

# Worst-Case Value at Risk of Nonlinear Portfolios

Steve Zymler, Daniel Kuhn, Berç Rustem

Department of Computing, Imperial College of Science, Technology and Medicine,  
London SW7 2AZ, United Kingdom {sz02@doc.ic.ac.uk, dkuhn@imperial.ac.uk, br@doc.ic.ac.uk}

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 nonconvex) 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

*History:* Received July 4, 2011; accepted June 4, 2012, by Dimitris Bertsimas, optimization. Published online in *Articles in Advance* October 8, 2012.

## 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 (1952), 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 (1952) 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 return is skewed. This shortcoming triggered extensive research on downside risk measures. Because of its intuitive appeal and because its use is enforced by financial regulators, value at risk (VaR) remains the most popular downside risk measure (Jorion 2001). The VaR at level  $\epsilon$  is defined as the least  $(1 - \epsilon)$ -quantile of the portfolio loss distribution.

Despite its popularity, VaR lacks some desirable theoretical properties. First, VaR is known to be a nonconvex 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. Second, VaR fails to satisfy the subadditivity property of coherent risk measures (Artzner et al. 1999). Thus, the VaR of a portfolio can exceed the weighted sum of the VaRs of its constituents. In other words, VaR may penalize diversification. Third, 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. (2003) proposed 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 (SOCP). El Ghaoui et al. (2003) also disclosed an interesting connection to robust optimization (Ben-Tal and Nemirovski 1998, 1999; Rustem and Howe 2002): 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 overpessimistic 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 because of scarcity of time series data.
- WVaR disregards perfect dependencies between the derivative returns and the underlying asset returns. These (typically nonlinear) dependencies are known in practice because 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 nonlinear 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. (2008) included asymmetric distributional information into the WVaR optimization to obtain a tighter approximation of VaR. However, their model requires forward- and backward-deviation measures as inputs, 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 that explicitly account for perfect nonlinear 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 relationship 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 (Alizadeh and Goldfarb 2003). We also establish the equivalence of the WPVaR model to a robust optimization model described by Zymler et al. (2011b).

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 nonconvex) 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 (2002) and Mina and Ulmer (1999). 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 relationship 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 (SDP), which can be solved efficiently (Vandenberghe and Boyd 1996), 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. (2003) as well as the new WPVaR and WQVaR measures proposed here are in fact coherent risk measures in the sense of Artzner et al. (1999). This will be established by proving that WPVaR (WQVaR) coincides with the worst-case *conditional* value at risk (WCVaR) of the underlying polyhedral (quadratic) loss function.

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

- (1) We generalize the WVaR model (El Ghaoui et al. 2003) to explicitly account for the nonlinear 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 nonlinear 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 WVVaR 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. (1999).

### 1.1. Notation

We use lowercase boldface letters to denote vectors and uppercase boldface 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  $\langle \mathbf{X}, \mathbf{Y} \rangle = \text{Tr}(\mathbf{X}\mathbf{Y})$  be the trace scalar product, whereas 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, whereas 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  $\mathbf{w} \in \mathbb{R}^m$ , whose elements add up to 1. The component  $w_i$  denotes the percentage of total wealth that is invested in the  $i$ th asset at time  $t = 0$ . Furthermore,  $\tilde{\mathbf{r}}$  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  $\mathbf{w}$  over the investment period is thus given by the random variable  $\tilde{r}_p = \mathbf{w}^T \tilde{\mathbf{r}}$ . Loosely speaking, we aim at finding an allocation vector  $\mathbf{w}$  that entails a high portfolio

return while 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 *value at risk*. The VaR at level  $\epsilon$  is defined as the  $(1 - \epsilon)$ -percentile of the portfolio loss distribution, where  $\epsilon$  is typically chosen as 1% or 5%. Put differently,  $\text{VaR}_\epsilon(\mathbf{w})$  is defined as the smallest real number  $\gamma$  with the property that  $-\mathbf{w}^T \tilde{\mathbf{r}}$  exceeds  $\gamma$  with a probability not larger than  $\epsilon$ , that is,

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

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

In this paper we investigate portfolio optimization problems of the type

$$\begin{aligned} & \underset{\mathbf{w} \in \mathbb{R}^m}{\text{minimize}} && \text{VaR}_\epsilon(\mathbf{w}) \\ & \text{subject to} && \mathbf{w} \in \mathcal{W}, \end{aligned} \quad (2)$$

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

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

$$\begin{aligned} & \underset{\mathbf{w} \in \mathbb{R}^m, \gamma \in \mathbb{R}}{\text{minimize}} && \gamma \\ & \text{subject to} && \mathbb{P}\{\gamma + \mathbf{w}^T \tilde{\mathbf{r}} \geq 0\} \geq 1 - \epsilon, \\ & && \mathbf{w} \in \mathcal{W}, \end{aligned} \quad (3)$$

which constitutes a chance-constrained stochastic program. Optimization problems of this kind are usually difficult to solve because they tend to have nonconvex 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

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: *nonparametric approaches* that approximate the asset return distribution by a discrete (sampled or empirical) distribution and *parametric approaches* that 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  $\tilde{\mathbf{r}}$  are governed by a Gaussian distribution with given mean vector  $\boldsymbol{\mu}_r \in \mathbb{R}^m$  and covariance matrix  $\boldsymbol{\Sigma}_r \in \mathbb{S}^m$ . This assumption has the advantage that the VaR can be calculated analytically as

$$\text{VaR}_\epsilon(\mathbf{w}) = -\boldsymbol{\mu}_r^T \mathbf{w} - \Phi^{-1}(\epsilon) \sqrt{\mathbf{w}^T \boldsymbol{\Sigma}_r \mathbf{w}}, \quad (4)$$

where  $\Phi$  is the standard normal distribution function. This model is sometimes referred to as *Normal VaR* (see, e.g., Natarajan et al. 2008). 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. (2003) suggested 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  $\mathcal{P}_r$  be the set of all probability distributions on  $\mathbb{R}^m$  with mean value  $\boldsymbol{\mu}_r$  and covariance matrix  $\boldsymbol{\Sigma}_r$ . We emphasize that  $\mathcal{P}_r$  contains also distributions that exhibit considerable skewness, so long as they match the given mean vector and covariance matrix. The *worst-case value at risk* for portfolio  $\mathbf{w}$  is now defined as

$$\text{WVaR}_\epsilon(\mathbf{w}) = \min \left\{ \gamma : \sup_{\mathbb{P} \in \mathcal{P}_r} \mathbb{P}\{\gamma \leq -\mathbf{w}^T \tilde{\mathbf{r}}\} \leq \epsilon \right\}. \quad (5)$$

El Ghaoui et al. (2003) demonstrated that WVAr has the closed-form expression

$$\text{WVaR}_\epsilon(\mathbf{w}) = -\boldsymbol{\mu}_r^T \mathbf{w} + \kappa(\epsilon) \sqrt{\mathbf{w}^T \boldsymbol{\Sigma}_r \mathbf{w}}, \quad (6)$$

where  $\kappa(\epsilon) = \sqrt{(1-\epsilon)/\epsilon}$ . WVAr represents a tight approximation for VaR in the sense that there exists a worst-case distribution  $\mathbb{P}^* \in \mathcal{P}_r$  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

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

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.

$$\begin{aligned} & \underset{\mathbf{w} \in \mathbb{R}^m, \gamma \in \mathbb{R}}{\text{minimize}} \quad \gamma \\ & \text{subject to} \quad \gamma + \mathbf{w}^T \tilde{\mathbf{r}} \geq 0, \\ & \quad \mathbf{w} \in \mathcal{W}. \end{aligned} \quad (8)$$

Because the asset return vector is uncertain, this model essentially represents a whole family of optimization problems, one for each possible realization of  $\tilde{\mathbf{r}}$ . 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 (Ben-Tal and Nemirovski 1998, Ben-Tal et al. 2009, Bertsimas and Sim 2004) 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  $\tilde{\mathbf{r}}$  within a prescribed *uncertainty set*  $\mathcal{U}$ . This set may cover only a subset of all possible realizations of  $\tilde{\mathbf{r}}$  and is chosen by the modeller. The *robust counterpart* of problem (8) is then defined as

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

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  $\tilde{\mathbf{r}}$ . 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 (Ben-Tal and Nemirovski 1998, 1999).

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

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

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

$$\begin{aligned} & \gamma + \mathbf{w}^T \mathbf{r} \geq 0 \quad \forall \mathbf{r} \in \mathcal{U} \\ & \iff -\boldsymbol{\mu}_r^T \mathbf{w} + \delta \|\boldsymbol{\Sigma}_r^{1/2} \mathbf{w}\|_2 \leq \gamma. \end{aligned} \quad (10)$$

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

$$\begin{aligned} & \underset{\mathbf{w} \in \mathbb{R}^m}{\text{minimize}} \quad \{-\boldsymbol{\mu}_r^T \mathbf{w} + \delta \|\boldsymbol{\Sigma}_r^{1/2} \mathbf{w}\|_2\} \\ & \text{subject to} \quad \mathbf{w} \in \mathcal{W}. \end{aligned} \quad (11)$$

By comparing (7) and (11), El Ghaoui et al. (2003) noticed that optimizing WVaR at level  $\epsilon$  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. 2008). 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 nonlinear 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  $\tilde{\mathbf{r}} = (\tilde{\boldsymbol{\xi}}, \tilde{\boldsymbol{\eta}})$ , 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  $\mathbf{w} \in \mathcal{W}$  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., Coval and Shumway 2002). Note that the moments of the basic assets' returns (i.e., stocks and bonds, etc.) can usually be estimated more accurately because of 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 nonlinear) measurable function  $f: \mathbb{R}^n \rightarrow \mathbb{R}^{m-n}$  such that  $\tilde{\mathbf{r}} = f(\tilde{\boldsymbol{\xi}})$ .<sup>1</sup> 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 nonlinear, WVaR tends to severely *overestimate* the actual VaR because the covariance matrix  $\boldsymbol{\Sigma}_r$  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 §2.2 is symmetrically oriented around the mean vector  $\boldsymbol{\mu}_r$ . If the underlying assets of the derivatives have approximately symmetrically distributed returns, then the derivative returns are heavily skewed. An ellipsoidal uncertainty set fails to capture this asymmetry. This geometric argument supports our conjecture that WVaR provides a poor (overpessimistic) 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 nonlinear dependencies between the basic asset and derivative returns. In contrast, we assume that the moments of the derivative returns are unknown. In the next sections we derive generic worst-case value at risk models that explicitly account for nonlinear (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  $\mathbf{w} = (\mathbf{w}^\xi, \mathbf{w}^\eta)$ , where  $\mathbf{w}^\xi \in \mathbb{R}^n$  and  $\mathbf{w}^\eta \in \mathbb{R}^{m-n}$  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  $\mathbf{w} \in \mathcal{W}$  implies  $\mathbf{w}^\eta \geq \mathbf{0}$ . Recall that

<sup>1</sup> For ease of exposition, we assume that the returns of the derivative underliers are the only risk factors determining the option returns.

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. Because the first  $n$  components of  $\tilde{\mathbf{r}}$  represent the basic asset returns  $\tilde{\xi}$ , we have  $f_j(\tilde{\xi}) = \tilde{\xi}_j$  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  $\tilde{\xi}_i$ ,

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

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,

$$\begin{aligned} f_j(\tilde{\xi}) &= \max\{-1, a_j + b_j \tilde{\xi}_i - 1\}, \\ &\text{where } a_j = \frac{k_j - s_i}{p_j} \text{ and } b_j = -\frac{s_i}{p_j}. \end{aligned} \quad (12b)$$

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

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

where  $\mathbf{a} \in \mathbb{R}^{m-n}$ ,  $\mathbf{B} \in \mathbb{R}^{(m-n) \times n}$  are known constants determined through (12a) and (12b),  $\mathbf{e} \in \mathbb{R}^{m-n}$  is the vector of 1s, and “max” denotes the componentwise maximization operator. Thus, the return  $\tilde{r}_p$  of some portfolio  $\mathbf{w} \in \mathcal{W}$  can be expressed as

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

#### 4.2. Worst-Case Polyhedral VaR Model

For any portfolio  $\mathbf{w} \in \mathcal{W}$ , we define the *worst-case polyhedral VaR* as

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

where  $\mathcal{P}$  denotes the set of all probability distributions of the *basic* asset returns  $\tilde{\xi}$  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 SDP.

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

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

where

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

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

**PROOF.** See Appendix B.

Even though (16) constitutes a tractable SDP that enables us to compute the WPVaR of a given portfolio  $\mathbf{w} \in \mathcal{W}$  in polynomial time, it would be desirable to obtain an equivalent SOCP because SOCPs exhibit better scalability properties than SDPs (Alizadeh and Goldfarb 2003). In Theorem 4.2 we demonstrate that such a reformulation exists.

**THEOREM 4.2.** *Problem (16) can be reformulated as*

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

which constitutes a tractable SOCP.

**PROOF.** The proof follows a similar reasoning as in El Ghaoui et al. (2003, Theorem 1) and is thus omitted.  $\square$

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

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

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

$$\begin{aligned} & \underset{\mathbf{w} \in \mathbb{R}^m}{\text{minimize}} \quad \text{WPVaR}_\epsilon(\mathbf{w}) \\ & \text{subject to} \quad \mathbf{w} \in \mathcal{W}, \end{aligned}$$

which is equivalent to the tractable SOCP

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

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

**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  $\mathcal{C} = \{\boldsymbol{\Omega} \in \mathbb{S}^{n+1}: \boldsymbol{\Omega} \succcurlyeq \mathbf{0}, \underline{\boldsymbol{\Omega}} \leq \boldsymbol{\Omega} \leq \bar{\boldsymbol{\Omega}}\}$  for some given  $\underline{\boldsymbol{\Omega}}$  and  $\bar{\boldsymbol{\Omega}}$  of appropriate dimensions. In this situation, WPVaR can be equipped with a second layer of robustness that hedges against uncertainty in  $\boldsymbol{\Omega}$ . If  $\mathcal{C}$  contains a strictly positive definite matrix, we can use methods proposed by Natarajan et al. (2010, §2.5) to reformulate the resulting doubly robust risk measure  $\sup_{\boldsymbol{\Omega} \in \mathcal{C}} \text{WPVaR}_\epsilon(\mathbf{w})$  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 that 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 Natarajan et al. 2010). An efficiently computable SDP-based upper bound on WPVaR with partitioned statistics information can be constructed by extending the techniques developed by Natarajan et al. (2010, §2.4) in a straightforward manner.

### 4.3. Robust Optimization Perspective on WPVaR

In §2 we highlighted a known relationship between WVaR optimization and robust optimization. Moreover, in §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 that generates WPVaR is no longer symmetric.

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

$$\begin{aligned} & \underset{\mathbf{w} \in \mathbb{R}^m, \gamma \in \mathbb{R}}{\text{minimize}} \quad \gamma \\ & \text{subject to} \quad -\mathbf{w}^T \mathbf{r} \leq \gamma \quad \forall \mathbf{r} \in \mathcal{U}_\epsilon^p, \\ & \quad \mathbf{w} \in \mathcal{W}, \end{aligned} \tag{20}$$

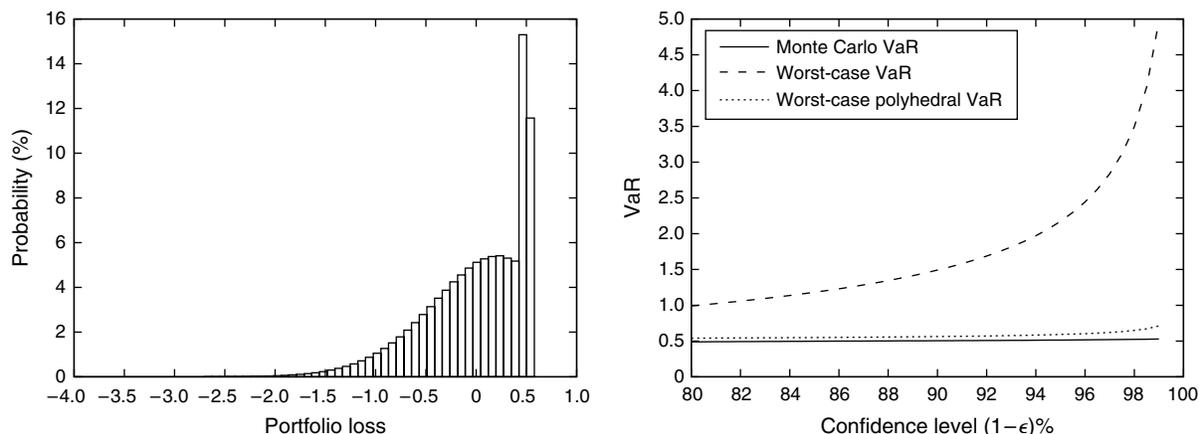
where the uncertainty set  $\mathcal{U}_\epsilon^p \subseteq \mathbb{R}^m$  is defined as

$$\begin{aligned} \mathcal{U}_\epsilon^p = \{ \mathbf{r} \in \mathbb{R}^m: \exists \boldsymbol{\xi} \in \mathbb{R}^n, \\ (\boldsymbol{\xi} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\xi} - \boldsymbol{\mu}) \leq \kappa(\epsilon)^2, \mathbf{r} = f(\boldsymbol{\xi}) \}. \end{aligned} \tag{21}$$

**PROOF.** The result is based on conic duality. We refer to Zymler et al. (2011b, Theorem 3.1) for a complete exposition of the proof.  $\square$

**EXAMPLE 4.1.** Consider a Black–Scholes economy consisting of stocks A and B, a European call option on stock A, and a European put option on stock B. Furthermore, let  $\mathbf{w}$  be an equally weighted portfolio of these  $m = 4$  assets, that is, set  $w_i = 1/m$  for  $i = 1, \dots, m$ . We assume that the prices of stocks A and B are governed by a bivariate geometric Brownian motion with drift coefficients of 12% and 8%, and volatilities of 30% and 20% per annum, respectively. The correlation between the instantaneous stock returns amounts to 20%. The initial prices of the stocks are \$100. The options mature in 21 days

**Figure 1** (Left) Portfolio Loss Distribution Obtained via Monte Carlo Simulation and (Right) VaR Estimates at Different Confidence Levels Obtained via Monte Carlo Sampling, WVaR, and WPVaR



Note. Negative values represent gains.

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 (Black and Scholes 1973), we obtain call and put option prices of \$3.58 and \$2.18, respectively.

We want to compute the VaR at confidence level  $\epsilon$  for portfolio  $w$  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  $\epsilon$ , the WVaR and WPVaR values exceed the Monte Carlo VaR estimate. This is not surprising because 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  $\epsilon$ , where the accuracy of the VaR estimate matters most. Indeed, for  $\epsilon = 1\%$ , the WVaR reports an unrealistically high value of 497%, which is seven times larger than the corresponding WPVaR value.

## 5. Worst-Case Quadratic VaR Optimization

The WPVaR model suffers from a number of weaknesses that may make it unattractive for certain

investors. First, 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 to obtain insurance benefits (see Zymler et al. 2011b), 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 that 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 §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 §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  $\tilde{\mathbf{s}}(t)$  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 nonbasic) asset  $i = 1, \dots, m$  is representable as  $v_i(\tilde{\mathbf{s}}(t), t)$ , where  $v_i: \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$  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  $i = 1, \dots, m$  we have

$$\begin{aligned} & v_i(\tilde{\mathbf{s}}(T), T) - v_i(\mathbf{s}(0), 0) \\ & \approx \bar{\theta}_i T + \bar{\Delta}_i^T (\tilde{\mathbf{s}}(T) - \mathbf{s}(0)) + \frac{1}{2} (\tilde{\mathbf{s}}(T) - \mathbf{s}(0))^T \bar{\Gamma}_i (\tilde{\mathbf{s}}(T) - \mathbf{s}(0)), \end{aligned}$$

where

$$\begin{aligned} \bar{\theta}_i &= \partial_i v_i(\mathbf{s}(0), 0) \in \mathbb{R}, \quad \bar{\Delta}_i = \nabla_s v_i(\mathbf{s}(0), 0) \in \mathbb{R}^n, \quad \text{and} \\ \bar{\Gamma}_i &= \nabla_s^2 v_i(\mathbf{s}(0), 0) \in \mathbb{S}^n. \end{aligned} \quad (22)$$

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  $(\mathbf{s}(0), 0)$  can be observed in the market. However, the values of  $v_i$  in a neighborhood of  $(\mathbf{s}(0), 0)$  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 Jaschke 2002). By using the *relative greeks*

$$\begin{aligned} \theta_i &= \frac{T}{v_i(\mathbf{s}(0), 0)} \bar{\theta}_i, \quad \Delta_i = \frac{1}{v_i(\mathbf{s}(0), 0)} \text{diag}(\mathbf{s}(0)) \bar{\Delta}_i, \\ \Gamma_i &= \frac{1}{v_i(\mathbf{s}(0), 0)} \text{diag}(\mathbf{s}(0))^T \bar{\Gamma}_i \text{diag}(\mathbf{s}(0)), \end{aligned}$$

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

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

Here we use the (possibly nonconvex) quadratic functions  $f_i$  to map the basic asset returns  $\tilde{\xi}$  to the asset returns  $\tilde{\mathbf{r}}$ . The return of a portfolio  $\mathbf{w} \in \mathcal{W}$  can therefore be approximated by

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

where we use the auxiliary functions

$$\begin{aligned} \theta(\mathbf{w}) &= \sum_{i=1}^m w_i \theta_i, \quad \Delta(\mathbf{w}) = \sum_{i=1}^m w_i \Delta_i, \quad \text{and} \\ \Gamma(\mathbf{w}) &= \sum_{i=1}^m w_i \Gamma_i, \end{aligned}$$

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

## 5.2. Worst-Case Quadratic VaR Model

We define the *worst-case quadratic VaR* of a fixed portfolio  $\mathbf{w} \in \mathcal{W}$  in terms of the Taylor expansion (24):

$$\begin{aligned} \text{WQVaR}_\epsilon(\mathbf{w}) &= \min \left\{ \gamma: \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{P}\{\gamma \leq -\mathbf{w}^T f(\tilde{\xi})\} \leq \epsilon \right\} \\ &= \min \left\{ \gamma: \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{P}\{\gamma \leq -\theta(\mathbf{w}) - \Delta(\mathbf{w})^T \tilde{\xi} \right. \\ &\quad \left. - \frac{1}{2} \tilde{\xi}^T \Gamma(\mathbf{w}) \tilde{\xi}\} \leq \epsilon \right\}. \end{aligned} \quad (25)$$

Note that the WQVaR approximates the portfolio return  $\mathbf{w}^T \tilde{\mathbf{r}}$  by a (possibly nonconvex) quadratic function of the basic asset returns  $\tilde{\xi}$ . Theorem 5.1 below shows how the WQVaR of a portfolio  $\mathbf{w}$  can be computed by solving a tractable SDP.

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

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

where  $\mathbf{\Omega}$  is the second-order moment matrix of  $\tilde{\xi}$ ; 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  $\mathbf{\Omega}$  by extending techniques proposed by Natarajan et al. (2010) 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  $\theta(\mathbf{w}) = 0$ ,  $\Delta(\mathbf{w}) = \mathbf{w}$ , and  $\Gamma(\mathbf{w}) = \mathbf{0}$ . In this special case, the WQVaR is computed by solving the following SDP:

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

El Ghaoui et al. (2003) showed (using similar arguments as in Theorem 4.2) that this SDP has the closed-form solution

$$\begin{aligned} \text{WVaR}_\epsilon(\mathbf{w}) = -\boldsymbol{\mu}^T \mathbf{w} + \kappa(\epsilon) \sqrt{\mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w}}, \\ \text{where } \kappa(\epsilon) = \sqrt{\frac{1-\epsilon}{\epsilon}}. \end{aligned}$$

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  $\mathbf{w} \in \mathcal{W}$ . Because the matrix inequality in (26) is linear in  $(\mathbf{M}, \tau, \gamma)$  and  $\mathbf{w}$ , one can reinterpret  $\mathbf{w}$  as a decision variable without impairing the problem's convexity. This observation reveals that we can efficiently minimize the WQVaR over all portfolios  $\mathbf{w} \in \mathcal{W}$  by solving the following SDP:

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

REMARK 5.3. Unlike in §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 linear matrix inequality (LMI) constraint is independent of  $\mathbf{w}$  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*

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

where

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

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

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

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 §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  $\Gamma(\mathbf{w}) \succcurlyeq \mathbf{0}$  is appended to the definition of the set  $\mathcal{W}$  of admissible portfolios, then the robust optimization problem (28) reduces to*

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

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

PROOF. See Appendix B.

REMARK 5.4. Note that the robust optimization problem (30) can be reformulated as

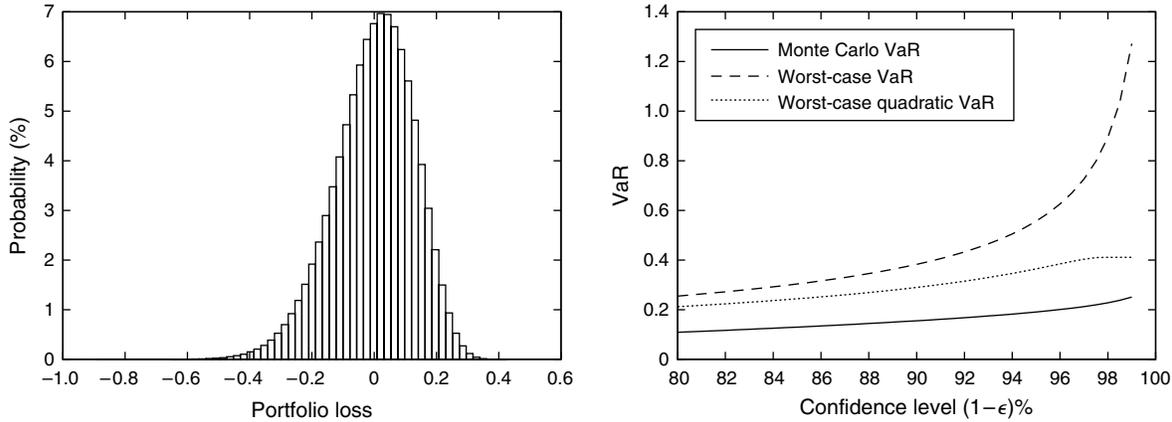
$$\begin{aligned} \text{minimize} \quad & \gamma \\ & \mathbf{w} \in \mathbb{R}^m, \gamma \in \mathbb{R} \\ \text{subject to} \quad & -\mathbf{w}^T \mathbf{r} \leq \gamma \quad \forall \mathbf{r} \in \mathcal{U}_\epsilon^{q2}, \\ & \mathbf{w} \in \mathcal{W}, \end{aligned} \quad (31)$$

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

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

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  $\mathbf{r}$ . As a result, the WQVaR model may provide a tighter

**Figure 2** (Left) Portfolio Loss Distribution Obtained via Monte Carlo Simulation and (Right) VaR Estimates at Different Confidence Levels Obtained via Monte Carlo Sampling, WVaR, and WQVaR



Note. Negative values represent gains.

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  $\Gamma(\mathbf{w}) \succcurlyeq \mathbf{0}$ , 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 nonconvexity 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. Because the VaR is no longer evaluated at the maturity time of the 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., MacMillan 1992). We use an analogous Monte Carlo simulation as in Example 4.1 to generate the stock and option returns over a two-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 three 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 §§4 and 5 and two conditional value at risk (CVaR) measures, which are coherent in the sense of Artzner et al. (1999). 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. (2009).

The classical CVaR is a quantile-based risk measure that evaluates the conditional expectation of the portfolio loss above VaR. For a given probability distribution  $\mathbb{P}$  of  $\tilde{\mathbf{r}}$  and tolerance  $\epsilon \in (0, 1)$  (Rockafellar and Uryasev 2002), define the CVaR of a portfolio  $\mathbf{w} \in \mathbb{R}^m$  as

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

where  $\mathbb{E}_{\mathbb{P}}(\cdot)$  denotes expectation with respect to  $\mathbb{P}$ . By construction, CVaR is convex in  $\mathbf{w}$  and—unlike VaR—constitutes a coherent risk measure (Acerbi and Tasche 2002). If only the mean vector  $\boldsymbol{\mu}_{\tilde{\mathbf{r}}}$  and covariance matrix  $\boldsymbol{\Sigma}_{\tilde{\mathbf{r}}} > \mathbf{0}$  of  $\tilde{\mathbf{r}}$  are known, CVaR does not admit exact evaluation. As in the case of VaR, it then proves useful to introduce the *worst-case CVaR*,

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

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

If a portfolio includes long positions in derivatives expiring at the end of the investment horizon as in §4, WCVaR may be an overly conservative risk measure.

In analogy to WPVaR, we thus introduce the worst-case polyhedral CVaR (WPCVaR),

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

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 that 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 §5, we can define the worst-case quadratic CVaR (WQCVaR) as follows:

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

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

**THEOREM 6.1.** *The following identities hold:*

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

**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 nonconvex) quadratic constraint functions Zymler et al. (2011a, Theorem 2.2).  $\square$

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

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

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

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

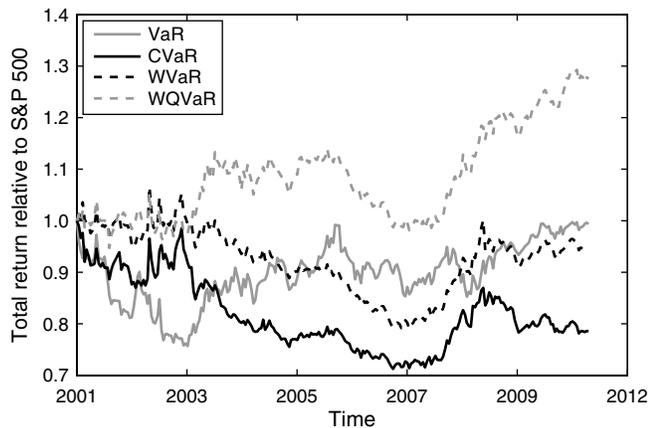
**PROOF.** The known coherence of the classical CVaR for any given distribution  $\mathbb{P}$  (see, e.g., Acerbi and Tasche 2002) implies via Zhu and Fukushima (2009, Proposition 1) that WPCVaR is coherent on  $R_P$ . By the equivalence of WPVaR and WPCVaR established in Theorem 6.1, WPVaR is thus a coherent risk measure on  $R_P$ . Coherence of WQVaR on  $R_Q$  is proved in a similar manner.  $\square$

Recall that the classical WVaR by El Ghaoui et al. (2003) 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  $\mathbf{w}^T \tilde{\mathbf{r}}$  for  $\mathbf{w} \in \mathbb{R}^m$ .

**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 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. 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, whereas 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., Gaivoronski and Pflug 2005, Rockafellar and Uryasev 2002). 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 §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 16 and 206 ms, respectively, using ILOG CPLEX, whereas the SOCPs

**Figure 3** Out-of-Sample Performance of Different Tracking Portfolios

and SDPs are solved in 91 and 195 ms, respectively, with SDPT3 (Tütüncü et al. 2003). 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 (Ye 1997). 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 the 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, whereas 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 the 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

**Table 1** (Top) Out-of-Sample Statistics of Excess Return over the Benchmark for Different Tracking Portfolios (in Basis Points) and (Bottom) 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 performance of the WVaR and WQVaR portfolios can only originate from the different treatment of the options in these models. The nonrobust 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 nonlinearly on their underlying assets. In this paper we explicitly incorporate this nonlinear 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, 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 nonlinear 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 nonrobust VaR.

## Acknowledgments

The authors are indebted to Aharon Ben-Tal for valuable discussions on the topics of this paper and to Grani Hanasusanto for his assistance with the numerical

experiments. This research was supported by the Engineering and Physical Sciences Research Council [Grant EP/I014640/1].

### Appendix 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. 2009). *Let  $\mathcal{S} \subseteq \mathbb{R}^n$  be any Borel measurable set (which is not necessarily convex), and define the worst-case probability  $\pi_{\text{wc}}$  as*

$$\pi_{\text{wc}} = \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{P}\{\tilde{\xi} \in \mathcal{S}\}, \quad (\text{A1})$$

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

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

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

### Appendix B. Proofs

PROOF OF THEOREM 4.1. To derive a manifestly tractable representation for WPVaR, we first simplify the maximization problem

$$\sup_{\mathbb{P} \in \mathcal{P}} \mathbb{P}\{\gamma \leq -(\mathbf{w}^\xi)^T \tilde{\xi} - (\mathbf{w}^\eta)^T \max\{-\mathbf{e}, \mathbf{a} + \mathbf{B}\tilde{\xi} - \mathbf{e}\}\}, \quad (\text{B1})$$

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

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

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

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

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

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

$$\mathcal{S}_\gamma = \left\{ \boldsymbol{\xi} \in \mathbb{R}^n : \max_{\mathbf{0} \leq \mathbf{y} \leq \mathbf{w}^\eta} \{\gamma + (\mathbf{w}^\xi)^T \boldsymbol{\xi} + \mathbf{y}^T (\mathbf{a} + \mathbf{B}\boldsymbol{\xi}) - \mathbf{e}^T \mathbf{w}^\eta\} \leq 0 \right\}. \quad (\text{B2})$$

The optimal value of problem (B1) is obtained by solving the worst-case probability problem  $\sup_{\mathbb{P} \in \mathcal{P}} \mathbb{P}\{\tilde{\xi} \in \mathcal{S}_\gamma\}$ , which, by Lemma A.1, can be expressed as

$$\begin{aligned} & \inf_{\mathbf{M} \in \mathbb{S}^{n+1}} \langle \boldsymbol{\Omega}, \mathbf{M} \rangle \\ & \text{s.t. } [\boldsymbol{\xi}^T \ 1] \mathbf{M} [\boldsymbol{\xi}^T \ 1]^T \geq 1 \ \forall \boldsymbol{\xi} : \\ & \quad \max_{\mathbf{0} \leq \mathbf{y} \leq \mathbf{w}^\eta} \{\gamma + (\mathbf{w}^\xi)^T \boldsymbol{\xi} + \mathbf{y}^T (\mathbf{a} + \mathbf{B}\boldsymbol{\xi}) - \mathbf{e}^T \mathbf{w}^\eta\} \leq 0, \\ & \quad \mathbf{M} \succcurlyeq \mathbf{0}. \end{aligned} \quad (\text{B3})$$

We will now argue that (B3) is equivalent to problem (B4),

$$\begin{aligned} & \inf \langle \boldsymbol{\Omega}, \mathbf{M} \rangle \\ & \text{s.t. } \mathbf{M} \in \mathbb{S}^{n+1}, \tau \in \mathbb{R}, \mathbf{M} \succcurlyeq \mathbf{0}, \tau \geq 0, \\ & \quad [\boldsymbol{\xi}^T \ 1] \mathbf{M} [\boldsymbol{\xi}^T \ 1]^T - 1 \\ & \quad + 2\tau \left( \max_{\mathbf{0} \leq \mathbf{y} \leq \mathbf{w}^\eta} \{\gamma + (\mathbf{w}^\xi)^T \boldsymbol{\xi} + \mathbf{y}^T (\mathbf{a} + \mathbf{B}\boldsymbol{\xi}) - \mathbf{e}^T \mathbf{w}^\eta\} \right) \geq 0 \\ & \quad \forall \boldsymbol{\xi} \in \mathbb{R}^n, \end{aligned} \quad (\text{B4})$$

for all but one value of  $\gamma$ . For ease of exposition, we first define

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

The equivalence of (B3) and (B4) is proved case by case. Assume first that  $h < 0$ . Then, the equivalence follows from the nonlinear Farkas lemma (see, e.g., Pólik and Terlaky 2007, Theorem 2.1). Assume next that  $h > 0$ . Then, the semi-infinite constraint in (B3) becomes redundant, and because  $\boldsymbol{\Omega} \succ \mathbf{0}$ , the optimal solution of (B3) is given by  $\mathbf{M} = \mathbf{0}$  with a corresponding optimal value of 0. The optimal value of problem (B4) is also equal to 0. Indeed, by choosing  $\tau = 1/h$ , the semi-infinite constraint in (B4) is satisfied independently of  $\mathbf{M}$ . Finally, (B3) and (B4) may differ for  $h = 0$ .

It can be seen that because  $\tau \geq 0$ , the semi-infinite constraint in (B4) is equivalent to the assertion that there exists some  $\mathbf{0} \leq \mathbf{y} \leq \mathbf{w}^\eta$  with

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

This semi-infinite constraint can be written as

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

Thus, problem (B4) can equivalently be formulated as

$$\begin{aligned} & \inf \langle \boldsymbol{\Omega}, \mathbf{M} \rangle \\ & \text{s.t. } \mathbf{M} \in \mathbb{S}^{n+1}, \mathbf{y} \in \mathbb{R}^{m-n}, \tau \in \mathbb{R}, \\ & \quad \mathbf{M} \succcurlyeq \mathbf{0}, \tau \geq 0, \mathbf{0} \leq \mathbf{y} \leq \mathbf{w}^\eta, \\ & \quad \mathbf{M} + \begin{bmatrix} \mathbf{0} & \tau(\mathbf{w}^\xi + \mathbf{B}^T \mathbf{y}) \\ \tau(\mathbf{w}^\xi + \mathbf{B}^T \mathbf{y})^T & -1 + 2\tau(\gamma + \mathbf{y}^T \mathbf{a} - \mathbf{e}^T \mathbf{w}^\eta) \end{bmatrix} \succcurlyeq \mathbf{0}. \end{aligned} \quad (\text{B5})$$

Because (B3) and (B5) are equivalent for all but one value of  $\gamma$ , and because their optimal values are nonincreasing in  $\gamma$ , we can express WPVaR in (15) as the optimal value of the following problem:

$$\begin{aligned} & \text{WPVaR}_\epsilon(\mathbf{w}) \\ &= \inf \gamma \\ & \text{s.t. } \mathbf{M} \in \mathbb{S}^{n+1}, \mathbf{y} \in \mathbb{R}^{m-n}, \tau \in \mathbb{R}, \gamma \in \mathbb{R}; \end{aligned}$$

$$\langle \mathbf{\Omega}, \mathbf{M} \rangle \leq \epsilon, \mathbf{M} \succcurlyeq \mathbf{0}, \tau \geq 0, \mathbf{0} \leq \mathbf{y} \leq \mathbf{w}^n, \\ \mathbf{M} + \begin{bmatrix} \mathbf{0} & \tau(\mathbf{w}^\xi + \mathbf{B}^T \mathbf{y}) \\ \tau(\mathbf{w}^\xi + \mathbf{B}^T \mathbf{y})^T & -1 + 2\tau(\gamma + \mathbf{y}^T \mathbf{a} - \mathbf{e}^T \mathbf{w}^n) \end{bmatrix} \succcurlyeq \mathbf{0}. \quad (\text{B6})$$

Problem (B6) is nonconvex because of 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  $\tau$ -component. However, because  $\epsilon < 1$ , 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  $\tau$ -component. This allows us to divide the matrix inequality in problem (B6) by  $\tau$ . Subsequently we perform variable substitutions in which we replace  $\tau$  by  $1/\tau$  and  $\mathbf{M}$  by  $\mathbf{M}/\tau$ . This shows that (B6) is equivalent to problem (16).  $\square$

**PROOF OF THEOREM 5.1** For the given portfolio  $\mathbf{w} \in \mathcal{W}$  and for any fixed  $\gamma \in \mathbb{R}$ , we introduce the set  $\mathcal{Q}_\gamma \subseteq \mathbb{R}^n$ , defined through

$$\mathcal{Q}_\gamma = \{ \xi \in \mathbb{R}^n : \gamma \leq -\theta(\mathbf{w}) - \mathbf{\Delta}(\mathbf{w})^T \xi - \frac{1}{2} \xi^T \mathbf{\Gamma}(\mathbf{w}) \xi \}. \quad (\text{B7})$$

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

$$\pi_{\text{wc}} = \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{P}\{ \tilde{\xi} \in \mathcal{Q}_\gamma \}, \quad (\text{B8})$$

which can be identified as the subordinate maximization problem in (25). Lemma A.1 implies that (B8) can equivalently be formulated as

$$\pi_{\text{wc}} = \inf_{\mathbf{M} \in \mathbb{S}^{n+1}} \{ \langle \mathbf{\Gamma}, \mathbf{M} \rangle : \mathbf{M} \succcurlyeq \mathbf{0}, [\xi^T \ 1] \mathbf{M} [\xi^T \ 1]^T \geq 1 \ \forall \xi \in \mathcal{Q}_\gamma \}. \quad (\text{B9})$$

By the definition of  $\mathcal{Q}$ , the semi-infinite constraint in problem (B9) is equivalent to

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

By using the  $\mathcal{S}$ -lemma (Pólik and Terlaky 2007) and by analogous reasoning as in §4.2, we can replace the semi-infinite constraint in problem (B9) by

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

without changing the optimal value of the problem. Thus, the worst-case probability problem (B8) can be rewritten as

$$\pi_{\text{wc}} = \inf \langle \mathbf{\Omega}, \mathbf{M} \rangle \\ \text{s.t. } \mathbf{M} \in \mathbb{S}^{n+1}, \tau \in \mathbb{R}, \mathbf{M} \succcurlyeq \mathbf{0}, \tau \geq 0; \\ \mathbf{M} + \begin{bmatrix} \tau \mathbf{\Gamma}(\mathbf{w}) & \tau \mathbf{\Delta}(\mathbf{w}) \\ \tau \mathbf{\Delta}(\mathbf{w})^T & -1 + 2\tau(\gamma + \theta(\mathbf{w})) \end{bmatrix} \succcurlyeq \mathbf{0}. \quad (\text{B10})$$

The WQVaR of the portfolio  $\mathbf{w}$  can therefore be obtained by solving the following nonconvex optimization problem:

$$\text{WQVaR}_\epsilon(\mathbf{w}) \\ = \inf \gamma \\ \text{s.t. } \mathbf{M} \in \mathbb{S}^{n+1}, \tau \in \mathbb{R}, \gamma \in \mathbb{R}; \quad (\text{B11})$$

$$\langle \mathbf{\Omega}, \mathbf{M} \rangle \leq \epsilon, \mathbf{M} \succcurlyeq \mathbf{0}, \tau \geq 0; \\ \mathbf{M} + \begin{bmatrix} \tau \mathbf{\Gamma}(\mathbf{w}) & \tau \mathbf{\Delta}(\mathbf{w}) \\ \tau \mathbf{\Delta}(\mathbf{w})^T & -1 + 2\tau(\gamma + \theta(\mathbf{w})) \end{bmatrix} \succcurlyeq \mathbf{0}.$$

By analogous reasoning as in §4.2, it can be shown that any feasible solution of problem (B11) has a strictly positive  $\tau$ -component. Thus we may divide the matrix inequality in (B11) by  $\tau$ . After the variable transformation  $\tau \rightarrow 1/\tau$  and  $\mathbf{M} \rightarrow \mathbf{M}/\tau$ , we obtain the postulated SDP (26).  $\square$

**PROOF OF THEOREM 5.2** For some fixed portfolio  $\mathbf{w} \in \mathcal{W}$ , the WQVaR can be computed by solving problem (26), which involves the LMI constraint

$$\mathbf{M} + \begin{bmatrix} \mathbf{\Gamma}(\mathbf{w}) & \mathbf{\Delta}(\mathbf{w}) \\ \mathbf{\Delta}(\mathbf{w})^T & -\tau + 2(\gamma + \theta(\mathbf{w})) \end{bmatrix} \succcurlyeq \mathbf{0}. \quad (\text{B12})$$

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

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

With this new notation, the LMI constraint (B12) is representable as

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

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

$$\inf \sup_{\xi \in \mathbb{R}^n} \{ -\frac{1}{2} \xi^T (\mathbf{V} + \mathbf{\Gamma}(\mathbf{w})) \xi \\ - \xi^T (\mathbf{v} + \mathbf{\Delta}(\mathbf{w})) - \theta(\mathbf{w}) - \frac{1}{2}(u - \tau) \} \\ \text{s.t. } \mathbf{V} \in \mathbb{S}^n, \mathbf{v} \in \mathbb{R}^n, \tau \in \mathbb{R}, u \in \mathbb{R}; \quad (\text{B13})$$

$$\begin{bmatrix} \mathbf{V} & \mathbf{v} \\ \mathbf{v}^T & u \end{bmatrix} \succcurlyeq \mathbf{0}, \tau \geq 0, \langle \mathbf{V}, \mathbf{\Sigma} + \mu \mu^T \rangle + 2\mathbf{v}^T \mu + u \leq \tau \epsilon.$$

Note that if  $\mathbf{V} + \mathbf{\Gamma}(\mathbf{w})$  is not positive semidefinite, the inner maximization problem in (B13) is unbounded. However, this implies that any  $\mathbf{V} \in \mathbb{S}^n$  is infeasible in the outer minimization problem unless  $\mathbf{V} + \mathbf{\Gamma}(\mathbf{w}) \succcurlyeq \mathbf{0}$ . Therefore, we can

add the constraint  $\mathbf{V} + \mathbf{\Gamma}(\mathbf{w}) \succcurlyeq \mathbf{0}$  to the minimization problem in (B13) without changing its feasible region. With this constraint appended, the min–max problem (B13) becomes a saddlepoint problem because its objective is concave in  $\xi$  for any fixed  $(\mathbf{V}, \mathbf{v}, u, \tau)$  and convex in  $(\mathbf{V}, \mathbf{v}, u, \tau)$  for any fixed  $\xi$ . 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:

$$\begin{aligned} \max_{\xi \in \mathbb{R}^n} \min \{ & -\frac{1}{2} \xi^T (\mathbf{V} + \mathbf{\Gamma}(\mathbf{w})) \xi - \xi^T (\mathbf{v} + \mathbf{\Delta}(\mathbf{w})) \\ & - \theta(\mathbf{w}) - \frac{1}{2}(u - \tau) \} \\ \text{s.t. } & \mathbf{V} \in \mathbb{S}^n, \mathbf{v} \in \mathbb{R}^n, \tau \in \mathbb{R}, u \in \mathbb{R}; \\ & \begin{bmatrix} \mathbf{V} & \mathbf{v} \\ \mathbf{v}^T & u \end{bmatrix} \succcurlyeq \mathbf{0}, \tau \geq 0, \langle \mathbf{V}, \mathbf{\Sigma} + \boldsymbol{\mu} \boldsymbol{\mu}^T \rangle \\ & + 2\mathbf{v}^T \boldsymbol{\mu} + u \leq \tau \epsilon. \end{aligned} \quad (\text{B14})$$

We proceed by dualizing the inner minimization problem in (B14). After a few elementary simplification steps, this dual problem reduces to

$$\begin{aligned} \max \{ & -\frac{1}{2} \langle \mathbf{\Gamma}(\mathbf{w}), \xi \xi^T + \mathbf{Y} \rangle - \xi^T \mathbf{\Delta}(\mathbf{w}) - \theta(\mathbf{w}) \} \\ \text{s.t. } & \mathbf{Y} \in \mathbb{S}^n, \alpha \in \mathbb{R}, \mathbf{Y} \succcurlyeq \mathbf{0}, 1 \leq \alpha \leq 1/\epsilon, \\ & \begin{bmatrix} \alpha(\mathbf{\Sigma} + \boldsymbol{\mu} \boldsymbol{\mu}^T) - (\xi \xi^T + \mathbf{Y}) & \alpha \boldsymbol{\mu} - \xi \\ (\alpha \boldsymbol{\mu} - \xi)^T & \alpha - 1 \end{bmatrix} \succcurlyeq \mathbf{0}. \end{aligned} \quad (\text{B15})$$

Note that strong duality holds because the inner problem in (B14) is strictly feasible for any  $\epsilon > 0$  (see Vandenberghe and Boyd 1996). This allows us to replace the inner minimization problem in (B14) by the maximization problem (B15), which yields the following equivalent formulation for the WQVaR problem (26):

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

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

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

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

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

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

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

Therefore, problem (B15) can be reformulated as

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

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

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

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

$$\min_{\mathbf{w} \in \mathcal{W}} \max_{\mathbf{Z} \in \mathcal{U}_\epsilon^{\mathbf{Z}}} \{ -\langle \mathbf{Q}(\mathbf{w}), \mathbf{Z} \rangle \}, \quad (\text{B16})$$

which is manifestly equivalent to the postulated semi-infinite program (28).  $\square$

**PROOF OF COROLLARY 5.1** The inner maximization problem in (B16) can be written as

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

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

$$\begin{aligned} \max \{ & -\theta(\mathbf{w}) - \mathbf{\Delta}(\mathbf{w})^T \xi - \frac{1}{2} \xi^T \mathbf{\Gamma}(\mathbf{w}) \xi - \frac{1}{2} \langle \mathbf{\Gamma}(\mathbf{w}), \mathbf{Y} \rangle \} \\ \text{s.t. } & \mathbf{Y} \in \mathbb{S}^n, \xi \in \mathbb{R}^n, \mathbf{Y} \succcurlyeq \mathbf{0}; \\ & \begin{bmatrix} (\mathbf{\Sigma} + \boldsymbol{\mu} \boldsymbol{\mu}^T) - \epsilon(\mathbf{Y} + \xi \xi^T) & \boldsymbol{\mu} - \epsilon \xi \\ (\boldsymbol{\mu} - \epsilon \xi)^T & 1 - \epsilon \end{bmatrix} \succcurlyeq \mathbf{0}. \end{aligned} \quad (\text{B17})$$

For all but one value of  $\gamma$ . 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  $\langle \mathbf{\Gamma}(\mathbf{w}), \mathbf{Y} \rangle \geq 0$  for all  $\mathbf{Y} \succcurlyeq \mathbf{0}$  because  $\mathbf{\Gamma}(\mathbf{w}) \succcurlyeq \mathbf{0}$  by assumption. Thus problem (B17) reduces to

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

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

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

Thus the original min-max formulation (B16) can be reexpressed as

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

which is equivalent to the postulated robust optimization problem.  $\square$

## References

- Acerbi C, Tasche D (2002) On the coherence of expected shortfall. *J. Banking Finance* 26(7):1487–1503.
- Alizadeh F, Goldfarb D (2003) Second-order cone programming. *Math. Programming* 95(1):3–51.
- Artzner P, Delbaen F, Eber J, Heath D (1999) Coherent measures of risk. *Math. Finance* 9(3):203–228.
- Ben-Tal A, Nemirovski A (1998) Robust convex optimization. *Math. Oper. Res.* 23(4):769–805.
- Ben-Tal A, Nemirovski A (1999) Robust solutions of uncertain linear programs. *Oper. Res. Lett.* 25(1):1–13.
- Ben-Tal A, El Ghaoui L, Nemirovski A (2009) *Robust Optimization* (Princeton University Press, Princeton, NJ).
- Bertsimas D, Sim M (2004) The price of robustness. *Oper. Res.* 52(1):35–53.
- Black F, Scholes M (1973) The pricing of options and corporate liabilities. *J. Political Econom.* 81(3):637–654.
- Calafiore G, Topcu U, El Ghaoui L (2009) Parameter estimation with expected and residual-at-risk criteria. *Systems Control Lett.* 58(1):39–46.
- Coval J, Shumway T (2002) Expected option returns. *J. Finance* 56(3):983–1009.
- El Ghaoui L, Oks M, Oustry F (2003) Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Oper. Res.* 51(4):543–556.
- Gaivoronski A, Pflug G (2005) Value-at-risk in portfolio optimization: Properties and computational approach. *J. Risk* 7(2):1–31.
- Jaschke S (2002) The cornish-fisher expansion in the context of delta-gamma-normal approximations. *J. Risk* 4(4):33–53.
- Jorion P (2001) *Value-at-Risk: The New Benchmark for Managing Financial Risk* (McGraw-Hill, New York).
- MacMillan L (1992) *Options as a Strategic Investment* (Prentice Hall, Upper Saddle River, NJ).
- Markowitz H (1952) Portfolio selection. *J. Finance* 7(1):77–91.
- Mina J, Ulmer A (1999) Delta-gamma four ways. Technical report, RiskMetrics Group, New York.
- Natarajan K, Pachamanova D, Sim M (2008) Incorporating asymmetric distributional information in robust value-at-risk optimization. *Management Sci.* 54(3):573–585.
- Natarajan K, Pachamanova D, Sim M (2009) Constructing risk measures from uncertainty sets. *Oper. Res.* 57(5):1129–1141.
- Natarajan K, Sim M, Uichanco J (2010) Tractable robust expected utility and risk models for portfolio optimisation. *Math. Finance* 20(4):695–731.
- Pólik I, Terlaky T (2007) A survey of the S-lemma. *SIAM Rev.* 49(3):371–481.
- Rockafellar R, Uryasev S (2002) Optimization of conditional value-at-risk. *J. Risk* 2:21–41.
- Rustem B, Howe M (2002) *Algorithms for Worst-Case Design and Applications to Risk Management* (Princeton University Press, Princeton, NJ).
- Tütüncü RH, Toh KC, Todd MJ (2003) Solving semidefinite-quadratic-linear programs using SDPT3. *Math. Programming Ser. B* 95(2):189–217.
- Vandenberghe L, Boyd S (1996) Semidefinite programming. *SIAM Rev.* 38(1):49–95.
- Ye Y (1997) *Interior Point Algorithms: Theory and Analysis* (John Wiley & Sons, New York).
- Zhu S, Fukushima M (2009) Worst-case conditional value-at-risk with application to robust portfolio management. *Oper. Res.* 57(5):1155–1168.
- Zymler S, Kuhn D, Rustem B (2011a) Distributionally robust joint chance constraints with second-order moment information. *Math. Programming*, ePub ahead of print November 10, <http://dx.doi.org/10.1007/s10107-011-0494-7>.
- Zymler S, Rustem B, Kuhn D (2011b) Robust portfolio optimization with derivative insurance guarantees. *Eur. J. Oper. Res.* 210(2):410–424.