exploits the popular delta-gamma approximation to model the portfolio return. In contrast to WCPVaR, WCQVaR does require knowledge of the option pricing models to determine the quadratic approximation. Through numerical experiments we demonstrate that the WCPVaR and WCQVaR models can provide much tighter VaR estimates of a portfolio containing options than the WCVaR model which does not explicitly account for non-linear dependencies between the asset returns.

We analyze the performance of the WCQVaR model in the context of an index tracking application and find that including options in the investment strategy significantly improves the out-of-sample performance. Although options are typically seen as a risky investments, our numerical results indicate that their use in a robust optimization framework can offer substantial benefits.

Acknowledgements. We are indebted to Prof. A. Ben-Tal for valuable discussions on the topic of this paper.

References

[1] C. Alexander. *Market Risk Analysis Volume IV: Value-at-Risk Models.* John Wiley & Sons, 2008.

[2] F. Alizadeh and D. Goldfarb. Second-order cone programming. *Mathematical Programming*, 95(1):3–51, 2003.

[3] P. Artzner, F. Delbaen, J. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.

[4] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust Optimization.* Princeton University Press, 2009.

[5] A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769–805, 1998.

[6] A. Ben-Tal and A. Nemirovski. Robust solutions of uncertain linear programs. *Operations Research Letters*, 25(1):1–13, 1999.

[7] F. Black and M. S. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–54, 1973.

[8] G. Calafiore, U. Topcu, and L. El Ghaoui. Parameter estimation with expected and residual-at-risk criteria. *Systems & Control Letters*, 58(1):39–46, 2009.

[9] J. Coval and T. Shumway. Expected option returns. *The Journal of Finance*, 56(3):983 – 1009, 2002.

[10] V. F. Dem'Yanov and V. N. Malozemov. *Introduction to minimax.* Keter Publishing House, Jerusalem, 1974.

[11] L. El Ghaoui, M. Oks, and F. Outstry. Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations Research*, 51(4):543–556, 2003.

[12] K. Isii. The extrema of probability determined by generalized moments (i) bounded random variables. *Annals of the Institute of Statistical Mathematics*, 12(2):119–134, 1960.

[13] S. R. Jaschke. The cornish-fisher expansion in the context of delta-gamma-normal approximations. *Journal of Risk*, 4(4):33–53, 2002.

[14] P. Jorion. *Value-at-Risk: The New Benchmark for Managing Financial Risk*. McGraw-Hill, 2001.

[15] J. Löfberg. YALMIP : A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[16] L. G. MacMillan. *Options as a strategic investment*. Prentice Hall, 1992.

[17] H. Markowitz. Portfolio selection. *Journal of Finance*, 7(1):77–91, 1952.

[18] J. Mina and A. Ulmer. Delta-gamma four ways. Technical report, RiskMetrics Group, 1999.

[19] K. Natarajan, D. Pachamanova, and M. Sim. Incorporating asymmetric distributional information in robust value-at-risk optimization. *Management Science*, 54(3):573–585, 2008.

[20] A. Nemirovski and A. Shapiro. Convex approximations of chance constrained programs. *SIAM Journal of Optimization*, 17(4):969–996, 2006.

[21] I. Polik and T. Terlaky. A survey of the S-lemma. *SIAM Review*, 49(3):371–481, 2007.

[22] B. Rustem and M. Howe. *Algorithms for Worst-Case Design and Applications to Risk Management*. Princeton University Press, 2002.

[23] A. Shapiro. *On duality theory of conic linear problems*. Semi-infinite programming: recent advances (M.A.Goberna and M.A. Lopez, eds.), Kluwer Academic Publishers, 2001.

[24] W. F. Sharpe. Mutual fund performance. *The Journal of Business*, 39(1):119–138, 1966.

[25] R. H. Tütüncü, K. C. Toh, and M. J. Todd. Solving semidefinite-quadratic-linear programs using SDPT3. *Mathematical Programming Ser. B*, 95:189–217, 2003.

[26] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38:49–95, 1996.

[27] S. Zymler, B. Rustem, and D. Kuhn. Robust portfolio optimization with derivative insurance guarantees. *Technical Report*, 2009.

A Appendix

A.1 Proof of Lemma 4.1

Define the indicator function of the set $\mathcal{S}$ as

$$\mathbb{I}_{\mathcal{S}}(\boldsymbol{\xi}) = \begin{cases} 1 & \text{if } \boldsymbol{\xi} \in \mathcal{S}, \\ 0 & \text{otherwise.} \end{cases}$$

The worst-case probability problem (20) can equivalently be expressed as

$$\begin{aligned} \pi_{\text{wc}} \quad = \quad \sup_{\mu \in \mathcal{M}_+} \quad & \int_{\mathbb{R}^n} \mathbb{I}_{\mathcal{S}}(\boldsymbol{\xi}) \mu(\mathrm{d}\boldsymbol{\xi}) \\ \text{s.t.} \quad & \int_{\mathbb{R}^n} \mu(\mathrm{d}\boldsymbol{\xi}) = 1 \\ & \int_{\mathbb{R}^n} \boldsymbol{\xi} \mu(\mathrm{d}\boldsymbol{\xi}) = \boldsymbol{\mu} \\ & \int_{\mathbb{R}^n} \boldsymbol{\xi}\boldsymbol{\xi}^T \mu(\mathrm{d}\boldsymbol{\xi}) = \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T, \end{aligned} \tag{54}$$

where $\mathcal{M}_+$ represents the cone of nonnegative Borel measures on $\mathbb{R}^n$. The optimization variable of the semi-infinite linear program (54) is the nonnegative measure μ. As can be seen, the first constraint forces μ to be a probability measure. The following constraints enforce consistency with the given first- and second-order moments, respectively.

We now assign dual variables $y_0 \in \mathbb{R}$, $\boldsymbol{y} \in \mathbb{R}^n$, and $\mathbf{Y} \in \mathbb{S}^n$ to the equality constraints in (54), respectively, and introduce the following dual problem (see, e.g., [23]).

$$\begin{aligned} \inf \quad & y_0 + \boldsymbol{y}^T \boldsymbol{\mu} + \langle \mathbf{Y}, \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T \rangle \\ \text{s.t.} \quad & y_0 \in \mathbb{R}, \quad \boldsymbol{y} \in \mathbb{R}^n, \quad \mathbf{Y} \in \mathbb{S}^n \\ & y_0 + \boldsymbol{y}^T \boldsymbol{\xi} + \langle \mathbf{Y}, \boldsymbol{\xi}\boldsymbol{\xi}^T \rangle \geq \mathbb{I}_{\mathcal{S}}(\boldsymbol{\xi}) \quad \forall \boldsymbol{\xi} \in \mathbb{R}^n \end{aligned} \tag{55}$$

Because $\boldsymbol{\Sigma} \succ \mathbf{0}$, it can be shown that strong duality holds [12]. Therefore the worst-case probability π_{wc} coincides with the optimal value of the dual problem (55).

By defining

$$\mathbf{M} = \begin{bmatrix} \mathbf{Y} & \frac{1}{2}\boldsymbol{y} \\ \frac{1}{2}\boldsymbol{y}^T & y_0 \end{bmatrix},$$

problem (55) can be reformulated as

$$
\begin{aligned}
\inf_{\mathbf{M}\in\mathbb{S}^{n+1}} \quad & \langle \boldsymbol{\Omega}, \mathbf{M} \rangle \\
\text{s.t.} \quad & \left[\boldsymbol{\xi}^T \ 1\right] \mathbf{M} \left[\boldsymbol{\xi}^T \ 1\right]^T \geq \mathbb{I}_{\mathcal{S}}(\boldsymbol{\xi}) \quad \forall \boldsymbol{\xi} \in \mathbb{R}^n.
\end{aligned} \tag{56}
$$

By definition of $\mathbb{I}_{\mathcal{S}}(\boldsymbol{\xi})$, the constraint in (56) can be expanded in terms of two semi-infinite constraints.

$$
\left[\boldsymbol{\xi}^T \ 1\right] \mathbf{M} \left[\boldsymbol{\xi}^T \ 1\right]^T \geq 0 \quad \forall \boldsymbol{\xi} \in \mathbb{R}^n \tag{57a}
$$

$$
\left[\boldsymbol{\xi}^T \ 1\right] \mathbf{M} \left[\boldsymbol{\xi}^T \ 1\right]^T \geq 1 \quad \forall \boldsymbol{\xi} \in \mathcal{S} \tag{57b}
$$

Since (57a) is equivalent to $\mathbf{M} \succcurlyeq \mathbf{0}$, the claim follows. ∎

A.2 Proof of Theorem 4.1

In order to obtain the postulated SOCP reformulation, we calculate the dual associated with problem (27), which, after some simplification steps, reduces to

$$
\begin{aligned}
\text{WCPVaR}_\epsilon(\boldsymbol{w}) = \max \quad & (\boldsymbol{e} - \boldsymbol{\delta})^T \boldsymbol{w}^{\boldsymbol{\eta}} - 2\boldsymbol{m}^T \boldsymbol{w}^{\boldsymbol{\xi}} \\
\text{s.t.} \quad & \alpha \in \mathbb{R}, \quad \boldsymbol{\delta} \in \mathbb{R}^{m-n}, \quad \boldsymbol{m} \in \mathbb{R}^n, \quad \mathbf{Z} \in \mathbb{S}^n \\
& 0 \leq \alpha \leq \frac{1}{2\epsilon}, \quad \alpha\boldsymbol{\Omega} \succcurlyeq \mathbf{Y} = \begin{bmatrix} \mathbf{Z} & \boldsymbol{m} \\ \boldsymbol{m}^T & 1/2 \end{bmatrix} \succcurlyeq \mathbf{0}, \\
& \boldsymbol{\delta} - 2\mathbf{B}\boldsymbol{m} - \boldsymbol{a} \geq \mathbf{0}, \quad \boldsymbol{\delta} \geq \mathbf{0}.
\end{aligned} \tag{58}
$$

Note that problem (58) is strictly feasible, which implies that strong conic duality holds [26]. This confirms that the optimal value of the dual problem (58) exactly matches the WCPVaR.

By the definition of $\boldsymbol{\Omega}$ in (22), we may conclude that

$$
\alpha\boldsymbol{\Omega} \succcurlyeq \mathbf{Y} \iff \begin{bmatrix} \alpha(\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T) - \mathbf{Z} & \alpha\boldsymbol{\mu} - \boldsymbol{m} \\ (\alpha\boldsymbol{\mu} - \boldsymbol{m})^T & \alpha - 1/2 \end{bmatrix} \succcurlyeq \mathbf{0} \implies \alpha \geq 1/2.
$$

This allows us to divide the matrix inequality in problem (58) by α. Subsequently, we apply the variable substitution $(\mathbf{Z}, \boldsymbol{m}, \alpha) \to (\mathbf{V}, \boldsymbol{v}, y)$ with $\mathbf{V} = \mathbf{Z}/\alpha$, $\boldsymbol{v} = \boldsymbol{m}/\alpha$, and $y = \frac{1}{2\alpha} \in [\epsilon, 1]$. We thus obtain the following problem reformulation.

$$
\begin{aligned}
\text{WCPVaR}_\epsilon(\boldsymbol{w}) = \max \quad & (\boldsymbol{e}-\boldsymbol{\delta})^T \boldsymbol{w}^{\boldsymbol{\eta}} - \frac{\boldsymbol{v}^T \boldsymbol{w}^{\boldsymbol{\xi}}}{y} \\
\text{s.t.} \quad & y \in \mathbb{R}, \quad \boldsymbol{\delta} \in \mathbb{R}^{m-n}, \quad \boldsymbol{v} \in \mathbb{R}^n, \quad \mathbf{V} \in \mathbb{S}^n \\
& \epsilon \leq y \leq 1, \quad \boldsymbol{\Omega} \succcurlyeq \begin{bmatrix} \mathbf{V} & \boldsymbol{v} \\ \boldsymbol{v}^T & y \end{bmatrix} \succcurlyeq \mathbf{0} \\
& \boldsymbol{\delta} \geq \frac{\mathbf{B}\boldsymbol{v}}{y} + \boldsymbol{a}, \quad \boldsymbol{\delta} \geq \mathbf{0}
\end{aligned}
\tag{59}
$$

Assume first that $y = 1$ at optimality. Then, by the definition of $\boldsymbol{\Omega}$ and the linear matrix inequality in problem (59), we find $\boldsymbol{v} = \boldsymbol{\mu}$, while (59) reduces to

$$
\begin{aligned}
& \max_{\boldsymbol{\delta} \in \mathbb{R}^{m-n}} \left\{ (\boldsymbol{e}-\boldsymbol{\delta})^T \boldsymbol{w}^{\boldsymbol{\eta}} - \boldsymbol{\mu}^T \boldsymbol{w}^{\boldsymbol{\xi}} \ : \ \boldsymbol{\delta} \geq \boldsymbol{a} + \mathbf{B}\boldsymbol{\mu}, \ \boldsymbol{\delta} \geq \mathbf{0} \right\} \\
&= -\,\boldsymbol{\mu}^T \boldsymbol{w}^{\boldsymbol{\xi}} - (\max\{-\boldsymbol{e}, \boldsymbol{a} + \mathbf{B}\boldsymbol{\mu} - \boldsymbol{e}\})^T \boldsymbol{w}^{\boldsymbol{\eta}} \\
&= -\, f(\boldsymbol{\mu})^T \boldsymbol{w}.
\end{aligned}
\tag{60}
$$

Assume now that $y < 1$ at optimality. By the definition of $\boldsymbol{\Omega}$ and by using Schur complements, we find

$$
\begin{aligned}
\boldsymbol{\Omega} \succcurlyeq \begin{bmatrix} \mathbf{V} & \boldsymbol{v} \\ \boldsymbol{v}^T & y \end{bmatrix} \iff & \begin{bmatrix} \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T - \mathbf{V} & \boldsymbol{\mu} - \boldsymbol{v} \\ (\boldsymbol{\mu} - \boldsymbol{v})^T & 1 - y \end{bmatrix} \succcurlyeq \mathbf{0} \\
\iff & \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T - \mathbf{V} - \frac{1}{1-y}(\boldsymbol{\mu} - \boldsymbol{v})(\boldsymbol{\mu} - \boldsymbol{v})^T \succcurlyeq \mathbf{0}.
\end{aligned}
\tag{61a}
$$

A similar argument yields the equivalence

$$
\begin{bmatrix} \mathbf{V} & \boldsymbol{v} \\ \boldsymbol{v}^T & y \end{bmatrix} \succcurlyeq \mathbf{0} \iff \mathbf{V} - \frac{1}{y}\boldsymbol{v}\boldsymbol{v}^T \succcurlyeq \mathbf{0}.
\tag{61b}
$$

By combining (61a) and (61b), the linear matrix inequality constraints in problem (59) are equivalent to

$$
\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T - \frac{1}{1-y}(\boldsymbol{\mu} - \boldsymbol{v})(\boldsymbol{\mu} - \boldsymbol{v})^T \succcurlyeq \mathbf{V} \succcurlyeq \frac{1}{y}\boldsymbol{v}\boldsymbol{v}^T.
$$

The decision variable $\mathbf{V}$ can now be eliminated from the problem, while the linear matrix inequality constraints in (59) can be replaced by

$$\begin{aligned} &\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T \succcurlyeq \frac{1}{1-y}(\boldsymbol{\mu} - \boldsymbol{v})(\boldsymbol{\mu} - \boldsymbol{v})^T + \frac{1}{y}\boldsymbol{v}\boldsymbol{v}^T \\ \iff &\boldsymbol{\Sigma} \succcurlyeq \frac{1}{y(1-y)}(\boldsymbol{v} - y\boldsymbol{\mu})(\boldsymbol{v} - y\boldsymbol{\mu})^T. \end{aligned} \tag{62}$$

The above arguments imply that problem (59) can be reformulated as

$$\mathrm{WCPVaR}_\epsilon(\boldsymbol{w}) = \max\{\phi(y) \ : \ y \in [\epsilon, 1]\},$$

where

$$\begin{aligned} \phi(y) = \max \quad & (\boldsymbol{e} - \boldsymbol{\delta})^T \boldsymbol{w}^{\boldsymbol{\eta}} - \frac{\boldsymbol{v}^T \boldsymbol{w}^{\boldsymbol{\xi}}}{y} \\ \text{s. t.} \quad & \boldsymbol{\delta} \in \mathbb{R}^{m-n}, \quad \boldsymbol{v} \in \mathbb{R}^n \\ & \boldsymbol{\Sigma} \succcurlyeq \frac{1}{y(1-y)}(\boldsymbol{v} - y\boldsymbol{\mu})(\boldsymbol{v} - y\boldsymbol{\mu})^T, \\ & \boldsymbol{\delta} \geq \frac{\mathbf{B}\boldsymbol{v}}{y} + \boldsymbol{a}, \quad \boldsymbol{\delta} \geq \mathbf{0}. \end{aligned} \tag{63}$$

For any fixed $y \in [\epsilon, 1)$, we have that $y^{-1}(1-y)^{-1} > 0$, and the linear matrix inequality in (63) can be rewritten as

$$\begin{bmatrix} \boldsymbol{\Sigma} & \boldsymbol{v} - y\boldsymbol{\mu} \\ (\boldsymbol{v} - y\boldsymbol{\mu})^T & y(1-y) \end{bmatrix} \succcurlyeq \mathbf{0}.$$

Since $\boldsymbol{\Sigma} \succ \mathbf{0}$, this linear matrix inequality holds if and only if

$$(\boldsymbol{v} - y\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{v} - y\boldsymbol{\mu}) \leq y(1-y),$$

which is equivalent to the second-order cone constraint

$$\left\| \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{v} - y\boldsymbol{\mu}) \right\|_2 \leq \sqrt{y(1-y)}.$$

For $y \in [\epsilon, 1)$, the value of $\phi(y)$ can thus be found by solving the following SOCP.

$$
\begin{aligned}
\phi(y) = \max \quad & (\boldsymbol{e} - \boldsymbol{\delta})^T \boldsymbol{w}^{\boldsymbol{\eta}} - \frac{\boldsymbol{v}^T \boldsymbol{w}^{\boldsymbol{\xi}}}{y} \\
\text{s.t.} \quad & \boldsymbol{\delta} \in \mathbb{R}^{m-n}, \quad \boldsymbol{v} \in \mathbb{R}^n \\
& \left\| \boldsymbol{\Sigma}^{-1/2} (\boldsymbol{v} - y\boldsymbol{\mu}) \right\|_2 \leq \sqrt{y(1-y)} \\
& \boldsymbol{\delta} \geq \frac{\mathbf{B}\boldsymbol{v}}{y} + \boldsymbol{a}, \quad \boldsymbol{\delta} \geq \mathbf{0}
\end{aligned}
\tag{64}
$$

Note that the above problem is strictly feasible for $y \in [\epsilon, 1)$. By strong conic duality the associated dual problem has the same optimal value [2]. We thus obtain that $\phi(y) = \phi'(y)$ for $y \in [\epsilon, 1)$, where

$$
\phi'(y) = \min_{\mathbf{0} \leq \boldsymbol{g} \leq \boldsymbol{w}^{\boldsymbol{\eta}}} \; -\boldsymbol{\mu}^T (\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T \boldsymbol{g}) + \sqrt{\frac{1-y}{y}} \left\| \boldsymbol{\Sigma}^{1/2} (\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T \boldsymbol{g}) \right\|_2 - \boldsymbol{a}^T \boldsymbol{g} + \boldsymbol{e}^T \boldsymbol{w}^{\boldsymbol{\eta}}.
$$

Note that for $y = 1$, we also have $\phi(1) = \phi'(1)$ since

$$
\begin{aligned}
\phi'(1) &= \min_{\mathbf{0} \leq \boldsymbol{g} \leq \boldsymbol{w}^{\boldsymbol{\eta}}} \; -\boldsymbol{\mu}^T (\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T \boldsymbol{g}) - \boldsymbol{a}^T \boldsymbol{g} + \boldsymbol{e}^T \boldsymbol{w}^{\boldsymbol{\eta}} \\
&= -\boldsymbol{\mu}^T \boldsymbol{w}^{\boldsymbol{\xi}} - (\max\{-\boldsymbol{e}, \boldsymbol{a} + \mathbf{B}\boldsymbol{\mu} - \boldsymbol{e}\})^T \boldsymbol{w}^{\boldsymbol{\eta}} \\
&= \phi(1),
\end{aligned}
$$

where the second equality follows from (60). Maximizing $\phi(y)$ over y yields the desired WCPVaR value. Since $\sqrt{(1-y)/y}$ is monotonically decreasing in y, we have $y = \epsilon$ at optimality. This results in the following optimization problem

$$
\text{WCPVaR}_\epsilon(\boldsymbol{w}) = \min_{\mathbf{0} \leq \boldsymbol{g} \leq \boldsymbol{w}^{\boldsymbol{\eta}}} \; -\boldsymbol{\mu}^T (\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T \boldsymbol{g}) + \sqrt{\frac{1-\epsilon}{\epsilon}} \left\| \boldsymbol{\Sigma}^{1/2} (\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T \boldsymbol{g}) \right\|_2 - \boldsymbol{a}^T \boldsymbol{g} + \boldsymbol{e}^T \boldsymbol{w}^{\boldsymbol{\eta}},
$$

which is the postulated reformulation of WCPVaR as the optimal value of a SOCP. ∎