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.

June 21, 2012

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 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 (WPVaR) and Worst-Case Quadratic VaR (WQVaR) approximations, respectively. WPVaR serves as a VaR approximation for portfolios containing long positions in European options expiring at the end of the investment horizon, whereas WQVaR is suitable for portfolios containing long and/or short positions in European and/or exotic options expiring beyond the investment horizon. We prove that—unlike VaR that may discourage diversification—WPVaR and WQVaR are in fact coherent risk measures. We also reveal connections to robust portfolio optimization.

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 [16], who observed that a prudent investor does not aim solely at maximizing the expected return of an investment, but 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 counterintuitive asset allocations when the portfolio

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

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 least $(1-\epsilon)$-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 (WVaR) represents a conservative (that is, pessimistic) approximation for the true (unknown) portfolio VaR. In contrast to VaR, WVaR 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, 23]: WVaR 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, WVaR 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 WVaR are the following.

- The calculation of WVaR 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.

- WVaR 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 WVaR model, fails to capture non-linear dependencies among the asset returns, and therefore WVaR tends to severely *overestimate* the true VaR of a portfolio containing derivatives.

Recall that WVaR 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 WVaR from a geometrical viewpoint. An intuitively appealing uncertainty set should be asymmetric to reflect the skewness of the derivative returns. Recently, Natarajan *et al.* [18] included asymmetric distributional information into the WVaR 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 new Worst-Case VaR models which explicitly account for perfect non-linear dependencies between the asset returns. We first introduce the *Worst-Case Polyhedral VaR* (WPVaR), which provides a 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. WPVaR 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 WPVaR optimization as a convex second-order cone program, which can be solved efficiently [2]. We also establish the equivalence of the WPVaR model to a robust optimization model described in [29].

Next, we introduce the *Worst-Case Quadratic VaR* (WQVaR) 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 WPVaR, WQVaR 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 [17]. However, to the best of our knowledge, the delta-gamma approximation has never been used in a VaR optimization model. We define WQVaR 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. WQVaR provides a tight conservative approximation for the true portfolio VaR if the delta-gamma approximation is accurate. We show that WQVaR optimization can be formulated as a convex semidefinite program, which can be solved efficiently [25], and we establish a connection to a new robust optimization problem.

In the aftermath of the financial crisis in 2008 VaR has been heavily criticized for failing to satisfy the subadditivity property of coherent risk measures, therefore occasionally discouraging diversification. One of the main insights of this work is that the WVaR by El Ghaoui *et al.* [11] as well as the new WPVaR and WQVaR measures proposed here are in fact coherent risk measures in the sense of Artzner *et al.* [3]. This will be established by proving that WPVaR (WQVaR) coincides with the worst-case *Conditional* Value-at-Risk of the underlying polyhedral (quadratic) loss function.

The main contributions of this paper can be summarized as follows:

(1) We generalize the WVaR 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 WPVaR and WQVaR models as described above. We show that in the absence of derivatives both models reduce to the WVaR model. Moreover, we formulate WPVaR optimization as a second-order cone program and WQVaR optimization as a semidefinite program. Both models are polynomial-time solvable.

(2) We show that both the WPVaR and the WQVaR 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 WVaR in the presence of derivatives.

(3) We establish that WPVaR (WQVaR) coincides with the Worst-Case *Conditional* Value-at-Risk of the underlying polyhedral (quadratic) loss function, and we demonstrate that WPVaR and WQVaR represent *coherent* risk measures in the sense of Artzner *et al.* [3].

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 $\tilde{r}_p = \boldsymbol{w}^T \tilde{\boldsymbol{r}}$. 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, $\text{VaR}_\epsilon(\boldsymbol{w})$ 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,

$$\text{VaR}_\epsilon(\boldsymbol{w}) = \min\left\{\gamma \ : \ \mathbb{P}\{\gamma \leq -\boldsymbol{w}^T\tilde{\boldsymbol{r}}\} \leq \epsilon\right\}, \tag{1}$$

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

In this paper we investigate portfolio optimization problems of the type

$$\begin{array}{ll} \underset{\boldsymbol{w}\in\mathbb{R}^m}{\text{minimize}} & \text{VaR}_\epsilon(\boldsymbol{w}) \\ \text{subject to} & \boldsymbol{w} \in \mathcal{W}, \end{array} \tag{2}$$

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 (1), the VaR optimization model (2) can be reformulated as

$$\begin{array}{ll} \underset{\boldsymbol{w}\in\mathbb{R}^m,\gamma\in\mathbb{R}}{\text{minimize}} & \gamma \\ \text{subject to} & \mathbb{P}\{\gamma + \boldsymbol{w}^T\tilde{\boldsymbol{r}} \geq 0\} \geq 1-\epsilon \\ & \boldsymbol{w} \in \mathcal{W}, \end{array} \tag{3}$$

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 (2) 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 $\boldsymbol{\mu}_{\boldsymbol{r}} \in \mathbb{R}^m$ and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{r}} \in \mathbb{S}^m$. This assumption has the advantage that the VaR can be calculated analytically as

$$\mathrm{VaR}_\epsilon(\boldsymbol{w}) = -\boldsymbol{\mu}_{\boldsymbol{r}}^T \boldsymbol{w} - \Phi^{-1}(\epsilon)\sqrt{\boldsymbol{w}^T \boldsymbol{\Sigma}_{\boldsymbol{r}} \boldsymbol{w}}, \tag{4}$$

where Φ is the standard normal distribution function. This model is sometimes referred to as *Normal VaR* (see, e.g., [18]). In practice, the distribution of the asset returns often fails to be Gaussian. In these cases, (4) 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

$$\mathrm{WVaR}_\epsilon(\boldsymbol{w}) = \min\left\{\gamma \ : \ \sup_{\mathbb{P}\in\mathcal{P}_{\boldsymbol{r}}} \mathbb{P}\{\gamma \le -\boldsymbol{w}^T\tilde{\boldsymbol{r}}\} \le \epsilon\right\}. \tag{5}$$

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

$$\mathrm{WVaR}_\epsilon(\boldsymbol{w}) = -\boldsymbol{\mu}_{\boldsymbol{r}}^T \boldsymbol{w} + \kappa(\epsilon)\sqrt{\boldsymbol{w}^T \boldsymbol{\Sigma}_{\boldsymbol{r}} \boldsymbol{w}}, \tag{6}$$

where $\kappa(\epsilon) = \sqrt{(1-\epsilon)/\epsilon}$. WVaR 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 WVaR.

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

$$\begin{aligned} \underset{\boldsymbol{w}\in\mathbb{R}^m}{\text{minimize}} \quad & -\boldsymbol{\mu}_{\boldsymbol{r}}^T \boldsymbol{w} + \kappa(\epsilon)\left\|\boldsymbol{\Sigma}_{\boldsymbol{r}}^{1/2}\boldsymbol{w}\right\|_2 \\ \text{subject to} \quad & \boldsymbol{w} \in \mathcal{W}, \end{aligned} \tag{7}$$

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.

$$\begin{aligned} \underset{\boldsymbol{w}\in\mathbb{R}^m,\gamma\in\mathbb{R}}{\text{minimize}} \quad & \gamma \\ \text{subject to} \quad & \gamma + \boldsymbol{w}^T\tilde{\boldsymbol{r}} \ge 0 \\ & \boldsymbol{w} \in \mathcal{W} \end{aligned} \tag{8}$$

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, (8) fails to provide a unique implementable investment decision. One way to disambiguate this model is to require that the explicit inequality constraint in (8) is satisfied with a given probability. By using this approach, we recover the chance-constrained stochastic program (3). Robust optimization [5, 4, 7] 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 (8) is then defined as

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

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 (9) is therefore closely related to the chance constraint in the stochastic program (3). 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 $(\boldsymbol{w}, \gamma) \in \mathcal{W} \times \mathbb{R}$.

$$\gamma + \boldsymbol{w}^T\boldsymbol{r} \geq 0 \quad \forall \boldsymbol{r} \in \mathcal{U} \iff -\boldsymbol{\mu}_{\boldsymbol{r}}^T\boldsymbol{w} + \delta \left\| \boldsymbol{\Sigma}_{\boldsymbol{r}}^{1/2}\boldsymbol{w} \right\|_2 \leq \gamma \tag{10}$$

Problem (9) can therefore be reformulated as the following second-order cone program.

$$\begin{aligned} \underset{\boldsymbol{w}\in\mathbb{R}^m}{\text{minimize}} \quad & -\boldsymbol{\mu}_{\boldsymbol{r}}^T\boldsymbol{w} + \delta \left\| \boldsymbol{\Sigma}_{\boldsymbol{r}}^{1/2}\boldsymbol{w} \right\|_2 \\ \text{subject to} \quad & \boldsymbol{w} \in \mathcal{W} \end{aligned} \tag{11}$$

By comparing (7) and (11), El Ghaoui *et al.* [11] noticed that optimizing WVaR at level ϵ is equivalent to solving the robust optimization problem (9) under an ellipsoidal uncertainty set with size parameter $\delta = \kappa(\epsilon)$, see also Natarajan *et al.* [18]. This uncertainty set will henceforth be denoted by $\mathcal{U}_\epsilon$.

In this paper we extend the WVaR model (6) and the equivalent robust optimization model (9) 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 the 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 WVaR model (6), which has the advantage of computational tractability and accounts for the absence of distributional information beyond first- and second-order moments. However, WVaR 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 WVaR model, are difficult to estimate reliably from historical data, see, e.g., [10]. 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, WVaR 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, WVaR 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, and disregards the perfect dependencies captured by the function f, which are typically known. When f is non-linear, WVaR tends to severely *overestimate* the actual VaR since the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{r}}$ accounts only for *linear* dependencies. The robust optimization perspective on WVaR 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 $\boldsymbol{\mu}_{\boldsymbol{r}}$. 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 WVaR 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

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

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 WVaR 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.

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 the 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 $j = 1, \dots, n$. Next, we investigate the option returns $\tilde{r}_j$ for $j = n+1, \dots, m$. 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 $\tilde{\xi}_i$,

$$\begin{aligned} f_j(\tilde{\boldsymbol{\xi}}) &= \frac{1}{c_j} \max\left\{0, s_i(1+\tilde{\xi}_i) - k_j\right\} - 1 \\ &= \max\left\{-1, a_j + b_j\tilde{\xi}_i - 1\right\}, \quad \text{where} \quad a_j = \frac{s_i - k_j}{c_j} \quad \text{and} \quad b_j = \frac{s_i}{c_j}. \end{aligned} \tag{12a}$$

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,

$$f_j(\tilde{\boldsymbol{\xi}}) = \max\left\{-1, a_j + b_j\tilde{\xi}_i - 1\right\}, \quad \text{where} \quad a_j = \frac{k_j - s_i}{p_j} \quad \text{and} \quad b_j = -\frac{s_i}{p_j}. \tag{12b}$$

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

$$\tilde{\boldsymbol{r}} = f(\tilde{\boldsymbol{\xi}}) = \begin{pmatrix} \tilde{\boldsymbol{\xi}} \\ \max\left\{-\boldsymbol{e}, \boldsymbol{a} + \mathbf{B}\tilde{\boldsymbol{\xi}} - \boldsymbol{e}\right\} \end{pmatrix}, \tag{13}$$

where $\boldsymbol{a} \in \mathbb{R}^{m-n}$, $\mathbf{B} \in \mathbb{R}^{(m-n)\times n}$ are known constants determined through (12a) and (12b), $\boldsymbol{e} \in \mathbb{R}^{m-n}$ 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

$$\begin{aligned} \tilde{r}_p &= \boldsymbol{w}^T\tilde{\boldsymbol{r}} = (\boldsymbol{w}^{\boldsymbol{\xi}})^T\tilde{\boldsymbol{\xi}} + (\boldsymbol{w}^{\boldsymbol{\eta}})^T\tilde{\boldsymbol{\eta}} \\ &= \boldsymbol{w}^T f(\tilde{\boldsymbol{\xi}}) = (\boldsymbol{w}^{\boldsymbol{\xi}})^T\tilde{\boldsymbol{\xi}} + (\boldsymbol{w}^{\boldsymbol{\eta}})^T \max\left\{-\boldsymbol{e}, \boldsymbol{a} + \mathbf{B}\tilde{\boldsymbol{\xi}} - \boldsymbol{e}\right\}. \end{aligned} \tag{14}$$

4.2 Worst-Case Polyhedral VaR Model

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

$$\begin{aligned} \mathrm{WPVaR}_\epsilon(\boldsymbol{w}) &= \min\left\{\gamma \ : \ \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\left\{\gamma \le -\boldsymbol{w}^T f(\tilde{\boldsymbol{\xi}})\right\} \le \epsilon\right\} \qquad (15)\\ &= \min\left\{\gamma \ : \ \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\left\{\gamma \le -(\boldsymbol{w}^{\boldsymbol{\xi}})^T\tilde{\boldsymbol{\xi}} - (\boldsymbol{w}^{\boldsymbol{\eta}})^T \max\left\{-\boldsymbol{e}, \boldsymbol{a} + \mathbf{B}\tilde{\boldsymbol{\xi}} - \boldsymbol{e}\right\}\right\} \le \epsilon\right\}, \end{aligned}$$

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 $\boldsymbol{\Sigma}$. WPVaR provides an accurate conservative approximation for the VaR of a portfolio whose return constitutes a convex piecewise linear (i.e., polyhedral) function of the basic asset returns. We now demonstrate that WPVaR can be evaluated efficiently as the optimal value of a tractable semidefinite program (SDP).

Theorem 4.1 *The* WPVaR *of a fixed portfolio* $\boldsymbol{w} \in \mathcal{W}$ *can be expressed as*

$$\begin{aligned} \mathrm{WPVaR}_\epsilon(\boldsymbol{w}) = \inf \quad & \gamma \\ \text{s.t.} \quad & \mathbf{M} \in \mathbb{S}^{n+1}, \quad \boldsymbol{y} \in \mathbb{R}^{m-n}, \quad \tau \in \mathbb{R}, \quad \gamma \in \mathbb{R} \\ & \langle \boldsymbol{\Omega}, \mathbf{M}\rangle \le \tau\epsilon, \quad \mathbf{M} \succcurlyeq \mathbf{0}, \quad \tau \ge 0, \quad \mathbf{0} \le \boldsymbol{y} \le \boldsymbol{w}^{\boldsymbol{\eta}} \\ & \mathbf{M} + \begin{bmatrix} \mathbf{0} & \boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T\boldsymbol{y} \\ (\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T\boldsymbol{y})^T & -\tau + 2(\gamma + \boldsymbol{y}^T\boldsymbol{a} - \boldsymbol{e}^T\boldsymbol{w}^{\boldsymbol{\eta}}) \end{bmatrix} \succcurlyeq \mathbf{0}, \end{aligned} \tag{16}$$

where

$$\boldsymbol{\Omega} = \begin{bmatrix} \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T & \boldsymbol{\mu} \\ \boldsymbol{\mu}^T & 1 \end{bmatrix} \tag{17}$$

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

Proof: See Appendix B.

Even though (16) constitutes a tractable SDP that enables us to compute the WPVaR of a given portfolio $\boldsymbol{w} \in \mathcal{W}$ in polynomial time, it would be desirable to obtain an equivalent second-order cone program (SOCP) because SOCPs exhibit better scalability properties than SDPs [2]. In Theorem 4.2 we demonstrate that such a reformulation exists.

Theorem 4.2 *Problem* (16) *can be reformulated as*

$$\mathrm{WPVaR}_\epsilon(\boldsymbol{w}) = \min_{\mathbf{0}\leq \boldsymbol{g} \leq \boldsymbol{w}^{\boldsymbol{\eta}}} \; -\boldsymbol{\mu}^T(\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T\boldsymbol{g}) + \kappa(\epsilon)\left\|\boldsymbol{\Sigma}^{1/2}(\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T\boldsymbol{g})\right\|_2 - \boldsymbol{a}^T\boldsymbol{g} + \boldsymbol{e}^T\boldsymbol{w}^{\boldsymbol{\eta}}, \quad (18)$$

which constitutes a tractable SOCP.

Proof: The proof follows a similar reasoning as in [11, Theorem 1] and is thus omitted. ∎

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*

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

Thus, the WPVaR *model encapsulates the* WVaR *model* (6) *as a special case.*

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

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

which is equivalent to the tractable SOCP

$$\begin{array}{ll} \text{minimize} & \gamma \\ \text{subject to} & \boldsymbol{w}^{\boldsymbol{\xi}} \in \mathbb{R}^n, \quad \boldsymbol{w}^{\boldsymbol{\eta}} \in \mathbb{R}^{m-n}, \quad \boldsymbol{g} \in \mathbb{R}^{m-n}, \quad \gamma \in \mathbb{R} \\ & -\boldsymbol{\mu}^T(\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T\boldsymbol{g}) + \kappa(\epsilon)\left\|\boldsymbol{\Sigma}^{1/2}(\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T\boldsymbol{g})\right\|_2 - \boldsymbol{a}^T\boldsymbol{g} + \boldsymbol{e}^T\boldsymbol{w}^{\boldsymbol{\eta}} \leq \gamma \\ & \mathbf{0} \leq \boldsymbol{g} \leq \boldsymbol{w}^{\boldsymbol{\eta}}, \quad \boldsymbol{w} = (\boldsymbol{w}, \boldsymbol{w}^{\boldsymbol{\eta}}), \quad \boldsymbol{w} \in \mathcal{W}. \end{array} \quad (19)$$

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

Remark 4.2 (General Maturities) *The returns of European call and put options expiring beyond the end of the investment horizon constitute convex and asymptotically linear functions of* $\tilde{\boldsymbol{\xi}}$ *that can be inferred from option pricing models. Hence, they can be uniformly approximated by convex piecewise linear functions with finitely many pieces. If a market consists only of basic assets and European-style options expiring at or beyond time* T, *we may thus assume, without much loss of accuracy, that the return function* $f(\tilde{\boldsymbol{\xi}})$ *is piecewise linear. Our definition* (15) *of* WPVaR *for long-only portfolios as well as the corresponding tractable SOCP reformulation* (18) *can be extended to this generalized setting in a straightforward manner. Using this generalized* WPVaR *model allows investors to benefit from a greater variety of options traded in the market.*

Remark 4.3 (Second Layer of Robustness) *So far we have assumed that the mean values and covariances of* $\tilde{\boldsymbol{\xi}}$ *are precisely known. However, if these statistics need to be estimated from noisy data, the moment matrix* $\boldsymbol{\Omega}$ *may only be known to lie within a box-type confidence set of the form* $\mathcal{O} = \left\{\boldsymbol{\Omega} \in \mathbb{S}^{n+1} \ : \ \boldsymbol{\Omega} \succcurlyeq \mathbf{0},\ \underline{\boldsymbol{\Omega}} \leq \boldsymbol{\Omega} \leq \overline{\boldsymbol{\Omega}}\right\}$ *for some given* $\underline{\boldsymbol{\Omega}}$ *and* $\overline{\boldsymbol{\Omega}}$ *of appropriate dimensions. In this situation,* WPVaR *can be equipped with a second layer of robustness that hedges against uncertainty in* $\boldsymbol{\Omega}$*. If* $\mathcal{O}$ *contains a strictly positive definite matrix, we can use methods proposed in [20, Section 2.5] to reformulate the resulting doubly robust risk measure* $\sup_{\boldsymbol{\Omega}\in\mathcal{O}} \mathrm{WPVaR}_\epsilon(\boldsymbol{w})$ *as the optimal value of an SDP, which can be solved efficiently.*

Remark 4.4 (Partitioned Statistics) *The* WPVaR *model developed in this section naturally extends to richer families of asset return distributions. For instance, one can represent the vector of asset returns as* $\tilde{\boldsymbol{\xi}} = \tilde{\boldsymbol{\xi}}_+ - \tilde{\boldsymbol{\xi}}_-$ *where* $\tilde{\boldsymbol{\xi}}_+$ *and* $\tilde{\boldsymbol{\xi}}_-$ *denote the positive and negative parts of* $\tilde{\boldsymbol{\xi}}$*, respectively. Using this convention, it is possible to introduce a variant of* WPVaR *which hedges against all distributions of* $\tilde{\boldsymbol{\xi}}_+$ *and* $\tilde{\boldsymbol{\xi}}_-$ *that are supported on the nonnegative orthant and share the same first and second-order moments. This sort of partitioned statistics enables us to capture distributional asymmetry, see [20]. An efficiently computable SDP-based upper bound on* WPVaR *with partitioned statistics information can be constructed by extending the techniques developed in [20, Section 2.4] in a straightforward manner.*

4.3 Robust Optimization Perspective on WPVaR

In Section 2 we highlighted a known relationship between WVaR optimization and robust optimization. Moreover, in Section 3 we argued that the ellipsoidal uncertainty set related to the WVaR 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 WPVaR minimization problem (19) can also be cast as a robust optimization problem of the type (9). However, the uncertainty set which generates WPVaR is no longer symmetric.

Theorem 4.3 *The* WPVaR *minimization problem* (19) *is equivalent to the robust optimization problem*

$$\begin{array}{ll} \underset{\boldsymbol{w}\in\mathbb{R}^m,\gamma\in\mathbb{R}}{\text{minimize}} & \gamma \\ \text{subject to} & -\boldsymbol{w}^T\boldsymbol{r} \leq \gamma \qquad \forall \boldsymbol{r} \in \mathcal{U}^p_\epsilon \\ & \boldsymbol{w} \in \mathcal{W}, \end{array} \tag{20}$$

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

$$\mathcal{U}^p_\epsilon = \left\{\boldsymbol{r} \in \mathbb{R}^m \ : \ \exists \boldsymbol{\xi} \in \mathbb{R}^n,\ (\boldsymbol{\xi}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{\xi}-\boldsymbol{\mu}) \leq \kappa(\epsilon)^2,\ \boldsymbol{r} = f(\boldsymbol{\xi})\right\}. \tag{21}$$

Proof: The result is based on conic duality. We refer to [29, Theorem 3.1] for a complete exposition of the proof. ∎

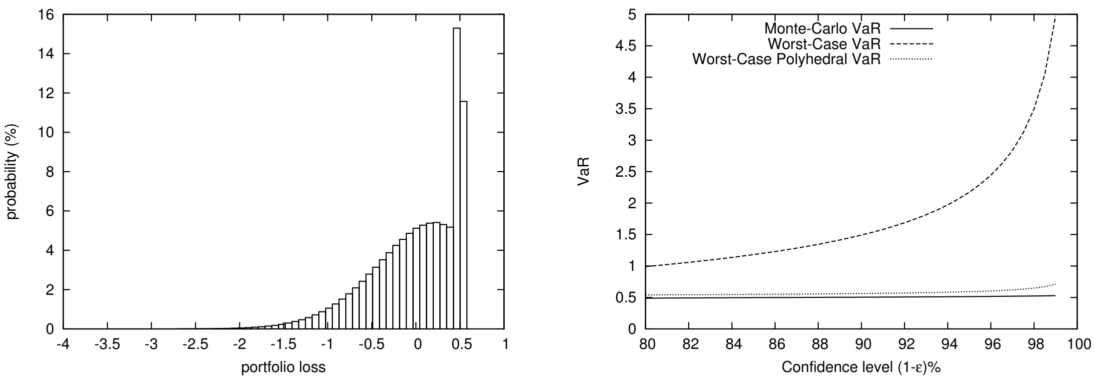

Figure 1: Left: The portfolio loss distribution obtained via Monte-Carlo simulation. Note that negative values represent gains. Right: The VaR estimates at different confidence levels obtained via Monte-Carlo sampling, WVaR, and WPVaR.

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* $\boldsymbol{w}$ *be an equally weighted portfolio of these* $m = 4$ *assets, that is, set* $w_i = 1/m$ *for* $i = 1, \ldots, m$. *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 [8], 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* $\boldsymbol{w}$ *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* WVaR (6) *and* WPVaR (18).

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* WVaR *and* WPVaR *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,* WVaR *vastly overestimates* WPVaR. *This effect is amplified for lower values of* ϵ*, where the accuracy of the VaR estimate matters most. Indeed, for* $\epsilon = 1\%$*, the* WVaR *reports an unrealistically high value of* 497%*, which is 7 times larger than the corresponding* WPVaR *value.*

5 Worst-Case Quadratic VaR Optimization

The WPVaR 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 [29]), 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 WPVaR model. It is important to note that WPVaR 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

$$\bar{\theta}_i = \partial_t v_i(\boldsymbol{s}(0), 0) \in \mathbb{R}, \quad \bar{\boldsymbol{\Delta}}_i = \nabla_{\boldsymbol{s}} v_i(\boldsymbol{s}(0), 0) \in \mathbb{R}^n, \quad \text{and} \quad \bar{\boldsymbol{\Gamma}}_i = \nabla^2_{\boldsymbol{s}} v_i(\boldsymbol{s}(0), 0) \in \mathbb{S}^n. \tag{22}$$

The values computed in (22) 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*

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

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

$$\tilde{r}_i \;\approx\; f_i(\tilde{\boldsymbol{\xi}}) = \theta_i + \boldsymbol{\Delta}_i^T\tilde{\boldsymbol{\xi}} + \frac{1}{2}\tilde{\boldsymbol{\xi}}^T\boldsymbol{\Gamma}_i\tilde{\boldsymbol{\xi}} \quad \forall i = 1,\dots,m. \tag{23}$$

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

$$\boldsymbol{w}^T\tilde{\boldsymbol{r}} \;\approx \theta(\boldsymbol{w}) + \boldsymbol{\Delta}(\boldsymbol{w})^T\tilde{\boldsymbol{\xi}} + \frac{1}{2}\tilde{\boldsymbol{\xi}}^T\boldsymbol{\Gamma}(\boldsymbol{w})\tilde{\boldsymbol{\xi}}, \tag{24}$$

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.

5.2 Worst-Case Quadratic VaR Model

We define the *Worst-Case Quadratic VaR* (WQVaR) of a fixed portfolio $\boldsymbol{w} \in \mathcal{W}$ in terms of the Taylor expansion (24).

$$\begin{aligned}\mathrm{WQVaR}_\epsilon(\boldsymbol{w}) \;&=\; \min\left\{\gamma \;:\; \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\left\{\gamma \le -\boldsymbol{w}^T f(\tilde{\boldsymbol{\xi}})\right\} \le \epsilon\right\} \\ &=\; \min\left\{\gamma \;:\; \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\left\{\gamma \le -\theta(\boldsymbol{w}) - \boldsymbol{\Delta}(\boldsymbol{w})^T\tilde{\boldsymbol{\xi}} - \frac{1}{2}\tilde{\boldsymbol{\xi}}^T\boldsymbol{\Gamma}(\boldsymbol{w})\tilde{\boldsymbol{\xi}}\right\} \le \epsilon\right\}\end{aligned} \tag{25}$$

Note that the WQVaR 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 WQVaR of a portfolio $\boldsymbol{w}$ can be computed by solving a tractable SDP.

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

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

where $\boldsymbol{\Omega}$ *is the second-order moment matrix of* $\tilde{\boldsymbol{\xi}}$*; see* (17)*.*

Proof: See Appendix B.

Remark 5.1 *The* WQVaR *model described here 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* WQVaR *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, all derivations presented here remain valid in this generalized setting.*

As in the case of WPVaR*, it is possible to extend the* WQVaR *model to account for partitioned statistics information and for box-type uncertainty in the moment matrix* $\boldsymbol{\Omega}$ *by extending techniques proposed in [20] in a straightforward manner; see also Remarks 4.3 and 4.4.*

Remark 5.2 *If the market only contains basic assets and no derivatives, then* $m = n$*, and the coefficient functions in the delta-gamma approximation* (24) *reduce to* $\theta(\boldsymbol{w}) = 0$*,* $\boldsymbol{\Delta}(\boldsymbol{w}) = \boldsymbol{w}$*, and* $\boldsymbol{\Gamma}(\boldsymbol{w}) = \mathbf{0}$*. In this special case, the* WQVaR *is computed by solving the following SDP.*

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

El Ghaoui et al. *[11] have shown (using similar arguments as in Theorem 4.2) that this SDP has the closed form solution*

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

Thus, the WQVaR *model is a direct extension of the* WVaR *model* (6)*.*

Problem (26) constitutes a convex SDP that facilitates the efficient computation of the WQVaR for any fixed portfolio $\boldsymbol{w} \in \mathcal{W}$. Since the matrix inequality in (26) 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 WQVaR over all portfolios $\boldsymbol{w} \in \mathcal{W}$ by solving the following SDP.

$$
\begin{aligned}
\inf \quad & \gamma \\
\text{s.t.} \quad & \mathbf{M} \in \mathbb{S}^{n+1}, \quad \tau \in \mathbb{R}, \quad \gamma \in \mathbb{R}, \quad \boldsymbol{w} \in \mathcal{W} \\
& \langle \boldsymbol{\Omega}, \mathbf{M} \rangle \leq \tau\epsilon, \quad \mathbf{M} \succcurlyeq \mathbf{0}, \quad \tau \geq 0 \\
& \mathbf{M} + \begin{bmatrix} \boldsymbol{\Gamma}(\boldsymbol{w}) & \boldsymbol{\Delta}(\boldsymbol{w}) \\ \boldsymbol{\Delta}(\boldsymbol{w})^T & -\tau + 2(\gamma + \theta(\boldsymbol{w})) \end{bmatrix} \succcurlyeq \mathbf{0}
\end{aligned}
\tag{27}
$$

Remark 5.3 *Unlike in Section 4, there seems to be no equivalent SOCP formulation for the SDP* (27)*. In particular, there is no simple way to adapt the arguments in the proof of Theorem 4.2 to the current setting. The reason for this is a fundamental difference between the corresponding SDP problems* (16) *and* (27)*. In fact, the top left principal submatrix in the last LMI constraint is independent of* $\boldsymbol{w}$ *in* (16) *but not in* (27)*.*

5.3 Robust Optimization Perspective on WQVaR

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

Theorem 5.2 *The* WQVaR *minimization problem* (27) *is equivalent to the robust optimization problem*

$$
\begin{aligned}
\underset{\boldsymbol{w} \in \mathbb{R}^m, \gamma \in \mathbb{R}}{\text{minimize}} \quad & \gamma \\
\text{subject to} \quad & -\langle \mathbf{Q}(\boldsymbol{w}), \mathbf{Z} \rangle \leq \gamma \qquad \forall \mathbf{Z} \in \mathcal{U}_\epsilon^q \\
& \boldsymbol{w} \in \mathcal{W},
\end{aligned}
\tag{28}
$$

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*

$$
\mathcal{U}_\epsilon^q = \left\{ \mathbf{Z} = \begin{bmatrix} \mathbf{X} & \boldsymbol{\xi} \\ \boldsymbol{\xi}^T & 1 \end{bmatrix} \in \mathbb{S}^{n+1} \; : \; \boldsymbol{\Omega} - \epsilon\mathbf{Z} \succcurlyeq \mathbf{0}, \; \mathbf{Z} \succcurlyeq \mathbf{0} \right\}. \tag{29}
$$

Proof: See Appendix B.

It may not be evident how the uncertainty set $\mathcal{U}_\epsilon^q$ (defined in (29)) associated with the WQVaR 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 admissible portfolios, then the robust optimization problem* (28) *reduces to*

$$\begin{aligned} \underset{\boldsymbol{w}\in\mathbb{R}^m,\gamma\in\mathbb{R}}{\text{minimize}} \quad & \gamma \\ \text{subject to} \quad & -\theta(\boldsymbol{w}) - \boldsymbol{\Delta}(\boldsymbol{w})^T\boldsymbol{\xi} - \frac{1}{2}\boldsymbol{\xi}^T\boldsymbol{\Gamma}(\boldsymbol{w})\boldsymbol{\xi} \leq \gamma \qquad \forall \boldsymbol{\xi} \in \mathcal{U}_\epsilon \\ & \boldsymbol{w} \in \mathcal{W}, \end{aligned} \tag{30}$$

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

Proof: See Appendix B.

Remark 5.4 *Note that the robust optimization problem* (30) *can be reformulated as*

$$\begin{aligned} \underset{\boldsymbol{w}\in\mathbb{R}^m,\gamma\in\mathbb{R}}{\text{minimize}} \quad & \gamma \\ \text{subject to} \quad & -\boldsymbol{w}^T\boldsymbol{r} \leq \gamma \qquad \forall \boldsymbol{r} \in \mathcal{U}_\epsilon^{q2} \\ & \boldsymbol{w} \in \mathcal{W}, \end{aligned} \tag{31}$$

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

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

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 WQVaR *model may provide a tighter approximation of the actual VaR of a portfolio containing derivatives than the* WVaR *model.*

It seems that a min-max formulation (31) 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 (28), 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 after 2 days instead of 21 days. Since the VaR is no longer evaluated at the maturity time of the*

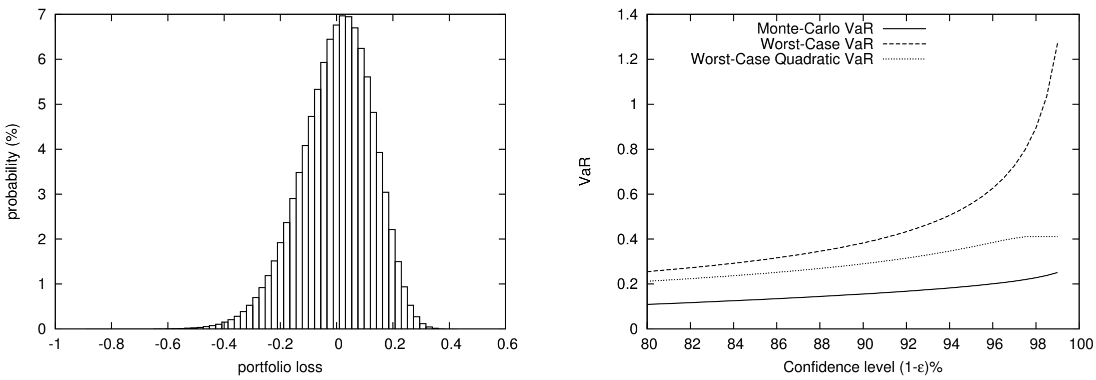

Figure 2: Left: The portfolio loss distribution obtained via Monte-Carlo simulation. Note that negative values represent gains. Right: The VaR estimates at different confidence levels obtained via Monte-Carlo sampling, WVaR, and WQVaR.

options, we use the WQVaR *model instead of the* WPVaR *model. The coefficients of the quadratic approximation function* (24) *are calculated using the standard Black-Scholes greek formulas (see, e.g., [15]). 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*,* WVaR*, and* WQVaR *for different confidence levels. Even for the short horizon time under consideration, the* WVaR *model still fails to give a realistic VaR estimate. At* $\epsilon = 1\%$*,* WVaR *is more than 3 times as large as the corresponding* WQVaR *value. This example demonstrates that* WQVaR *can offer significantly better VaR estimates than* WVaR *when the portfolio contains options.*

6 Relation to Worst-Case Conditional Value-at-Risk

We now establish a connection between the VaR-based risk measures proposed in Sections 4 and 5 and two Conditional Value-at-Risk (CVaR) measures, which are coherent in the sense of Artzner *et al.* [3]. This equivalence will then allow us to conclude that WPVaR and WQVaR are coherent risk measures over spaces of restricted portfolio returns. Coherence of risk measures from a robust optimization perspective has also been investigated by Natarajan *et al.* [19].

The classical CVaR is a quantile-based risk measure which evaluates the conditional expectation of the portfolio loss above VaR. For a given probability distribution $\mathbb{P}$ of $\tilde{\boldsymbol{r}}$ and tolerance $\epsilon \in (0, 1)$, Rockafellar and Uryasev [22] define the CVaR of a portfolio $\boldsymbol{w} \in \mathbb{R}^m$ as

$$\mathrm{CVaR}_\epsilon(\boldsymbol{w}) = \inf_{\beta \in \mathbb{R}} \left\{ \beta + \frac{1}{\epsilon} \mathbb{E}_{\mathbb{P}} \left((-\boldsymbol{w}^T \tilde{\boldsymbol{r}} - \beta)^+ \right) \right\},$$

where $\mathbb{E}_{\mathbb{P}}(\cdot)$ denotes expectation with respect to $\mathbb{P}$. By construction, CVaR is convex in $\boldsymbol{w}$ and—unlike VaR—constitutes a coherent risk measure [1]. If only the mean vector $\boldsymbol{\mu}_{\boldsymbol{r}}$ and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{r}} \succ \boldsymbol{0}$ of $\tilde{\boldsymbol{r}}$ are known, CVaR does not admit exact evaluation. As in the case of VaR, it then proves useful to introduce the *Worst-Case CVaR* (WCVaR)

$$\text{WCVaR}_\epsilon(\boldsymbol{w}) = \sup_{\mathbb{P}\in\mathcal{P}_{\boldsymbol{r}}} \inf_{\beta\in\mathbb{R}} \left\{ \beta + \frac{1}{\epsilon}\mathbb{E}_{\mathbb{P}}(-\boldsymbol{w}^T\tilde{\boldsymbol{r}} - \beta)^+ \right\},$$

where $\mathcal{P}_{\boldsymbol{r}}$ denotes the set of probability distributions of $\tilde{\boldsymbol{r}}$ consistent with the given mean vector $\boldsymbol{\mu}_{\boldsymbol{r}}$ and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{r}}$.

If a portfolio includes long positions in derivatives expiring at the end of the investment horizon as in Section 4, WCVaR may be an overly conservative risk measure. In analogy to WPVaR we thus introduce the Worst-Case Polyhedral CVaR (WPCVaR),

$$\text{WPCVaR}_\epsilon(\boldsymbol{w}) = \sup_{\mathbb{P}\in\mathcal{P}} \inf_{\beta\in\mathbb{R}} \left\{ \beta + \frac{1}{\epsilon}\mathbb{E}_{\mathbb{P}} \left(-(\boldsymbol{w}^{\boldsymbol{\xi}})^T\tilde{\boldsymbol{\xi}} - (\boldsymbol{w}^{\boldsymbol{\eta}})^T \max\{-\boldsymbol{e}, \boldsymbol{a} + \mathbf{B}\tilde{\boldsymbol{\xi}} - \boldsymbol{e}\} - \beta \right)^+ \right\},$$

which faithfully accounts for the known dependencies among the derivatives and their underlying assets. Here, $\mathcal{P}$ denotes the usual set of all distributions of the derivative underliers which share the same mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Similarly, for portfolios including derivatives that expire far beyond the investment horizon as in Section 5, we can define the Worst-Case Quadratic CVaR (WQCVaR) as follows.

$$\text{WQCVaR}_\epsilon(\boldsymbol{w}) = \sup_{\mathbb{P}\in\mathcal{P}} \inf_{\beta\in\mathbb{R}} \left\{ \beta + \frac{1}{\epsilon}\mathbb{E}_{\mathbb{P}} \left(-\theta(\boldsymbol{w}) - \boldsymbol{\Delta}(\boldsymbol{w})^T\tilde{\boldsymbol{\xi}} - \frac{1}{2}\tilde{\boldsymbol{\xi}}^T\boldsymbol{\Gamma}(\boldsymbol{w})\tilde{\boldsymbol{\xi}} - \beta \right)^+ \right\}$$

The next theorem establishes a connection between these new CVaR-based risk measures and the WPVaR and WQVaR measures introduced in Sections 4 and 5, respectively.

Theorem 6.1 *The following identities hold:*

(i) $\text{WPVaR}_\epsilon(\boldsymbol{w}) = \text{WPCVaR}_\epsilon(\boldsymbol{w})$ *for all* $\boldsymbol{w} = (\boldsymbol{w}^{\boldsymbol{\xi}}, \boldsymbol{w}^{\boldsymbol{\eta}}) \in \mathbb{R}^n \times \mathbb{R}^{m-n}_+$*;*

(ii) $\text{WQVaR}_\epsilon(\boldsymbol{w}) = \text{WQCVaR}_\epsilon(\boldsymbol{w})$ *for all* $\boldsymbol{w} \in \mathbb{R}^m$*.*

Proof: The claim is an immediate consequence of an exactness result about Worst-Case CVaR approximations for distributionally robust chance constraints with convex or (possibly non-convex) quadratic constraint functions [28, Theorem 2.2]. ■

Corollary 6.1 WPVaR *is a coherent risk measure on the cone of polyhedral portfolio returns*

$$R_{\text{P}} = \left\{ \tilde{r}_p = (\boldsymbol{w}^{\boldsymbol{\xi}})^T\tilde{\boldsymbol{\xi}} + (\boldsymbol{w}^{\boldsymbol{\eta}})^T \max\{-\boldsymbol{e}, \boldsymbol{a} + \mathbf{B}\tilde{\boldsymbol{\xi}} - \boldsymbol{e}\} \;:\; \boldsymbol{w}^{\boldsymbol{\xi}} \in \mathbb{R}^n,\ \boldsymbol{w}^{\boldsymbol{\eta}} \in \mathbb{R}^{m-n}_+ \right\}.$$

Similarly, WQVaR *is a coherent risk measure on the subspace of quadratic portfolio returns*

$$R_{\text{Q}} = \left\{ \tilde{r}_p = \theta(\boldsymbol{w}) + \boldsymbol{\Delta}(\boldsymbol{w})^T\tilde{\boldsymbol{\xi}} + \frac{1}{2}\tilde{\boldsymbol{\xi}}^T\boldsymbol{\Gamma}(\boldsymbol{w})\tilde{\boldsymbol{\xi}} \;:\; \boldsymbol{w} \in \mathbb{R}^m \right\}.$$

Proof: The known coherence of the classical CVaR for any given distribution $\mathbb{P}$, see e.g. [1], implies via [27, Proposition 1] that WPCVaR is coherent on R_{P}. By the equivalence of WPVaR and WPCVaR established in Theorem 6.1, WPVaR is thus a coherent risk measure on R_P. Coherence of WQVaR on R_Q is proved in a similar manner. ∎

Recall that the classical WVaR by El Ghaoui *et al.* [11] represents a special case of WPVaR (or WQVaR). Thus, an important consequence of Corollary 6.1 is that WVaR is a coherent risk measure on the space of all portfolio returns of the form $\boldsymbol{w}^T\tilde{\boldsymbol{r}}$ for $\boldsymbol{w} \in \mathbb{R}^m$.

Example 6.1 (Index Tracking) *We compare the out-of-sample performance of* VaR*,* CVaR*,* WVaR *and* WQVaR *in an index tracking experiment with real data taken from the Optionmetrics IvyDB database. The aim is to replicate the S&P 500 (SPX) index with a portfolio consisting of the NASDAQ-100 (NDX) and Russell 2000 (RUT) indices as well as two at-the-money European call options on NASDAQ-100 and Russell 2000, respectively. The experiment covers 265 biweekly rebalancing intervals from February 6 2001 to March 30 2011. In each rebalancing interval and for both underliers we select the option with minimum bid-ask spread among all available options expiring within the next two to three months. This choice of maturity dates ensures that the option returns admit a delta-gamma approximation. The high liquidity of the index options allows us to assume that they can be purchased or sold anytime during the backtest period. The historical bid-ask spreads average at 3.8% and 3.2% for the chosen options on NASDAQ-100 and Russell 2000, respectively. In order to ensure a fair reflection of liquidity risk, we artificially inflate the bid-ask spreads of the options to 5%. The bid-ask spreads of the option underliers are assumed to vanish, while the confidence level of the risk measures is set to* $\epsilon = 5\%$.

At the start of each rebalancing interval we minimize the risk of the tracking portfolio's excess return over the benchmark. We use a rolling estimation window of three years (178 data points) to calibrate the arising optimization models. VaR *and* CVaR *are evaluated under the empirical (discrete) distribution implied by the return observations in the estimation window. Thus, the underlying optimization problems can be reformulated as mixed-integer linear programs and linear programs, respectively, see e.g. [12, 22]. Similarly,* WVaR *and* WQVaR *are evaluated using the sample means and covariances implied by the data in the rolling estimation window. The underlying optimization models can be reformulated as SOCPs and SDPs, respectively, see Section 2.1 and Theorem 5.1. The* WQVaR *model additionally requires information about the options' delta, gamma and theta sensitivities. These are obtained from the implied volatilities reported in the Optionmetrics database and are calculated using the Black-Scholes formula.*

On average, the arising linear and mixed-integer linear programs are solved in 16ms and 206ms using ILOG CPLEX, while the SOCPs and SDPs are solved in 91ms and 195ms with SDPT3 [24], respectively. In general, the optimal CVaR*,* WVaR *and* WQVaR*-portfolios can be computed in time polynomial in the number of stocks and options via efficient interior point algorithms [26]. For example, we have found that instances of the* WQVaR *optimization model involving 180 stocks and 180 options can conveniently be solved in less than one hour. In contrast, the* VaR *optimization problems involve binary variables and permit no polynomial time solution. Figure 3 visualizes the cumulative returns of the four tracking portfolios relative to the cumulative return of S&P 500. In this example the* WQVaR*-portfolio exceeds the benchmark over 80.4% of the time. In contrast, the* WVaR*-portfolio exceeds the benchmark only 7.8% of the time, while*

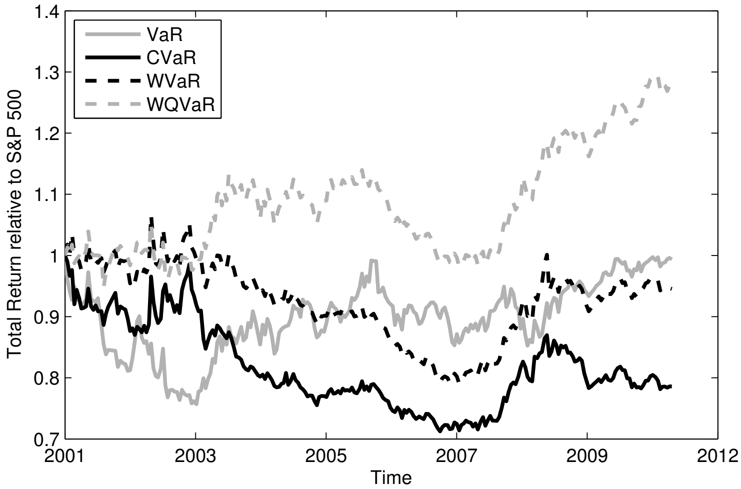

Figure 3: Out-of-sample performance of different tracking portfolios

Table 1: Left: Out-of-sample statistics of excess return over the benchmark for different tracking portfolios (in basis points). Right: Quadratic variation of the hedging instruments' portfolio weights in the different tracking portfolios.

Statistic	VaR	CVaR	WVaR	WQVaR
Mean	1.3	-10.5	-4.1	10.0
Std Dev	183.5	150.6	143.8	145.4
VaR	313.8	236.3	240.2	208.6
CVaR	412.0	365.2	342.7	311.3

Instrument	VaR	CVaR	WVaR	WQVaR
NDX	72.733	7.252	4.711	3.730
RUT	68.073	7.326	4.864	3.738
Call on NDX	0.586	0.108	0.062	0.069
Call on RUT	0.257	0.121	0.090	0.080

the portfolios based on CVaR *and* VaR *fall consistently short of the benchmark. Table 1 reports out-of-sample statistics of the portfolios' excess returns over S&P 500. We observe that the excess return of the* WQVaR*-portfolio has the highest out-of-sample mean and the lowest out-of-sample* VaR *and* CVaR *(consistently computed at $\epsilon = 5\%$) among all tracking portfolios. We remark that any discrepancies in the performance of the* WVaR *and* WQVaR*-portfolios can only originate from the different treatment of the options in these models. The non-robust tracking portfolios based on* VaR *and* CVaR *tend to be overfitted to the empirical return distributions and are therefore prone to error maximisation phenomena. Moreover, the corresponding portfolio weights are less stable than those of the* WVaR *and* WQVaR*-portfolios. Table 1 reports the quadratic variations, that is, the summed squared differences of all successive portfolio weights throughout the backtest period. The quadratic variation of a stochastic process can be regarded as a measure of its smoothness. We observe that the quadratic variations corresponding to the portfolio weights in the* WQVaR*-portfolio are almost two orders of magnitude smaller than those of the* VaR*-portfolio, and they are also significantly smaller than those of the* CVaR *and* WVaR*-portfolios. We therefore conclude that the* WQVaR*-portfolio is least affected by data noise.*

7 Conclusions

Derivatives depend non-linearly on their underlying assets. In this paper we explicitly incorporate this non-linear relationship into the WVaR model, which results in two new models. The

WPVaR model expresses the option returns as convex-piecewise linear functions of the underlying assets and is therefore suited for portfolios containing European options maturing at the investment horizon. 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 WPVaR model precludes short-sales of options. The WQVaR model, on the other hand, can handle portfolios containing non-European options and does not rely on short-sales restrictions. It exploits the popular delta-gamma approximation to model the portfolio return. In contrast to WPVaR, however, WQVaR does require knowledge of suitable option pricing models to determine the delta-gamma approximation. Through numerical experiments we demonstrate that the WPVaR and WQVaR models can provide much tighter VaR estimates of an options portfolio than the WVaR model, which disregard any non-linear dependencies between the asset returns. We also demonstrate that WQVaR, WPVaR and WVaR are equivalent to the corresponding worst-case CVaR risk measures, which implies that they are actually coherent—in marked contrast to the classical non-robust VaR.

Acknowledgments. We are indebted to Prof. A. Ben-Tal for valuable discussions on the topics of this paper and to Grani Hanasusanto for his assistance with the numerical experiments. This research has been supported by EPSRC under grant EP/I014640/1.

References

[1] C. Acerbi and D. Tasche. On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7):1487–1503, 2002.

[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] D. Bertsimas and M. Sim. The price of robustness. *Operations Research*, 52(1):35–53, 2004.

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

[9] 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.

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

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

[12] A. Gaivoronski and G. Pflug. Value-at-Risk in portfolio optimization: Properties and computational approach. *Journal of Risk*, 7(2):1–31, 2005.

[13] S. 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] L. MacMillan. *Options as a strategic investment*. Prentice Hall, 1992.

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

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

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

[19] K. Natarajan, D. Pachamanova, and M. Sim. Constructing risk measures from uncertainty sets. *Operations Research*, 57(5):1129–1141, 2009.

[20] K. Natarajan, M. Sim, and J. Uichanco. Tractable robust expected utility and risk models for portfolio optimisation. *Mathematical Finance*, 20(4):695–731, 2010.

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

[22] R. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–41, 2002.

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

[24] 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.

[25] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, 1996.

[26] Yinyu Ye. *Interior point algorithms: theory and analysis*. John Wiley & Sons, Inc., New York, NY, USA, 1997.

[27] S. Zhu and M. Fukushima. Worst-case conditional value-at-risk with application to robust portfolio management. *Operations Research*, 57(5):1155–1168, 2009.

[28] S. Zymler, D. Kuhn, and B. Rustem. Distributionally robust joint chance constraints with second-order moment information. *Mathematical Programming*, 2011. Forthcoming.

[29] S. Zymler, B. Rustem, and D. Kuhn. Robust portfolio optimization with derivative insurance guarantees. *European Journal of Operational Research*, 210(2):410–424, 2011.

A Worst-Case Probability Problems

In this appendix we review a general result about worst-case probability problems that plays a key role for many of the derivations in this paper.

Lemma A.1 (Calafiore *et al.* [9]) *Let $\mathcal{S} \subseteq \mathbb{R}^n$ be any Borel measurable set (which is not necessarily convex), and define the worst-case probability $\pi_{\rm wc}$ as*

$$\pi_{\rm wc} = \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\{\tilde{\boldsymbol{\xi}} \in \mathcal{S}\}, \tag{32}$$

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,

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

where $\boldsymbol{\Omega}$ is the second-order moment matrix of $\tilde{\boldsymbol{\xi}}$; see (17)*.*

B Proofs

Proof of Theorem 4.1 In order to derive a manifestly tractable representation for WPVaR, we first simplify the maximization problem

$$\sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\left\{ \gamma \leq -(\boldsymbol{w}^{\boldsymbol{\xi}})^T \tilde{\boldsymbol{\xi}} - (\boldsymbol{w}^{\boldsymbol{\eta}})^T \max\left\{ -\boldsymbol{e}, \boldsymbol{a} + \mathbf{B}\tilde{\boldsymbol{\xi}} - \boldsymbol{e} \right\} \right\}, \tag{33}$$

which can be identified as the subordinate optimization problem in (15).

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

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

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

$$\mathcal{S}_\gamma = \left\{ \boldsymbol{\xi} \in \mathbb{R}^n \ : \ \max_{\mathbf{0} \leq \boldsymbol{y} \leq \boldsymbol{w}^{\boldsymbol{\eta}}} \left\{ \gamma + (\boldsymbol{w}^{\boldsymbol{\xi}})^T \boldsymbol{\xi} + \boldsymbol{y}^T (\boldsymbol{a} + \mathbf{B}\boldsymbol{\xi}) - \boldsymbol{e}^T \boldsymbol{w}^{\boldsymbol{\eta}} \right\} \leq 0 \right\}. \tag{34}$$

The optimal value of problem (33) is obtained by solving the worst-case probability problem $\sup_{\mathbb{P} \in \mathcal{P}} \mathbb{P}\{\tilde{\boldsymbol{\xi}} \in \mathcal{S}_\gamma\}$, which, by Lemma A.1 can be expressed as

$$\begin{array}{ll} \inf\limits_{\mathbf{M} \in \mathbb{S}^{n+1}} & \langle \boldsymbol{\Omega}, \mathbf{M} \rangle \\ \text{s.t.} & \left[\boldsymbol{\xi}^T \ 1\right] \mathbf{M} \left[\boldsymbol{\xi}^T \ 1\right]^T \geq 1 \ \forall \boldsymbol{\xi} : \max\limits_{\mathbf{0} \leq \boldsymbol{y} \leq \boldsymbol{w}^{\boldsymbol{\eta}}} \{\gamma + (\boldsymbol{w}^{\boldsymbol{\xi}})^T \boldsymbol{\xi} + \boldsymbol{y}^T (\boldsymbol{a} + \mathbf{B}\boldsymbol{\xi}) - \boldsymbol{e}^T \boldsymbol{w}^{\boldsymbol{\eta}}\} \leq 0 \\ & \mathbf{M} \succcurlyeq \mathbf{0}. \end{array} \tag{35}$$

We will now argue that (35) is equivalent to problem (36) below for all but one value of γ.

$$\begin{array}{ll} \inf & \langle \boldsymbol{\Omega}, \mathbf{M} \rangle \\ \text{s.t.} & \mathbf{M} \in \mathbb{S}^{n+1}, \quad \tau \in \mathbb{R}, \quad \mathbf{M} \succcurlyeq \mathbf{0}, \quad \tau \geq 0 \\ & \left[\boldsymbol{\xi}^T \ 1\right] \mathbf{M} \left[\boldsymbol{\xi}^T \ 1\right]^T - 1 + 2\tau \left(\max\limits_{\mathbf{0} \leq \boldsymbol{y} \leq \boldsymbol{w}^{\boldsymbol{\eta}}} \{\gamma + (\boldsymbol{w}^{\boldsymbol{\xi}})^T \boldsymbol{\xi} + \boldsymbol{y}^T (\boldsymbol{a} + \mathbf{B}\boldsymbol{\xi}) - \boldsymbol{e}^T \boldsymbol{w}^{\boldsymbol{\eta}}\} \right) \geq 0 \ \forall \boldsymbol{\xi} \in \mathbb{R}^n \end{array} \tag{36}$$

For ease of exposition, we 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 (35) and (36) is proved case by case. Assume first that $h < 0$. Then, the equivalence follows from the non-linear Farkas Lemma, see, e.g., [21, Theorem 2.1]. Assume next that $h > 0$. Then, the semi-infinite constraint in (35) becomes redundant and, since $\boldsymbol{\Omega} \succ \mathbf{0}$, the optimal solution of (35) is given by $\mathbf{M} = \mathbf{0}$ with a corresponding optimal value of 0. The optimal value of problem (36) is also equal to 0. Indeed, by choosing $\tau = 1/h$, the semi-infinite constraint in (36) is satisfied independently of $\mathbf{M}$. Finally, (35) and (36) may differ for $h = 0$.

It can be seen that since $\tau \geq 0$, the semi-infinite constraint in (36) is equivalent to the assertion that there exists some $\mathbf{0} \leq \boldsymbol{y} \leq \boldsymbol{w}^{\boldsymbol{\eta}}$ with

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

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, problem (36) can equivalently be formulated as

$$
\begin{array}{ll}
\inf & \langle \mathbf{\Omega}, \mathbf{M} \rangle \\
\text{s.t.} & \mathbf{M} \in \mathbb{S}^{n+1}, \quad \boldsymbol{y} \in \mathbb{R}^{m-n}, \quad \tau \in \mathbb{R} \\
& \mathbf{M} \succcurlyeq \mathbf{0}, \quad \tau \geq 0, \quad \mathbf{0} \leq \boldsymbol{y} \leq \boldsymbol{w}^{\boldsymbol{\eta}} \\
& \mathbf{M} + \begin{bmatrix} \mathbf{0} & \tau(\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T \boldsymbol{y}) \\ \tau(\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T \boldsymbol{y})^T & -1 + 2\tau(\gamma + \boldsymbol{y}^T \boldsymbol{a} - \boldsymbol{e}^T \boldsymbol{w}^{\boldsymbol{\eta}}) \end{bmatrix} \succcurlyeq \mathbf{0}.
\end{array}
\tag{37}
$$

Since (35) and (37) are equivalent for all but one value of γ and since their optimal values are nonincreasing in γ, we can express WPVaR in (15) as the optimal value of the following problem.

$$
\begin{array}{lll}
\text{WPVaR}_\epsilon(\boldsymbol{w}) = & \inf & \gamma \\
& \text{s.t.} & \mathbf{M} \in \mathbb{S}^{n+1}, \quad \boldsymbol{y} \in \mathbb{R}^{m-n}, \quad \tau \in \mathbb{R}, \quad \gamma \in \mathbb{R} \\
& & \langle \mathbf{\Omega}, \mathbf{M} \rangle \leq \epsilon, \quad \mathbf{M} \succcurlyeq \mathbf{0}, \quad \tau \geq 0, \quad \mathbf{0} \leq \boldsymbol{y} \leq \boldsymbol{w}^{\boldsymbol{\eta}} \\
& & \mathbf{M} + \begin{bmatrix} \mathbf{0} & \tau(\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T \boldsymbol{y}) \\ \tau(\boldsymbol{w}^{\boldsymbol{\xi}} + \mathbf{B}^T \boldsymbol{y})^T & -1 + 2\tau(\gamma + \boldsymbol{y}^T \boldsymbol{a} - \boldsymbol{e}^T \boldsymbol{w}^{\boldsymbol{\eta}}) \end{bmatrix} \succcurlyeq \mathbf{0}
\end{array}
\tag{38}
$$

Problem (38) is non-convex due to the bilinear terms in the matrix inequality constraint. It can easily be shown that $\langle \mathbf{\Omega}, \mathbf{M} \rangle \geq 1$ 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 (38) by τ. Subsequently we perform variable substitutions in which we replace τ by $1/\tau$ and $\mathbf{M}$ by $\mathbf{M}/\tau$. This shows that (38) is equivalent to problem (16). ∎

Proof of Theorem 5.1 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

$$
\mathcal{Q}_\gamma = \left\{ \boldsymbol{\xi} \in \mathbb{R}^n \; : \; \gamma \leq -\theta(\boldsymbol{w}) - \mathbf{\Delta}(\boldsymbol{w})^T \boldsymbol{\xi} - \frac{1}{2} \boldsymbol{\xi}^T \mathbf{\Gamma}(\boldsymbol{w}) \boldsymbol{\xi} \right\}. \tag{39}
$$

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

$$
\pi_{\text{wc}} = \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{P}\{\tilde{\boldsymbol{\xi}} \in \mathcal{Q}_\gamma\}, \tag{40}
$$

which can be identified as the subordinate maximization problem in (25). Lemma A.1 implies that (40) can equivalently be formulated as

$$
\pi_{\text{wc}} = \inf_{\mathbf{M} \in \mathbb{S}^{n+1}} \left\{ \langle \mathbf{\Gamma}, \mathbf{M} \rangle \; : \; \mathbf{M} \succcurlyeq \mathbf{0}, \quad \begin{bmatrix} \boldsymbol{\xi}^T & 1 \end{bmatrix} \mathbf{M} \begin{bmatrix} \boldsymbol{\xi}^T & 1 \end{bmatrix}^T \geq 1 \quad \forall \boldsymbol{\xi} \in \mathcal{Q}_\gamma \right\}. \tag{41}
$$

By the definition of $\mathcal{Q}$, the semi-infinite constraint in problem (41) is equivalent to

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

By using the $\mathcal{S}$-lemma [21] and by analogous reasoning as in Section 4.2, we can replace the semi-infinite constraint in problem (41) by

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

without changing the optimal value of the problem. Thus, the worst-case probability problem (40) can be rewritten as

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

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

$$\begin{aligned} \text{WQVaR}_\epsilon(\boldsymbol{w}) = \inf \quad & \gamma \\ \text{s.t.} \quad & \mathbf{M} \in \mathbb{S}^{n+1}, \quad \tau \in \mathbb{R}, \quad \gamma \in \mathbb{R} \\ & \langle \boldsymbol{\Omega}, \mathbf{M} \rangle \leq \epsilon, \quad \mathbf{M} \succcurlyeq \mathbf{0}, \quad \tau \geq 0 \\ & \mathbf{M} + \begin{bmatrix} \tau \boldsymbol{\Gamma}(\boldsymbol{w}) & \tau \boldsymbol{\Delta}(\boldsymbol{w}) \\ \tau \boldsymbol{\Delta}(\boldsymbol{w})^T & -1 + 2\tau(\gamma + \theta(\boldsymbol{w})) \end{bmatrix} \succcurlyeq \mathbf{0} \end{aligned} \tag{43}$$

By analogous reasoning as in Section 4.2, it can be shown that any feasible solution of problem (43) has a strictly positive τ-component. Thus we may divide the matrix inequality in (43) by τ. After the variable transformation $\tau \to 1/\tau$ and $\mathbf{M} \to \mathbf{M}/\tau$, we obtain the postulated SDP (26). ∎

Proof of Theorem 5.2 For some fixed portfolio $\boldsymbol{w} \in \mathcal{W}$, the WQVaR can be computed by solving problem (26), which involves the LMI constraint

$$\mathbf{M} + \begin{bmatrix} \boldsymbol{\Gamma}(\boldsymbol{w}) & \boldsymbol{\Delta}(\boldsymbol{w}) \\ \boldsymbol{\Delta}(\boldsymbol{w})^T & -\tau + 2(\gamma + \theta(\boldsymbol{w})) \end{bmatrix} \succcurlyeq \mathbf{0}. \tag{44}$$

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 (44) is representable as

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

Thus, the WQVaR problem (26) can be rewritten as

$$
\begin{aligned}
\inf \quad & \sup_{\boldsymbol{\xi} \in \mathbb{R}^n} \ -\frac{1}{2}\boldsymbol{\xi}^T(\mathbf{V} + \boldsymbol{\Gamma}(\boldsymbol{w}))\boldsymbol{\xi} - \boldsymbol{\xi}^T(\boldsymbol{v} + \boldsymbol{\Delta}(\boldsymbol{w})) - \theta(\boldsymbol{w}) - \frac{1}{2}(u - \tau) \\
\text{s.t.} \quad & \mathbf{V} \in \mathbb{S}^n, \quad \boldsymbol{v} \in \mathbb{R}^n, \quad \tau \in \mathbb{R}, \quad u \in \mathbb{R} \\
& \begin{bmatrix} \mathbf{V} & \boldsymbol{v} \\ \boldsymbol{v}^T & u \end{bmatrix} \succcurlyeq \mathbf{0}, \quad \tau \geq 0, \quad \langle \mathbf{V}, \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T \rangle + 2\boldsymbol{v}^T\boldsymbol{\mu} + u \leq \tau\epsilon.
\end{aligned} \tag{45}
$$

Note that if $\mathbf{V} + \boldsymbol{\Gamma}(\boldsymbol{w})$ is not positive semidefinite, the inner maximization problem in (45) 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 (45) without changing its feasible region. With this constraint appended, the min-max problem (45) 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.

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

We proceed by dualizing the inner minimization problem in (46). After a few elementary simplification steps, this dual problem reduces to

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

Note that strong duality holds because the inner problem in (46) is strictly feasible for any $\epsilon > 0$, see [25]. This allows us to replace the inner minimization problem in (46) by the maximization problem (47), which yields the following equivalent formulation for the WQVaR problem (26).

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

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

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

By definition of $\mathbf{\Omega}$ as the second-order moment matrix of the basic asset returns, see (17), 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 (47) can be reformulated as

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

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

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

The WQVaR minimization problem (27) can therefore be expressed as the min-max problem

$$\min_{\boldsymbol{w} \in \mathcal{W}} \max_{\mathbf{Z} \in \mathcal{U}_\epsilon^q} -\langle \mathbf{Q}(\boldsymbol{w}), \mathbf{Z} \rangle, \tag{48}$$

which is manifestly equivalent to the postulated semi-infinite program (28). ■

Proof of Corollary 5.1 The inner maximization problem in (48) can be written as

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

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

$$\begin{array}{ll} \max & -\theta(\boldsymbol{w}) - \mathbf{\Delta}(\boldsymbol{w})^T \boldsymbol{\xi} - \frac{1}{2}\boldsymbol{\xi}^T \mathbf{\Gamma}(\boldsymbol{w}) \boldsymbol{\xi} - \frac{1}{2}\langle \mathbf{\Gamma}(\boldsymbol{w}), \mathbf{Y} \rangle \\ \text{s. t.} & \mathbf{Y} \in \mathbb{S}^n, \quad \boldsymbol{\xi} \in \mathbb{R}^n, \quad \mathbf{Y} \succcurlyeq \mathbf{0} \\ & \begin{bmatrix} (\mathbf{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T) - \epsilon (\mathbf{Y} + \boldsymbol{\xi}\boldsymbol{\xi}^T) & \boldsymbol{\mu} - \epsilon \boldsymbol{\xi} \\ (\boldsymbol{\mu} - \epsilon \boldsymbol{\xi})^T & 1 - \epsilon \end{bmatrix} \succcurlyeq \mathbf{0}. \end{array} \tag{49}$$

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 $\langle \mathbf{\Gamma}(\boldsymbol{w}), \mathbf{Y} \rangle \geq 0$ for all $\mathbf{Y} \succcurlyeq \mathbf{0}$ because $\mathbf{\Gamma}(\boldsymbol{w}) \succcurlyeq \mathbf{0}$ by assumption. Thus problem (49) reduces to

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

Using similar arguments as in Theorem 4.2, we can show that the semidefinite constraint in the above problem is equivalent to

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

Thus the original min-max formulation (48) 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. ∎