

# Two-stage Stochastic Programming under Multivariate Risk Constraints with an Application to Humanitarian Relief Network Design

Nilay Noyan

Industrial Engineering Program, Sabancı University, Turkey  
nnoyan@sabanciuniv.edu

Merve Meraklı, Simge Küçükyavuz

Industrial and Systems Engineering, University of Washington, U.S.A.  
merakli@uw.edu, simge@uw.edu

January 10, 2017

**ABSTRACT:** In this study, we consider two classes of multicriteria two-stage stochastic programs in finite probability spaces with multivariate risk constraints. The first-stage problem features a multivariate stochastic benchmarking constraint based on a vector-valued random variable representing multiple and possibly conflicting stochastic performance measures associated with the second-stage decisions. In particular, the aim is to ensure that the associated random outcome vector of interest is preferable to a specified benchmark with respect to the multivariate polyhedral conditional value-at-risk (CVaR) or a multivariate stochastic order relation. In this case, the classical decomposition methods cannot be used directly due to the complicating multivariate stochastic benchmarking constraints. We propose an exact unified decomposition framework for solving these two classes of optimization problems and show its finite convergence. We apply the proposed approach to a stochastic network design problem in a pre-disaster humanitarian logistics context and conduct a computational study concerning the threat of hurricanes in the Southeastern part of the United States. Our numerical results on these large-scale problems show that our proposed algorithm is computationally scalable.

*Keywords:* stochastic programming; multicriteria optimization; risk-averse two-stage; multivariate risk; conditional value-at-risk; Benders decomposition; branch-and-cut; network design; pre-disaster; humanitarian relief

---

**1. Introduction** Two-stage stochastic programming is a vibrant research area, which provides a natural and widely applicable modeling framework for decision making problems under uncertainty in a large variety of fields. Such models are well-suited for the situations where decisions are made in two stages; the first-stage decisions are made before the uncertainty is revealed, and the second-stage (recourse) decisions are made given the predetermined first-stage decisions and the observed realization of the random parameters. For many practical decision making problems, it is essential to evaluate the decisions according to multiple and possibly conflicting stochastic criteria. Along these lines, we focus on two-stage stochastic programming models involving first-stage decisions leading to uncertain outcomes that can be evaluated according to multiple stochastic performance measures associated with the corresponding second-stage decisions.

In multicriteria stochastic optimization, it is common to employ a weighted-sum approach, which aggregates the specified multiple stochastic objective criteria to obtain a single objective function. Following this mainstream approach, we formulate the second-stage problem as a multiobjective optimization problem and focus on its expected optimal objective value in the first stage. In addition, one may also prefer to impose constraints on some performance measures associated with the second-stage decisions. To this end, we employ stochastic benchmarking preference relations between the vector-valued random outcomes of the second-stage decisions. In particular, in the first class of problems we study, the first-stage problem enforces the multivariate polyhedral CVaR constraint introduced by [Noyan and Rudolf \(2013\)](#). This approach considers a family of linear scalarization functions (a polyhedral set of scalarization weight vectors) and requires all scalarized versions of the decision-based random vector of outcomes to be preferable to the corresponding scalarizations of an existing reference (benchmark) outcome according to the univariate CVaR relation. Furthermore, allowing arbitrary polyhedra as scalarization sets also provides a

good balance between flexibility and computational tractability. To the best of our knowledge, our study is a first in developing such a risk-averse two-stage model featuring multivariate risk measure-based constraints. We also note that our proposed modeling framework is not limited to the risk-neutral objective function; it can be easily extended to the case where the objective function features also a risk measure representing the decision makers' risk preferences regarding the random recourse cost. For acceptable risk-averse objectives, which lead to computationally tractable two-stage stochastic programming models, we refer the reader to [Ahmed \(2006\)](#) and [Dupáčová \(2008\)](#).

Optimization problems with the multivariate stochastic benchmarking constraints have been receiving increasing attention in the literature. Using multivariate stochastic relations is essential to capture the correlation between the multiple random outcomes. The majority of the existing studies extend a univariate stochastic preference relation, based on second-order stochastic dominance (SSD) or a coherent risk measure such as CVaR, to the multivariate case by considering a family of scalarization functions. While multivariate SSD-constrained problems are more frequently studied (see, e.g., [Dentcheva and Ruszczyński, 2009](#); [Homem-de-Mello and Mehrotra, 2009](#); [Hu et al., 2012](#); [Dentcheva and Wolfhagen, 2015; 2016](#)), there is an increasing interest in risk measure-constrained variants (see, e.g., [Noyan and Rudolf, 2013](#); [Liu et al., 2015](#)), which provide natural relaxations to overly demanding and conservative SSD-based models. More recently, studies focusing on both multivariate SSD- and risk measure-constrained models also appear in literature ([Küçükyavuz and Noyan, 2016](#); [Noyan and Rudolf, 2016](#)). In these existing studies, the scalarization functions are almost exclusively of the linear type with the exception of [Noyan and Rudolf \(2016\)](#) who extend the multivariate risk-constrained models by incorporating a general class of scalarization functions (including a variety of non-linear scalarization functions). Moreover, we point out that the scalarization-based line of research has been dedicated almost entirely to single-stage (static) decision-making problems. The only exception is the work of [Dentcheva and Wolfhagen \(2016\)](#), which introduces a two-stage stochastic optimization problem involving a multivariate SSD (more precisely, referred to as the increasing convex ordering in the context of cost minimization) constraint on a random outcome vector of the second-stage decisions with respect to the unit simplex of scalarization vectors. The authors present two approximate decomposition-based solution algorithms which rely on the Lagrangian relaxation of the multivariate stochastic ordering constraint and show their finite convergence to an  $\varepsilon$ -feasible  $\varepsilon$ -optimal solution, even if the probability space is not finite; their results and algorithms are adaptations of those presented in their earlier study ([Dentcheva and Wolfhagen, 2015](#)) for the single-stage case. We contribute to this line of research by introducing an alternative two-stage model with multivariate CVaR constraints. In addition, we consider the SSD-based counterpart (as in [Dentcheva and Wolfhagen, 2015](#)) and provide a new computationally tractable and exact solution algorithm for this problem class. We defer a detailed discussion on the advantages of our solution method compared to the existing approaches for the multivariate SSD-constrained two-stage models to Section 3.

Although, it is not directly related to the multicriteria stochastic optimization problems of our interest, we note that there are a few studies on risk-averse two-stage models with univariate (first- or second-order) stochastic dominance and CVaR-based constraints (see, e.g., [Fábián, 2008](#); [Gollmer et al., 2011](#); [Dentcheva and Martinez, 2012](#)). However, they employ the risk constraints to compare scalar-based random variables. Our study extends such risk-averse modeling approaches to the multicriteria case, allowing us to consider additional stochastic criteria associated with the second-stage decisions other than the random recourse cost.

The proposed modeling approach with the multivariate risk constraints is a fairly recent research area, and it has promise to be applied in a wide variety of fields. This approach is particularly well-suited for the field of humanitarian logistics, since incorporating risk is crucial for rarely occurring disaster events

(e.g., Noyan, 2012), and considering multiple conflicting performance criteria (such as efficiency and equity) is often essential for the effectiveness of the relief response systems (e.g., Vitoriano et al., 2011; Huang et al., 2012; Gutjahr and Nolz, 2016). Motivated by the significance of the long-term pre-disaster planning, we apply the proposed framework to a stochastic pre-disaster relief network design problem. Hence, our study also contributes to the humanitarian relief literature by introducing a new risk-averse two-stage optimization model, which provides a flexible and tractable way of considering decision makers' risk preferences based on multiple stochastic criteria.

Next we summarize our contributions and give an outline of our paper. In Section 2, we describe a novel two-stage stochastic program with multivariate CVaR constraints. It is well-known that risk-neutral two-stage stochastic programs are generally hard to solve due to a potentially large number of scenario-dependent recourse decisions. Thus, introducing a multivariate stochastic benchmarking relation, which enforces a collection of risk constraints associated with a scalarization set, further complicates the solution of these optimization models. A common approach in solving such large-scale stochastic optimization models is to employ a Benders-type scenario decomposition approach (e.g., Ruszczyński and Shapiro, 2003). The classical decomposition methods cannot be directly applied to our model due to the complicating risk constraints. Utilizing successive relaxations of the multivariate polyhedral CVaR relation, we develop two types of exact and finitely convergent delayed cut generation solution algorithms; the iteratively generated cuts – common in both algorithms – are associated with the scalarization vectors for which the risk constraints are relaxed. In addition, the second algorithm adapts a scenario decomposition approach that exploits the decomposable structure of CVaR and the second-stage problems; this further decomposition proves to be useful in solving larger problem instances. We also develop strong duality results and optimality conditions under certain linearity assumptions. Despite our focus on the CVaR-based models, in Section 3, we show that our proposed solution methods also apply to the multivariate SSD-constrained two-stage models. In Section 4, we apply the proposed modeling approach to a stochastic pre-disaster relief network design problem and present the corresponding mathematical programming formulations. Section 5 is dedicated to the computational study.

**2. Two-Stage Optimization with Multivariate CVaR Constraints** In this section, we introduce the multicriteria stochastic decision-making framework of interest, and present the proposed multivariate CVaR-constrained two-stage stochastic programming model and the corresponding delayed cut generation solution algorithms.

We consider a finite probability space  $(\Omega, 2^\Omega, \Pi)$  with  $\Omega = \{\omega_1, \dots, \omega_m\}$  and  $\Pi(\omega_s) = p_s$ . Let  $S = \{1, \dots, m\}$  be the index set of the elementary events (also referred to as scenarios). We assume that a benchmark random outcome vector and a polyhedron of scalarization vectors, each component of which represents the relative importance of each decision criterion, are given (specified by decision makers). Let  $\mathbf{Z}$  be the benchmark vector with realizations  $\mathbf{z}_1, \dots, \mathbf{z}_{|T|}$  and associated probabilities  $\tilde{p}_i$ ,  $i \in T$ ; it is often constructed from some benchmark decision in which case we have  $T = S$  and  $\tilde{p}_i = p_i$ ,  $i \in S$ . The scalarization set  $\mathcal{C}$  is naturally assumed to be a subset of the unit simplex without loss of generality. In our setting, smaller values of random variables and risk measures are considered to be preferable. We now introduce the class of two-stage stochastic programming problems with multivariate CVaR constraints, where the general form of the first-stage problem is given by

$$\min f(\mathbf{x}) + \mathbb{E}(Q(\mathbf{x}, \boldsymbol{\xi}(\omega))) \quad (1a)$$

$$\text{s.t. } \text{CVaR}_\alpha(\mathbf{c}^\top \mathbf{G}(\mathbf{x}, \boldsymbol{\xi}(\omega))) \leq \text{CVaR}_\alpha(\mathbf{c}^\top \mathbf{Z}), \quad \forall \mathbf{c} \in \mathcal{C}, \quad (1b)$$

$$\mathbf{x} \in \mathcal{X}. \quad (1c)$$

Here,  $f(\mathbf{x})$  is a convex objective function,  $\mathcal{X} \subset \mathbb{R}_+^{\bar{n}_1} \times \mathbb{Z}_+^{n_1 - \bar{n}_1}$  is a non-empty convex set (such as a polyhedron) of the first-stage decision variables, defined by the deterministic constraints of the problem.

$\xi(\omega) \subset \mathbb{R}^r$  and  $Q(\mathbf{x}, \xi(\omega_s))$  designate the vector of the random input parameters of the second-stage problem and the optimal second-stage objective value under scenario  $s$ , respectively. More specifically, introducing the notation  $\xi(\omega_s) = (\mathbf{q}(\omega_s), T(\omega_s), W(\omega_s))$ ,  $\mathbf{h}(\omega_s) = (\mathbf{q}_s, T_s, W_s, \mathbf{h}_s)$ ,

$$Q(\mathbf{x}, \xi(\omega_s)) = \min_{\mathbf{y} \in \mathcal{Y}(\mathbf{x}, \xi(\omega_s))} \mathbf{q}_s^\top \mathbf{y}, \quad (2)$$

where  $\mathcal{Y}(\mathbf{x}, \xi(\omega_s)) = \{\mathbf{y} \in \mathbb{R}_+^{n_2} : T_s \mathbf{x} + W_s \mathbf{y} \geq \mathbf{h}_s\}$ . The second-stage objective function  $\mathbf{q}_s^\top \mathbf{y}$  could be a weighted sum of multiple objectives as in [Dentcheva and Wolflagen \(2016\)](#), or a single objective, such as cost, that is compatible with the first-stage objective. In addition, for some mapping  $\mathbf{G} : \mathcal{X} \times \mathbb{R}^r \rightarrow \mathbb{R}^d$ ,  $\mathbf{G}(\mathbf{x}, \xi(\omega))$  is the decision-based random vector featured in (1b). In particular,  $\mathbf{G}(\mathbf{x}, \xi(\omega_s))$  represents the multiple random performance measures of interest associated with the optimal second-stage decisions  $\mathbf{y}(\omega_s) \in \arg \min_{\mathbf{y} \in \mathcal{Y}(\mathbf{x}, \xi(\omega_s))} \{\mathbf{q}_s^\top \mathbf{y}\}$ ,  $s \in S$ , for given  $\mathbf{x}$  and  $\xi(\omega_s)$ . To highlight this correspondence we introduce the random vector  $\hat{\mathbf{G}}(\mathbf{x}, \mathbf{y})$  given by  $[\hat{\mathbf{G}}(\mathbf{x}, \mathbf{y})](\omega) = \hat{\mathbf{G}}(\mathbf{x}, \mathbf{y}(\omega), \omega) = \mathbf{G}(\mathbf{x}, \xi(\omega))$ . Finally, we let  $\hat{\mathbf{G}}(\mathbf{x}, \mathbf{y}(\omega_s), \omega_s) = \hat{\mathbf{g}}_s(\mathbf{x}, \mathbf{y}_s) = (\mathbf{g}_s^1(\mathbf{x}, \mathbf{y}_s), \mathbf{g}_s^2(\mathbf{x}, \mathbf{y}_s), \dots, \mathbf{g}_s^d(\mathbf{x}, \mathbf{y}_s))^\top$ ,  $s \in S$ , and assume that  $\mathbf{g}_s^i(\mathbf{x}, \mathbf{y}_s)$  is an affine function of  $\mathbf{x}$  and  $\mathbf{y}_s$  such that  $\mathbf{g}_s^i(\mathbf{x}, \mathbf{y}_s) = \bar{\mathbf{g}}_s^i \mathbf{x} + \tilde{\mathbf{g}}_s^i \mathbf{y}_s$  for all  $i \in \{1, \dots, d\}$  and  $s \in S$ . Throughout the paper, we assume that problem (1), if feasible, has a finite objective value, i.e.,  $f(\mathbf{x}), \mathbb{E}(Q(\mathbf{x}, \xi(\omega))) > -\infty$  for all  $\mathbf{x} \in \mathcal{X}$ .

The multivariate risk constraint (1b) ensures that the decision-based random outcome vector  $\mathbf{G}(\mathbf{x}, \xi(\omega))$  is preferable to the specified benchmark  $\mathbf{Z}$  according to the multivariate CVaR relation introduced by [Noyan and Rudolf \(2013\)](#). According to this relation, all scalarized versions of  $\mathbf{G}(\mathbf{x}, \xi(\omega))$  are preferable to the corresponding scalarizations of the benchmark outcome according to the univariate CVaR-based preference relation. We note that using a scalarization-based risk constraint such as (1b) is useful to address ambiguities and inconsistencies in the weight vectors; for detailed discussions we refer to [Hu and Mehrotra \(2012\)](#) and [Liu et al. \(2015\)](#). Moreover, CVaR is a widely applied popular risk measure with several desirable properties; it is a law invariant coherent risk measure ([Artzner et al., 1999](#)), serves as a fundamental building block for other coherent risk measures, and can be used to capture a wide range of risk preferences, including risk-neutral and worst-case approaches. In addition, [Noyan and Rudolf \(2013\)](#) highlight that CVaR-based relations provide flexible, meaningful, and computationally tractable relaxations for overly conservative SSD relations. In a broad sense, CVaR can be viewed as a weighted sum of the least favorable outcomes (those exceeding the  $\alpha$ -quantile – also known as value-at-risk at confidence level  $\alpha$ ).

Using the above notation, the proposed risk-averse two-stage stochastic programming model can alternatively be formulated as follows:

$$\min f(\mathbf{x}) + \sum_{s \in S} p_s Q(\mathbf{x}, \xi(\omega_s)) \quad (3a)$$

$$\text{s.t. } \text{CVaR}_\alpha(\mathbf{c}^\top \hat{\mathbf{G}}(\mathbf{x}, \mathbf{y})) \leq \text{CVaR}_\alpha(\mathbf{c}^\top \mathbf{Z}), \quad \forall \mathbf{c} \in \mathcal{C}, \quad (3b)$$

$$\mathbf{x} \in \mathcal{X}, \quad (3c)$$

$$Q(\mathbf{x}, \xi(\omega_s)) = \mathbf{q}_s^\top \mathbf{y}(\omega_s), \quad \forall s \in S, \quad (3d)$$

$$\mathbf{y}(\omega_s) \in \arg \min_{\mathbf{y} \in \mathcal{Y}(\mathbf{x}, \xi(\omega_s))} \{\mathbf{q}_s^\top \mathbf{y}\}, \quad \forall s \in S, \quad (3e)$$

where for a given  $\hat{\mathbf{c}}_{(l)} \in \mathcal{C}$ ,  $\text{CVaR}_\alpha(\hat{\mathbf{c}}_{(l)}^\top \hat{\mathbf{G}}(\mathbf{x}, \mathbf{y}))$  is calculated as (see [Rockafellar and Uryasev, 2000](#))

$$\min \left\{ \eta_l + \frac{1}{1-\alpha} \sum_{s \in S} p_s w_{sl} : w_{sl} \geq \hat{\mathbf{c}}_{(l)}^\top \hat{\mathbf{G}}(\mathbf{x}, \mathbf{y}(\omega_s), \omega_s) - \eta_l \right. \\ \left. \forall s \in S, w_{sl} \geq 0 \quad \forall s \in S, \eta_l \in \mathbb{R} \right\}. \quad (4)$$

For the risk-neutral version of the proposed two-stage optimization model (1)-(2), which does not feature the risk constraints (1b), it is well-known that the corresponding deterministic equivalent formulation (DEF) can simply be obtained from (3) by dropping the optimality condition on the second-stage decisions, i.e., by replacing (3e) with  $\mathbf{y}(\omega_s) \in \mathcal{Y}(\mathbf{x}, \boldsymbol{\xi}(\omega_s))$ ,  $\forall s \in S$ . DEF provides both the optimal first-stage and second-stage decisions without any issue of the so-called second-stage sub-optimality; this result relies on the interchangeability of the order of expectation and minimization operators associated with the second-stage objective function, and it is known as interchangeability principle (see, e.g., Ruszczyński and Shapiro, 2003). Such a result also exists for the risk-averse two-stage stochastic programming models, where the objective function features a non-decreasing functional of the recourse cost (Takriti and Ahmed, 2004). The next proposition shows that we can similarly drop the optimality condition in (3e) to obtain the DEF of our model even if it features a set of risk constraints.

**PROPOSITION 2.1** *The following deterministic formulation is equivalent to the two-stage model given by (1)-(2) (and consequently, equivalent to (3)):*

$$\min \quad f(\mathbf{x}) + \sum_{s \in S} p_s \mathbf{q}_s^\top \mathbf{y}_s \quad (5a)$$

$$\text{s.t.} \quad \eta_l + \frac{1}{1-\alpha} \sum_{s \in S} p_s w_{sl} \leq \text{CVaR}_\alpha(\hat{\mathbf{c}}_{(l)}^\top \mathbf{Z}), \quad \forall l = 1, \dots, \bar{L}, \quad (5b)$$

$$w_{sl} \geq \hat{\mathbf{c}}_{(l)}^\top \hat{\mathbf{g}}_s(\mathbf{x}, \mathbf{y}_s) - \eta_l, \quad \forall s \in S, l = 1, \dots, \bar{L}, \quad (5c)$$

$$w_{sl} \geq 0, \quad \forall s \in S, l = 1, \dots, \bar{L}, \quad (5d)$$

$$\mathbf{x} \in \mathcal{X}, \quad \boldsymbol{\eta} \in \mathbb{R}^{\bar{L}}, \quad (5e)$$

$$T_s \mathbf{x} + W_s \mathbf{y}_s \geq \mathbf{h}_s, \quad \forall s \in S, \quad (5f)$$

$$\mathbf{y}_s \in \mathbb{R}_+^{n_2}, \quad \forall s \in S, \quad (5g)$$

where  $\bar{L}$  is a finite integer and  $\hat{\mathbf{c}}_{(l)}$ ,  $l = 1, \dots, \bar{L}$ , are the projections of the vertices of the polyhedron  $P = \{(\mathbf{c}, \boldsymbol{\eta}, \mathbf{w}) \in \mathcal{C} \times \mathbb{R} \times \mathbb{R}_+^{|T|} : w_i \geq \eta - \mathbf{c}^\top \mathbf{z}_i, i \in T\}$ .

**PROOF.** First, observe that it is sufficient to consider a finite number of scalarization vectors from set  $\mathcal{C}$ , specifically the projections of the extreme points (given by  $\hat{\mathbf{c}}_{(1)}, \dots, \hat{\mathbf{c}}_{(\bar{L})}$ ) of the above defined polyhedron  $P$  (Noyan and Rudolf, 2013), since  $P$  is only characterized by the given fixed (decision-independent) benchmark outcome vector. Hence, constraints (1b) can be replaced by

$$\text{CVaR}_\alpha(\hat{\mathbf{c}}_{(l)}^\top \mathbf{G}(\mathbf{x}, \boldsymbol{\xi}(\omega))) \leq \text{CVaR}_\alpha(\hat{\mathbf{c}}_{(l)}^\top \mathbf{Z}), \quad \forall l = 1, \dots, \bar{L}. \quad (6)$$

Second, the univariate CVaR-relations in (6) can be represented by linear inequalities of type (5b)-(5d), ignoring the issue of optimality of the  $\mathbf{y}_s$ ,  $s \in S$ , decisions for now. This observation is a simple consequence of the LP representation of CVaR in (4), and the fact that  $\text{CVaR}_\alpha(\mathbf{c}^\top \mathbf{Z})$  is a known constant given a vector  $\mathbf{c}$  and the benchmark  $\mathbf{Z}$ . It is also easy to see that CVaR constraints (5c) can be decomposed over scenarios. This allows us to consider some modified versions of the first- and second-stage problems given in (1)-(2); in particular,  $w_{sl}$  and  $\eta_l$ , for  $s \in S, l = 1, \dots, \bar{L}$  are considered as first-stage decision variables and (1b) is replaced by (5b) and (5d), whereas the inequalities associated with  $s \in S$  in (5c) are added to the constraints of the second-stage problem in (2). Solving (2) ensures that  $\mathbf{y}_s$  featured in (5c) satisfies the optimality condition in (3e). As a result, we obtain the following model which is equivalent to (1)-(2):

$$\min \quad f(\mathbf{x}) + \sum_{s \in S} p_s \tilde{Q}(\mathbf{x}, \boldsymbol{\eta}, \mathbf{w}, \boldsymbol{\xi}(\omega_s)) \quad (7a)$$

$$\text{s.t.} \quad (5b), (5d) - (5e), \quad (7b)$$

where

$$\tilde{Q}(\mathbf{x}, \boldsymbol{\eta}, \mathbf{w}, \boldsymbol{\xi}(\omega_s)) = \min \mathbf{q}_s^\top \mathbf{y} \quad (8a)$$

$$\text{s.t. } w_{sl} \geq \hat{\mathbf{c}}_{(l)}^\top \hat{\mathbf{g}}_s(\mathbf{x}, \mathbf{y}) - \eta_l, \quad \forall l = 1, \dots, \bar{L}, \quad (8b)$$

$$T_s \mathbf{x} + W_s \mathbf{y} \geq \mathbf{h}_s, \quad \mathbf{y} \in \mathbb{R}_+^{n_2}. \quad (8c)$$

Thus, we can reformulate the proposed risk-averse model (1)-(2) as a risk-neutral one (as seen from (7)-(8)). Then, by the well-known interchangeability principle of risk-neutral two-stage models, the assertion directly follows.  $\square$

**REMARK 2.1 Mean-Risk Objective Function.** *The proposed model can be easily extended to the case where the random objective outcomes are compared according to a mean-risk functional representing the decision makers' risk preferences. In particular,  $\mathbb{E}(Q(\mathbf{x}, \boldsymbol{\xi}(\omega)))$  can be replaced by  $\mathbb{E}(Q(\mathbf{x}, \boldsymbol{\xi}(\omega))) + \lambda \rho(Q(\mathbf{x}, \boldsymbol{\xi}(\omega)))$  in (1), where  $\rho$  is a risk functional such as  $\text{CVaR}_\alpha$  and  $\lambda$  is a non-negative trade-off coefficient representing the exchange rate of mean cost for risk. When  $\rho$  is a non-decreasing function and  $\lambda \geq 0$ , the mean-risk function preserves the convexity, and consequently, slight variants of our cutting plane solution algorithms, which we will describe in the next section, are applicable (for similar developments see, e.g., Ahmed, 2006; Noyan, 2012). Thus, our proposed modeling and solution framework is not limited to the risk-neutral objective function.*

In the appendix we provide strong duality results and optimality conditions for an important special case, which generalize the previous results for single-stage multivariate CVaR-constrained problems (Noyan and Rudolf, 2013). Related results for two-stage multivariate SSD-constrained optimization are given in Dentcheva and Wolflagen (2016). Next we present two types of algorithms to solve the computationally challenging two-stage stochastic programming model (3).

**2.1 Delayed Cut Generation for Deterministic Equivalent Formulation** The deterministic equivalent formulation (5) features finite but potentially an exponential number of scalarization vectors, obtained as projections of the vertices of the polyhedron  $P$ . Hence, it is natural to develop a delayed cut generation algorithm, which avoids the impractical vertex enumeration approach required for explicitly constructing the risk constraints in advance. The proposed algorithms rely on solving successive relaxations of the multivariate polyhedral CVaR relation, and they iteratively generate cuts associated with the scalarization vectors for which the risk constraints are relaxed.

In our first cutting plane approach, at an intermediate iteration, the relaxed master problem (RMP) includes the constraint (3b) for a subset of  $\mathcal{C}$ , say  $\{\tilde{\mathbf{c}}_{(1)}, \dots, \tilde{\mathbf{c}}_{(L)}\}$ . Accordingly, RMP takes the form of (5), where  $\bar{L}$  is replaced by  $L$ . Given an optimal solution to the RMP  $(\bar{\mathbf{x}}, \bar{\boldsymbol{\eta}}, \bar{\mathbf{w}})$ , the following separation problem (SP) is solved to identify if there is any violated inequality (3b):

$$\max_{\mathbf{c} \in \mathcal{C}} \text{CVaR}_\alpha(\mathbf{c}^\top \hat{\mathbf{G}}(\bar{\mathbf{x}}, \bar{\mathbf{y}})) - \text{CVaR}_\alpha(\mathbf{c}^\top \mathbf{Z}). \quad (9)$$

Noyan and Rudolf (2013), Küçükyavuz and Noyan (2016), and Liu et al. (2015) provide alternative mixed-integer programming (MIP) formulations for the cut generation problem (9). Because we can treat the DEF as a large-scale single-stage multivariate CVaR-constrained problem, the convergence of the resulting delayed cut generation algorithm follows from Noyan and Rudolf (2013).

**REMARK 2.2** *The cut generation formulations in Noyan and Rudolf (2013), Küçükyavuz and Noyan (2016) and Liu et al. (2015) are based on the opposite convention that larger values of random variables, as well as larger values of risk measures, are considered to be preferable. Those existing MIP formulations should be altered to reflect this difference in convention.*



**2.2 Delayed Cut Generation with Scenario Decomposition** The deterministic equivalent formulation (5) contains a vector of  $\mathbf{y}_s$  variables for each scenario  $s \in S$  and its size grows very fast as the number of scenarios increases. Because the number of scenarios is typically large, it is generally impractical to solve the DEF directly, even without the multivariate stochastic benchmarking constraints. The L-shaped method (Van Slyke and Wets, 1969), which is a Benders-type scenario decomposition algorithm, is arguably the most commonly used computationally viable method to solve the classical (risk-neutral) two-stage stochastic programs. We propose a non-trivial extension of this approach to solve our multivariate CVaR-constrained two-stage model.

In the DEF (5), constraints (5b)–(5c), which capture the CVaR relation (1b) along with the associated auxiliary variables  $\boldsymbol{\eta}$  and  $\mathbf{w}$ , create further coupling of the scenarios in addition to the original coupling constraint (5f). Considering this structure, we handle the  $\mathbf{x}$ ,  $\boldsymbol{\eta}$ , and  $\mathbf{w}$  variables in the first stage, decompose the second-stage problems over scenarios once the first-stage variables are fixed, and solve iteratively the RMP involving only the first-stage decision variables and auxiliary decision variables ( $\theta_s$ ,  $s \in S$ ) for approximating the optimal second-stage objective function values. We obtain the following RMP at an intermediate iteration, where a subset of the scalarization vectors of cardinality  $L$  is generated so far:

$$(\text{MP}^L) \quad \min \quad f(\mathbf{x}) + \sum_{s \in S} p_s \theta_s \quad (10a)$$

$$\text{s.t.} \quad (\mathbf{x}, \boldsymbol{\eta}, \mathbf{w}, \boldsymbol{\theta}) \in \mathcal{O}, \quad (10b)$$

$$(\mathbf{x}, \boldsymbol{\eta}, \mathbf{w}) \in \mathcal{F}, \quad (10c)$$

$$\eta_l + \frac{1}{1 - \alpha} \sum_{s \in S} p_s w_{sl} \leq \text{CVaR}_\alpha(\tilde{\mathbf{c}}_{(l)}^\top \mathbf{Z}), \quad \forall l = 1, \dots, L, \quad (10d)$$

$$(5d) - (5e) \text{ (with } \bar{L} \text{ is replaced by } L\text{)}. \quad (10e)$$

Here  $\mathcal{O}$  is a polyhedral set defined by the constraints (referred to as the *optimality cuts*) that give valid lower bounding approximations of  $\boldsymbol{\theta}$ , and  $\mathcal{F}$  is a polyhedral set defined by the constraints (referred to as the *feasibility cuts*) that represent the conditions for the first-stage decision vectors to yield feasible second-stage problems. In what follows, we give the explicit forms of the inequalities in the sets  $\mathcal{O}$  and  $\mathcal{F}$ .

We start the algorithm with a small subset (could be empty) of the scalarization set  $\mathcal{C}$ . At each iteration, given the subset  $\{\tilde{\mathbf{c}}_{(1)}, \dots, \tilde{\mathbf{c}}_{(L)}\}$ , we first solve the relaxed master problem (MP<sup>L</sup>) to obtain an optimal solution  $\bar{\mathbf{v}} := (\bar{\mathbf{x}}, \bar{\boldsymbol{\eta}}, \bar{\mathbf{w}}, \bar{\boldsymbol{\theta}})$ . Using this solution of the RMP, we solve the second-stage subproblem (PS<sub>s</sub><sup>L</sup>) for each scenario  $s \in S$  given by

$$(\text{PS}_s^L) \quad \bar{Q}_s^L := \min \quad \mathbf{q}_s^\top \mathbf{y} \quad (11a)$$

$$\text{s.t.} \quad -\tilde{\mathbf{c}}_{(l)}^\top \tilde{\mathbf{g}}_s \mathbf{y} \geq \tilde{\mathbf{c}}_{(l)}^\top \tilde{\mathbf{g}}_s \bar{\mathbf{x}} - \bar{\eta}_l - \bar{w}_{sl}, \quad \forall l = 1, \dots, L, \quad (11b)$$

$$W_s \mathbf{y} \geq \mathbf{h}_s - T_s \bar{\mathbf{x}}, \quad (11c)$$

$$\mathbf{y} \in \mathbb{R}_+^{n_2}, \quad (11d)$$

where constraint (11b) is equivalent to constraint (5c) after partial substitution of the values of the first-stage variables with  $\bar{\mathbf{x}}$ ,  $\bar{\boldsymbol{\eta}}$  and  $\bar{\mathbf{w}}$ . Let  $\beta_{sl}$ ,  $l = 1, \dots, L$ , and  $\boldsymbol{\gamma}_s$  for  $s \in S$ , be the dual variables associated with constraints (11b) and (11c), respectively. If the primal subproblem (PS<sub>s</sub><sup>L</sup>) is infeasible for some  $s \in S$ , then let  $(\boldsymbol{\gamma}_s, \boldsymbol{\beta}_s)$  provide an extreme ray (direction of unboundedness) of the corresponding dual feasible region and add the following Benders feasibility cut to the set  $\mathcal{F}$ :

$$0 \geq \boldsymbol{\gamma}_s^\top (\mathbf{h}_s - T_s \mathbf{x}) + \sum_{l=1}^L \beta_{sl}^\top \left( \tilde{\mathbf{c}}_{(l)}^\top \tilde{\mathbf{g}}_s \mathbf{x} - \eta_l - w_{sl} \right). \quad (12)$$

If the primal subproblem is feasible and the current estimate of the optimal second-stage objective value ( $\bar{\theta}_s$ ) is less than the actual optimal second-stage objective value ( $\bar{Q}_s^L$ ), then let  $(\gamma_s, \beta_s)$  provide an optimal dual extreme point and add the following Benders optimality cut to the set  $\mathcal{O}$ :

$$\theta_s \geq \gamma_s^\top (\mathbf{h}_s - T_s \mathbf{x}) + \sum_{l=1}^L \beta_{sl}^\top \left( \tilde{\mathbf{c}}_{(l)}^\top \bar{\mathbf{g}}_s \mathbf{x} - \eta_l - w_{sl} \right). \quad (13)$$

Now suppose that all primal subproblems are feasible with the finite objective values  $\bar{Q}_s^L, s \in S$  (recall our assumption that the original problem is not unbounded). In this case, given  $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ , we solve the separation problem (9) to obtain an optimal solution  $\mathbf{c}^*$ . If the corresponding optimal objective value is non-positive, then there is no violation in the CVaR constraint (3b). Furthermore, recall that in this case, all second-stage problems are feasible as well. In other words, the current first-stage solution is feasible for the original problem, so we update the upper bound on the optimal objective value (denoted by  $u$ ). On the other hand, if the optimal objective value of the SP is positive, then  $\mathbf{c}^* \in \mathcal{C}$  creating the most violation in constraint (11b) is added as the  $L + 1$ th scalarization vector both to the RMP and the second-stage subproblems with  $L \leftarrow L + 1$ . We repeat these iterations until either an infeasible RMP is detected, or the RMP objective function value is within the given tolerance  $\epsilon$  of the upper bound  $u$ . The optimality tolerance is specified as a very small positive number such as  $10^{-10}$ . The pseudo-code of the proposed decomposition algorithm is provided in Algorithm 1. Note that the feasible regions of the dual subproblems change depending on the scalarization vectors included in the formulations at that iteration. In addition, the RMP also grows both in terms of the number of variables and constraints as we add more scalarization vectors to its formulation due to the addition of the  $\eta_l$  and  $\mathbf{w}_l$  variables for each scalarization vector  $\tilde{\mathbf{c}}_{(l)}$ . As a result, our proposed algorithm can be seen as a *delayed column and cut generation algorithm*.

**PROPOSITION 2.2** *Suppose that the relaxed master problem (MP<sup>L</sup>) is given by (10), the second-stage subproblems (PS<sub>s</sub><sup>L</sup>) are given by (11), the separation problem is given by (9), the feasibility and optimality cuts, respectively, are given by (12) and (13), and the first-stage solution vector is given by  $\bar{\mathbf{v}} = (\bar{\mathbf{x}}, \bar{\boldsymbol{\eta}}, \bar{\mathbf{w}}, \bar{\boldsymbol{\theta}})$ . Then, Algorithm 1 provides either an optimal solution to problem (1) or a proof of infeasibility in finitely many iterations.*

**PROOF.** First, observe that the optimality cuts (13) generated from subproblems (PS<sub>s</sub><sup>L</sup>) for  $s \in S$  and the subset  $\{\tilde{\mathbf{c}}_{(1)}, \dots, \tilde{\mathbf{c}}_{(L)}\}$  of  $\mathcal{C}$  are valid, because subproblems (PS<sub>s</sub><sup>L</sup>) are relaxations of the original subproblems (8) given by (PS<sub>s</sub><sup>L̄</sup>), and hence the duals of the relaxed subproblems provide valid lower bounds on the optimal second-stage objective values. Similarly, the feasibility cuts (12) obtained from the relaxed subproblems are valid for the original subproblems (8). At an intermediate iteration, if there is a violation in the CVaR constraint (3b), then the exact solution of the SP guarantees that a corresponding violating scalarization vector is identified and the associated inequalities and variables are added to the RMP. Furthermore, Noyan and Rudolf (2013) provide a procedure to guarantee that the  $\mathbf{c}^*$  vector generated from the SP is a vertex solution as defined in Proposition 2.1. Next, recall that the number of scalarization vectors of interest,  $\bar{L}$ , is finite. Hence, in the worst case, in a finite number of iterations, all  $\bar{L}$  scalarization vectors will be added, and the second-stage problems become exact. At this point, the convergence of the algorithm follows directly from the convergence of the classical L-shaped method (Van Slyke and Wets, 1969).  $\square$

**3. Two-Stage Optimization with a Multivariate Stochastic Ordering** The solution algorithms proposed in Section 2 can also be applied to the two-stage stochastic programs with the multivariate preference constraints based on increasing convex order (ICO). The counterpart of ICO in the opposite convention, where larger values of random variables are preferred, is referred to as the SSD. Similar to



---

**Algorithm 1:** Decomposition algorithm for the multivariate risk-constrained two-stage optimization

---

```

1  Given  $L$  initial scalarization vectors  $(\{\tilde{\mathbf{c}}_{(i)}\}_1^L \subset \mathcal{C})$  and a small tolerance parameter  $\epsilon > 0$ , set
    $converge = false$  and  $u \leftarrow \infty$ ;
2  while  $converge = false$  do
3      Solve  $(MP^L)$  and get the optimal solution  $\bar{\mathbf{v}}$  and objective value  $F(MP^L)$ ;
4      if  $(MP^L)$  is infeasible then
5           $converge = true$ . The original risk-averse problem is infeasible;
6      else
7          if  $(1 - \epsilon)u \leq F(MP^L) \leq (1 + \epsilon)u$  then
8              Set  $converge = true$ . The optimal solution is  $\bar{\mathbf{v}}$ ;
9          else
10             for each scenario  $s \in S$  do
11                 Solve the primal subproblem  $(PS_s^L)$  and get the optimal objective value  $\bar{Q}_s^L$ ;
12                 if  $(PS_s^L)$  is infeasible then
13                     Obtain an extreme ray of the dual feasible region associated with  $(PS_s^L)$ ,
14                     given by  $(\boldsymbol{\gamma}_s, \boldsymbol{\beta}_s)$ , and add the corresponding feasibility cut to  $\mathcal{F}$ ;
15                 else
16                     Obtain an optimal extreme point of the dual of  $(PS_s^L)$ , given by  $(\boldsymbol{\gamma}_s, \boldsymbol{\beta}_s)$ , and
17                     if  $\bar{\theta}_s < Q_s^L$ , then add the corresponding optimality cut to  $\mathcal{O}$ ;
18             if  $(PS_s^L)$  is feasible for all  $s \in S$  then
19                 Solve the separation problem (SP);
20                 Obtain the optimal solution  $\mathbf{c}^*$  and optimal objective value  $v(\mathbf{c}^*)$ ;
21                 if  $v(\mathbf{c}^*) \leq 0$  then
22                     Current solution is feasible, set  $u \leftarrow f(\bar{\mathbf{x}}) + \sum_{s \in S} p_s \bar{Q}_s^L$ ;
23                 else
24                     Set  $L \leftarrow L + 1$  and  $\tilde{\mathbf{c}}_{(L)} \leftarrow \mathbf{c}^*$ ;

```

---

the CVaR case, the univariate ICO relation can be extended to the multivariate case by considering a family of linear scalarization functions; more specifically, a random vector  $\mathbf{X}$  is said to dominate another random vector  $\mathbf{Z}$  with respect to the ICO and the scalarization set  $\mathcal{C}$  if

$$\mathbb{E} \left( [\mathbf{c}^\top \mathbf{X} - \eta]_+ \right) \leq \mathbb{E} \left( [\mathbf{c}^\top \mathbf{Z} - \eta]_+ \right), \quad \forall \eta \in \mathbb{R}, \mathbf{c} \in \mathcal{C}. \quad (14)$$

The risk-averse two-stage optimization model of Dentcheva and Wolfhagen (2016) utilizes this multivariate stochastic preference relation to compare the decision-based random outcome vector  $\hat{\mathbf{G}}(\mathbf{x}, \mathbf{y})$  with the benchmark outcome vector  $\mathbf{Z}$ . In particular, the two-stage stochastic programming model with the multivariate ICO constraints takes the form of (3), where the multivariate CVaR relation (3b) is replaced by (14) with  $\mathbf{X} = \hat{\mathbf{G}}(\mathbf{x}, \mathbf{y})$ . Assuming again that  $\mathbf{Z}$  has finitely many realizations  $\mathbf{z}_1, \dots, \mathbf{z}_{|T|}$  with associated probabilities  $\bar{p}_i$ ,  $i \in T$ , the inequality (14) can be equivalently stated as

$$\mathbb{E} \left( [\mathbf{c}^\top \mathbf{X} - \mathbf{c}^\top \mathbf{z}_i]_+ \right) \leq \mathbb{E} \left( [\mathbf{c}^\top \mathbf{Z} - \mathbf{c}^\top \mathbf{z}_i]_+ \right), \quad \forall i \in T, \mathbf{c} \in \mathcal{C}. \quad (15)$$

Relation (15) directly follows from (14) due to the well-known result (Dentcheva and Ruszczyński, 2003) that for finite probability spaces the univariate SSD relation remains valid if the corresponding expected shortfall constraints (in our convention, constraints (14) featuring the expected excess values for a particular  $\mathbf{c}$ ) only required for the realizations of the benchmark random variable (instead of all  $\eta \in \mathbb{R}$ ).

For finite probability spaces, [Homem-de-Mello and Mehrotra \(2009\)](#) show that it is sufficient to consider a finite number of scalarization vectors  $\{\hat{\mathbf{c}}_{(1)}, \dots, \hat{\mathbf{c}}_{(\bar{L})}\}$  in relation (15), specifically the projections of the extreme points of polyhedra  $P_t$ ,  $t \in T$ , defined for each  $t \in T$  as  $P_t = \{w_i \geq \mathbf{c}^\top \mathbf{z}_i - \mathbf{c}^\top \mathbf{z}_t, i \in T, \mathbf{c} \in \mathcal{C}, \mathbf{w} \in \mathbb{R}_+^{|T|}\}$ . For an alternative finite representation of (15), which is based on the vertices of the polyhedron  $P$  defined in Proposition 2.1, we refer to [Noyan and Rudolf \(2013\)](#). The fact that these finite representations of (15) are only characterized by the given benchmark vector allows us to develop computationally tractable DEFs of the optimization models featuring the multivariate ICO relation as a benchmarking constraint. In particular, for the two-stage multivariate ICO-constrained problem we can obtain the following DEF with finitely many constraints:

$$\min f(\mathbf{x}) + \sum_{s \in S} p_s \mathbf{q}_s^\top \mathbf{y}_s \quad (16a)$$

$$\text{s.t. } \sum_{s \in S} p_s w_{slt} \leq \sum_{i \in T} \tilde{p}_i \left[ \hat{\mathbf{c}}_{(l)}^\top \mathbf{z}_i - \hat{\mathbf{c}}_{(l)}^\top \mathbf{z}_t \right]_+, \quad \forall t \in T, l = 1, \dots, \bar{L}, \quad (16b)$$

$$w_{slt} \geq \hat{\mathbf{c}}_{(l)}^\top \hat{\mathbf{g}}_s(\mathbf{x}, \mathbf{y}_s) - \hat{\mathbf{c}}_{(l)}^\top \mathbf{z}_t, \quad \forall s \in S, t \in T, l = 1, \dots, \bar{L}, \quad (16c)$$

$$w_{slt} \geq 0, \quad \forall s \in S, t \in T, l = 1, \dots, \bar{L}, \quad (16d)$$

$$\mathbf{x} \in \mathcal{X}, \quad (5f) - (5g). \quad (16e)$$

Although the number of inequalities (16b)–(16c) is finite, it is possibly exponential. Hence, as in the CVaR-constrained optimization, the ICO constraints can be added as needed using a delayed cut generation approach. As in Section 2.1, we start with solving the DEF with a subset of constraints (16b)–(16c) for  $l = 1, \dots, L$ . For a given solution  $\bar{\mathbf{v}} := (\bar{\mathbf{x}}, \bar{\mathbf{w}}, \bar{\mathbf{y}})$  to this relaxed problem, we check if there is a violation in (15) (with  $\mathbf{X} = \hat{\mathbf{G}}(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ ). Note that constraints (15) are defined for each scalarization vector and for each realization of  $\mathbf{Z}$ , as such possible violations in constraints (15) should be checked for all  $\mathbf{z}_t$ ,  $t \in T$ . In line with these discussions, given  $(\bar{\mathbf{x}}, \bar{\mathbf{w}}, \bar{\mathbf{y}})$ , the separation problem corresponding to the  $t^{\text{th}}$  realization of  $\mathbf{Z}$  becomes

$$\min_{\mathbf{c} \in \mathcal{C}} \mathbb{E} \left( \left[ \mathbf{c}^\top \mathbf{Z} - \mathbf{c}^\top \mathbf{z}_t \right]_+ \right) - \mathbb{E} \left( \left[ \mathbf{c}^\top \hat{\mathbf{G}}(\bar{\mathbf{x}}, \bar{\mathbf{y}}) - \mathbf{c}^\top \mathbf{z}_t \right]_+ \right). \quad (17)$$

This cut generation problem can be solved using the adaptations of the MIP formulations provided in [Homem-de-Mello and Mehrotra \(2009\)](#), [Küçükyavuz and Noyan \(2016\)](#), and [Noyan and Rudolf \(2016\)](#) for the opposite convention that larger outcomes are preferred.

As in the CVaR case, the structure of the ICO constraints can be exploited to further decompose the formulation (16) over scenarios. Observe that constraints (16b) and (5f) couple the scenarios together. However, if we handle the original first-stage variables  $\mathbf{x}$  and the auxiliary  $\mathbf{w}$  variables in the first stage, then the second-stage problems decompose over scenarios once the first-stage variables are fixed. Therefore, we can apply Algorithm 1 with the updated definitions of the problems  $(\text{MP}^L)$ ,  $(\text{PS}_s^L)$ , and  $(\text{SP})$ , the feasibility and optimality cuts, and the first-stage solution vector  $\mathbf{v}$ , which we describe next.

The RMP at an intermediate iteration, where a subset of the scalarization vectors of cardinality  $L$  is generated so far, is formulated as

$$(\text{MP}^L) \min f(\mathbf{x}) + \sum_{s \in S} p_s \theta_s \quad (18a)$$

$$\text{s.t. } (\mathbf{x}, \mathbf{w}, \boldsymbol{\theta}) \in \mathcal{O}, \quad (18b)$$

$$(\mathbf{x}, \mathbf{w}) \in \mathcal{F}, \quad (18c)$$

$$\mathbf{x} \in \mathcal{X}, \quad (16b), (16d) \text{ (with } \bar{L} \text{ is replaced by } L). \quad (18d)$$

Given a first-stage solution vector  $\bar{\mathbf{v}} = (\bar{\mathbf{x}}, \bar{\mathbf{w}}, \bar{\boldsymbol{\theta}})$  of RMP and the  $L$  scalarization vectors, the second-stage subproblem under  $s \in S$  takes the form of

$$(\text{PS}_s^L) \quad \bar{Q}_s^L := \min \mathbf{q}_s^\top \mathbf{y} \quad (19a)$$

$$\text{s.t.} \quad -\tilde{\mathbf{c}}_{(l)}^\top \tilde{\mathbf{g}}_s \mathbf{y} \geq \tilde{\mathbf{c}}_{(l)}^\top \tilde{\mathbf{g}}_s \bar{\mathbf{x}} - \tilde{\mathbf{c}}_{(l)}^\top \mathbf{z}_t - \bar{w}_{slt}, \quad \forall t \in T, l = 1, \dots, L, \quad (19b)$$

$$(11c) - (11d). \quad (19c)$$

Here, constraint (19b) is equivalent to constraint (16c) with partial substitution of the values of the first-stage variables  $\bar{\mathbf{x}}$  and  $\bar{\mathbf{w}}$ . Denoting the dual vectors associated with constraints (19b) and (11c) by  $\beta_{stl}, t \in T, l = 1, \dots, L$ , and  $\gamma_s, s \in S$ , the feasibility and optimality cuts are given by

$$0 \geq \gamma_s(\mathbf{h}_s - T_s \mathbf{x}) + \sum_{t \in T} \sum_{l=1}^L \beta_{stl} \left( \tilde{\mathbf{c}}_{(l)}^\top \tilde{\mathbf{g}}_s \mathbf{x} - \tilde{\mathbf{c}}_{(l)}^\top \mathbf{z}_t - w_{slt} \right) \quad \text{and} \quad (20)$$

$$\theta_s \geq \gamma_s(\mathbf{h}_s - T_s \mathbf{x}) + \sum_{t \in T} \sum_{l=1}^L \beta_{stl} \left( \tilde{\mathbf{c}}_{(l)}^\top \tilde{\mathbf{g}}_s \mathbf{x} - \tilde{\mathbf{c}}_{(l)}^\top \mathbf{z}_t - w_{slt} \right). \quad (21)$$

**PROPOSITION 3.1** *Suppose that the relaxed master problem (MP<sup>L</sup>) is given by (18), the second-stage subproblems (PS<sub>s</sub><sup>L</sup>) are given by (19), the separation problem is given by (17), the feasibility and optimality cuts, respectively, are given by (20) and (21), and the first-stage solution vector is given by  $\bar{\mathbf{v}} = (\bar{\mathbf{x}}, \bar{\mathbf{w}}, \bar{\theta})$ . Then, Algorithm 1 provides either an optimal solution to problem (16) or a proof of infeasibility in finitely many iterations.*

**PROOF.** The proof is similar to that of Proposition 2.2. □

Finally, we highlight four major advantages of our proposed solution framework over the existing methods (Dentcheva and Wolflagen, 2016). First, our method reformulates the risk-averse model of interest as a risk-neutral two-stage stochastic program without using the Lagrangian relaxation of the multivariate ICO constraints, and consequently, manages to solve the problem within a single Benders decomposition framework. It maintains a single master problem to solve successively redefined (based on the iteratively enlarged subset of the scalarization vectors) risk-neutral relaxations of the problem. Recall that using a subset of  $\mathcal{C}$  provides a relaxation of the problem. These risk-neutral formulations are constructed by using the decomposable structure of the ICO constraints over scenarios. Dentcheva and Wolflagen (2016) also construct successive relaxations by considering a subset of  $\mathcal{C}$ ; however, they obtain approximate risk-neutral relaxations by using the Lagrangian relaxation of the multivariate ICO constraints. In particular, at each iteration of their algorithms, a risk-neutral approximation with an updated Lagrangian objective function is solved (to calculate the dual function given a set of Lagrangian multipliers) using decomposition methods such as Benders decomposition. The need for solving a separate risk-neutral two-stage model at each iteration, in addition to the well-known computational challenges in solving the non-differentiable Lagrangian dual problem, could impose significant computational difficulties. Second, the Lagrangian-based existing algorithms are not exact even for the finite probability spaces; they are shown to finitely converge to an  $\varepsilon$ -feasible (with respect to the ICO constraint)  $\varepsilon$ -optimal solution both for continuous and finite probability spaces. In contrast, our finitely-convergent algorithm provides an exact (optimal and feasible) solution for finite probability spaces. Third, the existing algorithms do not provide valid upper bounds on the optimal objective value at intermediate stages, because the Lagrangian dual at an iteration is constructed using the information on a subset of  $\mathcal{C}$ . In contrast, Algorithm 1 provides valid upper bounds (in line 20) at intermediate stages. Fourth, our method applies to problems that contain discrete variables in the first stage. In contrast, the Lagrangian-based methods cannot guarantee the integer feasibility (see the primal solution recovery step of the algorithms in Dentcheva and Wolflagen, 2016). In summary, we contribute to the literature by providing a new computationally tractable and exact solution algorithm for the multivariate ICO-constrained two-stage models with integer variables in the first stage under finite probability spaces.

**4. A Stochastic Optimization Model for Pre-disaster Relief Network Design** An effective and sound pre-disaster relief network design calls for modeling the risk associated with the high level of uncertainty inherent in rarely occurring disaster events (see, e.g., [Noyan, 2012](#)) and considering multiple and possibly conflicting performance criteria such as responsiveness and equity (see, e.g., [Vitoriano et al., 2011](#); [Huang et al., 2012](#); [Gutjahr and Nolz, 2016](#)). Moreover, a two-stage decision making framework is beneficial in this context, since the pre-positioning design decisions must be made before a disaster strikes, whereas the relief distribution decisions should be made in the post-disaster stage. Motivated by the significance of developing such effective pre-disaster plans (see, e.g., [Balcik and Beamon, 2008](#); [Salmerón and Apte, 2010](#)), we apply our proposed approach to a stochastic pre-disaster relief network design problem. In this section, we first briefly review the relevant literature, and then describe the problem setting and present the corresponding mathematical programming formulations.

There is a rich and growing literature on developing stochastic programming models for humanitarian logistics (e.g., [Çelik et al., 2012](#); [Liberatore et al., 2013](#)). The majority of the existing studies focus on pre-disaster relief network design and develop risk-neutral two-stage stochastic programs (e.g., [Balcik and Beamon, 2008](#); [Rawls and Turnquist, 2010](#); [Salmerón and Apte, 2010](#); [Döyen et al., 2012](#)). However, this extensive literature includes only a few studies (e.g., [Rawls and Turnquist, 2011](#); [Noyan, 2012](#); [Hong et al., 2015](#); [Elçi, 2016](#)) that provide risk-averse stochastic models, and lacks models with multivariate risk constraints. Here, we mainly discuss the studies that are most closely related to ours. [Rawls and Turnquist \(2010\)](#) develop a risk-neutral two-stage stochastic programming model; in the first-stage it determines the cost-wise optimal decisions on locations and capacities of the response facilities, and the inventory levels of relief supplies under uncertainty in demand, damage levels of pre-stocked supplies, and transportation network conditions. Their second-stage problem is formulated as a classical network flow model, which involves detailed distribution decisions representing the flow of relief supplies on each arc. [Noyan \(2012\)](#) obtains a risk-averse version of this model by incorporating CVaR as the risk measure on the total cost in addition to its expectation. There also exist chance-constrained variants ([Rawls and Turnquist, 2011](#); [Hong et al., 2015](#)). A recent study [Elçi \(2016\)](#) proposes a more elaborate risk-averse extension, which features a mean-risk objective function on the random total cost with CVaR being the risk measure (as in [Noyan, 2012](#)), and enforces a joint chance constraint on the feasibility of the second-stage problem (as in [Hong et al., 2015](#)). In contrast to the other variants, the model of [Elçi \(2016\)](#) relies on an alternative formulation of the second-stage problem, which focuses on assigning the demand points to the facilities instead of determining the detailed arc-flow decisions. In our study, we follow this practically meaningful assignment-based modeling approach for the second-stage problem, and develop a new risk-averse variant of the widely-cited model of [Rawls and Turnquist \(2010\)](#).

**Problem Description and Mathematical Formulations.** We extend the model of [Rawls and Turnquist \(2011\)](#) by incorporating the multivariate CVaR relation into the first-stage to evaluate the decisions based on additional multiple stochastic criteria. In particular, we consider the following additional stochastic measures: the maximum proportion of unsatisfied demand and a responsiveness measure based on the total delivery amount-based average travel time. With the term responsiveness we refer to the expeditiousness of the response. A large variety of criteria have been employed in humanitarian logistics and they can be grouped in three main categories (see, e.g., [Huang et al., 2012](#); [Gutjahr and Nolz, 2016](#)): *efficiency* (related to cost), *efficacy or effectiveness* (related to providing quick and sufficient distribution) and *equity* (related to providing a fair service). According to this classification, our model addresses the issues about efficiency and efficacy of the relief operations using a weighted-sum based objective, which aggregates the costs of opening facilities, demand shortages, and purchasing and shipping the relief supplies. In addition, it addresses the issues about responsiveness and equity (in terms of supply allocation) via the multivariate CVaR constraints. The CVaR-based relation is preferred,

since it is a natural relaxation of its SSD-counterpart. Multivariate SSD constraints are typically very demanding and they can be excessively hard to satisfy in practice.

We model the relief distribution system using a network, where  $I$  and  $J$  denote the sets of nodes representing the demand and candidate facility locations, respectively; we assume without loss of generality that  $J \subseteq I$ . The set  $L$  denotes the multiple facility types, a facility of type  $l$  has a given capacity level  $K_l$ . Setting up a facility of type  $l$  at node  $j$  has a fixed cost  $F_{jl}$  and a unit variable acquisition cost  $a_j$ . We consider a single commodity (as in, e.g., [Hong et al., 2015](#)), since the relief items can be supplied as bundles. The demand values, travel times, demand shortage penalty costs, and damage levels of supplies are assumed to be random and their realizations are represented by a finite set of scenarios  $S$  with probabilities  $p_s$ ,  $s \in S$ . We next introduce notation for the realizations of these random parameters under scenario  $s$ : demand at node  $i$  is  $d_i^s$ , travel time from node  $j$  to node  $i$  is  $\nu_{ji}^s$ , unit shortage penalty cost is  $\pi^s$ , unit shipment cost from node  $j$  to node  $i$  is  $c_{ji}^s$ , proportion of undamaged supplies at node  $j$  is  $\gamma_j^s$ . In contrast to [Rawls and Turnquist \(2010\)](#), each demand location can be served only by a facility node, which satisfies a responsiveness requirement. In particular, we consider a common upper bound on travel times to construct the scenario-dependent coverage sets:  $N_i^s$  and  $M_j^s$ , respectively, denote the set of facility nodes that can cover demand node  $i$  and the set of demand nodes that can be served by facility node  $j$  under scenario  $s$ . Enforcing a common threshold serves the objective of providing equitable service in terms of response times.

In our two-stage decision making framework, the first-stage decisions are represented using the following notation:  $x_{jl} = 1$  if a facility of size category  $l \in L$  is located at node  $j \in J$ , and  $x_{jl} = 0$  otherwise;  $R_j$  denotes the amount of supplies pre-located at node  $j \in J$ . The notation for the second-stage decisions under scenario  $s \in S$  is as follows:  $y_{ji}^s$  denotes the amount of supplies delivered to demand node  $i \in I$  from node  $j \in J$ ;  $u_i^s$  designates the amount of the unsatisfied demand at node  $i \in I$ ;  $u_{\max}^s$  is the maximum proportion of unsatisfied demand (maximum is calculated over the demand nodes), i.e.,  $u_{\max}^s = \max_{i \in I} u_i^s/d_i^s$ . We next present our risk-averse two-stage stochastic programming model, where the first-stage problem is formulated as follows:

$$\min \sum_{j \in J} \sum_{l \in L} F_{jl} x_{jl} + \sum_{j \in J} a_j R_j + \sum_{s \in S} p_s Q(\mathbf{x}, \mathbf{R}, \boldsymbol{\xi}(\omega_s)) \quad (22a)$$

$$\text{s.t. } \text{CVaR}_\alpha(\mathbf{c}^\top \mathbf{G}(\mathbf{x}, \mathbf{R}, \boldsymbol{\xi}(\omega))) \leq \text{CVaR}_\alpha(\mathbf{c}^\top \mathbf{Z}), \quad \forall \mathbf{c} \in \mathcal{C}, \quad (22b)$$

$$\sum_{l \in L} x_{jl} \leq 1, \quad \forall j \in J, \quad (22c)$$

$$R_j \leq \sum_{l \in L} K_l x_{jl}, \quad \forall j \in J, \quad (22d)$$

$$x_{jl} \in \{0, 1\}, \quad \forall j \in J, l \in L, \quad R_j \geq 0, \quad \forall j \in J. \quad (22e)$$

Here,  $\boldsymbol{\xi}(\omega_s) = (\mathbf{d}^s, \boldsymbol{\gamma}^s, \mathbf{c}^s, \boldsymbol{\nu}^s, \pi^s)$  is the vector of the random parameter realizations under scenario  $s \in S$ . Given the first-stage decision vectors  $\mathbf{x}$  and  $\mathbf{R}$ , the optimal second-stage objective value under scenario  $s$ ,  $Q(\mathbf{x}, \mathbf{R}, \boldsymbol{\xi}(\omega_s))$ , is calculated by solving the following second-stage problem under scenario  $s$ :

$$\min \sum_{j \in J} \sum_{i \in M_j^s \setminus \{j\}} c_{ji}^s y_{ji}^s + \sum_{i \in I} \pi^s u_i^s \quad (23a)$$

$$\text{s.t. } u_i^s \geq d_i^s - \sum_{j \in N_i^s} y_{ji}^s, \quad \forall i \in I \setminus J, \quad (23b)$$

$$u_j^s \geq d_j^s + \sum_{i \in M_j^s \setminus \{j\}} y_{ji}^s - (R_j \gamma_j^s + \sum_{i \in N_j^s \setminus \{j\}} y_{ij}^s), \quad \forall j \in J, \quad (23c)$$

$$\sum_{i \in M_j^s \setminus \{j\}} y_{ji}^s \leq R_j \gamma_j^s, \quad \forall j \in J, \quad (23d)$$

$$\sum_{i \in N_j^s \setminus \{j\}} y_{ij}^s \leq (1 - \sum_{l \in L} x_{jl}) d_j^s, \quad \forall j \in J, \quad (23e)$$

$$u_{\max}^s \geq \frac{u_i^s}{d_i^s}, \quad \forall i \in I, \quad (23f)$$

$$y_{ji}^s \geq 0 \quad \forall j \in J, i \in M_j^s \setminus \{j\}, \quad (23g)$$

$$u_i^s \geq 0 \quad \forall i \in I, \quad u_{\max}^s \geq 0. \quad (23h)$$

The objective, defined by (22a) and (23a), is to minimize the expected total cost of opening facilities, purchasing and shipping the relief supplies, and demand shortages. As discussed in Section 2, constraint (22b) ensures that the decision-based random outcome vector  $\mathbf{G}(\mathbf{x}, \mathbf{R}, \boldsymbol{\xi}(\omega))$  is preferable to the benchmark outcome  $\mathbf{Z}$  according to the multivariate polyhedral CVaR relation. Here  $\mathbf{G}(\mathbf{x}, \mathbf{R}, \boldsymbol{\xi}(\omega_s)) = \hat{\mathbf{g}}_s(\mathbf{x}, \mathbf{R}, \mathbf{y}^s, \mathbf{u}^s, u_{\max}^s)$  is a two-dimensional vector, the components of which are given by

$$\mathbf{g}_s^1(\mathbf{x}, \mathbf{R}, \mathbf{y}^s, \mathbf{u}^s, u_{\max}^s) = u_{\max}^s, \quad (24)$$

$$\mathbf{g}_s^2(\mathbf{x}, \mathbf{R}, \mathbf{y}^s, \mathbf{u}^s, u_{\max}^s) = \frac{\sum_{i \in I} \text{ATS}_i^s}{\sum_{i \in I} \max_{j \in N_i^s \setminus \{i\}} \nu_{ij}^s}, \text{ where } \text{ATS}_i^s = \sum_{j \in N_i^s \setminus \{i\}} \frac{\nu_{ij}^s y_{ji}^s}{d_i^s}. \quad (25)$$

The maximum proportion of unsatisfied demand (24) is crucial for equity in terms of supply allocation, whereas the second measure (25), the total delivery amount-based average travel time score (ATS), is related to responsiveness. The delivery-based unit average travel time to demand node  $i$  could ideally be calculated by replacing  $d_i^s$  with the corresponding exact delivery amount (i.e.,  $y_{ji}^s$ ) in (25); we prefer not to use this calculation to avoid nonlinear terms. Moreover, (25) relies on a scaling with an upper bound ( $\max_{j \in N_i^s \setminus \{i\}} \nu_{ij}^s$ ) of  $\text{ATS}_i^s$ ; this scaling ensures that both measures take values between 0 and 1 and prevents biased solutions due to the differences in the magnitude of the outcomes. We also note that considering a set of scalarization vectors in (22b) is essential to deal with the ambiguities and inconsistencies in the weight vectors. Eliciting the relative weights of even a single decision maker/expert can be challenging (for a related review, see, e.g., Liu et al., 2015), and is further exacerbated if multiple experts are involved as in the case of humanitarian logistics applications.

We next elaborate on the remaining constraints of the proposed model. Constraints (22c) ensure that at most one facility is opened at any node  $j \in J$ . Constraints (22d) guarantee that the inventory levels at open facilities do not exceed the corresponding capacity limits and there is a facility located at node  $j$  if its inventory level is positive. Constraints (23b)-(23e) correspond to a restricted version of the classical network flow formulation, which provides a more structured flow considering the coverage issues. More specifically, these constraints enforce that the nodes without any pre-stocked inventory receive service directly from the open facilities that are sufficiently close (according to the upper bound on travel times). Constraints (23d) ensure that the amount delivered from a facility node does not exceed its available inventory level and outgoing flow is not possible if there is no facility located. Constraints (23e) assure that there is no delivery to a node if a facility is located at that node, and the amount of delivery is limited by its demand, otherwise. Constraints (23f) calculate the maximum proportion of demand shortage.

We develop computationally effective solution methods for the model (22)-(23) by applying our cut generation-based algorithms presented in Section 2. Observing that the original second-stage problem (23) is always feasible, it is interesting to note that our scenario decomposition-based algorithm may need to generate the Benders feasibility cuts (12). This is the case since the algorithm solves iteratively modified versions of (23) with the constraints of type (11b), which may lead to infeasibility in a second-stage subproblem.



**5. Computational Study** In this section, we illustrate the computational effectiveness of the two types of algorithms proposed in Section 2 by implementing them to solve the two-stage model (22)-(23) under varying parameter settings. We first describe how the test problem instances are generated and then provide numerical results.

We used the data sets of Elçi (2016), which are based on a case study concerning the threat of hurricanes in the Southeastern part of the United States (Rawls and Turnquist, 2010); in this case study, the region is represented by a network with 30 nodes and 55 arcs, and only a single set of hurricane scenarios with a cardinality of 51 is considered. In order to generate instances with a larger number of scenarios, Hong et al. (2015) propose a scenario generation method which takes into account the dependency structures inherent in disaster relief networks. Their approach randomly identifies the link degradation and node damage levels depending on the proximity to the center of the disaster, and generates the realizations of the random input parameters accordingly. Elçi (2016) follows this elaborate scenario generation approach, but makes the necessary modifications according to their assignment-based formulation of the second-stage problem. More specifically, the author ignores 55 links of the original network and introduces the links according to the coverage-based responsiveness requirement. In our computational study, we consider the most dense (fully-connected) network structure to distribute water (in units of 1000 gallons). We refer the reader to Hong et al. (2015) and Elçi (2016) for further details on the values of the relevant parameters, such as cost, facility capacity limits, base values of the travel times and water demand, and the random distribution information, etc. Finally, we note that all scenarios are assumed to be equally likely.

We next discuss the additional issues that arise due to the different features of our new risk-averse model. The stochastic benchmarking constraints of interest additionally require us to specify a benchmark random outcome vector and a scalarization set. For illustrative purposes, we obtain a benchmark vector by identifying a reasonable and feasible benchmark first-stage solution and calculating the performance measures of interest associated with the corresponding optimal second-stage decisions. In particular, we consider the first-stage solution, where a single large-sized facility with the maximum possible inventory level (5394 units) is located at Houston, Baton Rouge, Biloxi, Atlanta and Orlando. In real life applications, a benchmark decision vector can be constructed by considering the existing practices of relief organizations. In addition, the scalarization set is assumed to have an ordered preference structure as follows:  $\mathcal{C} = \{\mathbf{c} \in \mathbb{R}_+^2 : c_1 + c_2 = 1, c_2 \geq c_1\}$ , where  $c_1$  and  $c_2$  correspond to the relative importance of the demand satisfaction criterion (24) and responsiveness criterion (25), respectively. This particular scalarization set description can be easily modified according the preferences of the decision makers.

**Computational Performance of the Solution Methods.** We evaluate the computational performance of the proposed solution algorithms in terms of the run times and the number of operations performed at each iteration.

All experiments are performed on a single thread of a Windows server with Intel(R) Xeon(R) CPU E5-2630 processor at 2.40 GHz and 32 GB of RAM using Java and Cplex 12.6.0. CPLEX is invoked with a time limit of 3600 seconds. MIP gap and feasibility tolerances are set to  $10^{-5}$  and  $10^{-9}$ , respectively; otherwise, we use the default CPLEX settings. The results are obtained for the instances with  $\alpha = 0.99, 0.95, 0.90$  and up to 1500 scenarios. For each setting, we report the average over three instances. We implement the cutting plane algorithms using the *lazy constraint callback* function of CPLEX. Considering the fact that using callbacks prevents CPLEX from utilizing dynamic search, we specify the MIP search method for both algorithms as *traditional branch-and-cut*. Following an acceleration technique suggested in Bodur and Luedtke (2016), we initialize the RMP of the scenario decomposition-based algorithm with optimality and feasibility cuts associated with an initial first-stage solution to benefit from the possible

improvement obtained from the automatically created CPLEX cuts. In particular, our initial solution locates two facilities at Mobile and Orlando with the smallest size. Additionally, we employ the *user cut callback* function of CPLEX for adding the Benders cuts at the root node for the fractional incumbent solutions. The *usercut callback* is invoked at each iteration at the root node until the improvement in the relative MIP gap is less than  $10^{-6}$  for the last five consecutive iterations.

In Table 1, we report several statistics on the performance of the delayed cut generation algorithm for DEF (DCG-DEF) and its scenario decomposition-based counterpart (DCG-SD). Under the “Time” columns, the solution times for only the instances solved within the time limit are reported. For both solution algorithms, we report the total run time and the time spent solving the separation problems under columns “Total/[%gap]”, and “Sep.”, respectively. If the algorithm terminated at the time limit with a feasible solution for some replications of the instance, then the average of the relative optimality gaps are reported under the column “Total/[%gap]”, inside square brackets. The relative optimality gap reported by CPLEX corresponds to the relative gap between the objective function value for the best available integer solution ( $z^{IP}$ ) and the best LP relaxation bound at the time of termination ( $z^{LP}$ ), i.e.  $100|z^{IP} - z^{LP}|/z^{IP}$ . For both algorithms, we report the average number of separation problems solved and the average number of scalarization vectors added under columns “Sep.” and “L”, respectively. For DCG-SD, we report the average total number of cuts ((10d), (12), and (13)) added to the master problem under column “Cuts”. We do not report the number of cuts for DCG-DEF, because in this case, constraints of type (5b) and (5c) are added when a new scalarization vector is included in the RMP. Hence, the total number of cuts added to the RMP for DCG-DEF is  $(|S| + 1)L$ , where  $|S|$  is the number of scenarios and  $L$  is the number of scalarization vectors generated.

The results show that DCG-SD outperforms DCG-DEF in terms of the computation times and number of instances that can be solved to optimality. Out of 63 instances, DCG-DEF is able to find an optimal solution for only 35 instances and it terminates with a feasible solution after one hour for seven instances. On the other hand, DCG-SD provides at least a feasible solution for all instances even if it fails to prove optimality for five of them. Considering the instances that can be solved to optimality by both methods, it can be seen that DCG-SD proves optimality in a shorter amount of time than DCG-DEF.

We observe that unlike the single-stage multivariate CVaR-constrained problems in Noyan and Rudolf (2013), the separation problem is no longer the bottleneck taking over 95% of the solution time. In fact, for DCG-DEF, the time spent on separation is negligible when compared to the overall solution time, with a small number of separation problems solved. On the other hand, a large number of separation problems are solved in DCG-SD, hence it is important to use the strongest formulations available for the separation problem (i.e., those provided in Liu et al., 2015 for the equiprobable scenario case) to reduce the time spent on generating the scalarization vectors. Nevertheless, a few instances cannot be solved with DCG-SD within the time limit mainly due to the long solution times of the separation MIPs. However, for these instances, our algorithm is able to find a feasible solution during its course, and hence it provides optimality gaps at termination. To avoid long solution times, one can implement heuristic separation methods. In addition, if  $\varepsilon$ -feasible solutions are satisfactory, then the violation thresholds can be updated accordingly.

We close this section by noting that the problem sizes we consider are significantly larger (in terms of first- and second-stage variables and constraints and/or the number of scenarios) when compared to the data sets in Noyan and Rudolf (2013) and Dentcheva and Wolfhagen (2015) for related optimization problems with multivariate stochastic benchmarking constraints. In addition, our first-stage problem involves discrete decisions, which results in the most difficult data sets considered for this problem class to date. In conclusion, our computational experiments showcase the scalability of the proposed scenario decomposition-based method with respect to the number of scenarios.

Table 1: Performance of the proposed delayed cut generation algorithms ( $|I| = |J| = 30$ )

$\alpha$	# Scs.	DCG-DEF				DCG-SD				
		Time (s)		Number		Time (s)		Number		
		Total / [%gap]	Sep.	Sep.	$L$	Total / [%gap]	Sep.	Cuts	Sep.	$L$
0.99	200	704.9	0.2	4.3	2.3	46.8	3.1	6101.3	62.0	3.0
	300	1483.5	0.7	5.3	3.0	108.7	5.1	9697.3	78.0	3.3
	400	2622.4	1.2	5.3	3.7	351.5	13.9	18569.7	122.0	5.3
	500	3071.3	0.8	4.0	3.0	363.3	17.7	20805.0	105.3	4.3
	600	[0.2] <sup>†</sup> **	-	-	-	662.1	42.4	27068.7	157.0	5.0
	800	***	-	-	-	987.2	51.7	31994.7	129.7	4.7
	1000	***	-	-	-	1594.2	110.3	43934.0	129.3	5.3
	1500	***	-	-	-	3450.6 [0.4] <sup>††</sup>	506.0	62436.0	172.0	4.0
0.95	200	408.9	0.7	2.7	0.7	148.6	17.4	9524.7	137.0	4.0
	300	1476.4	1.6	3.3	1.0	277.6	26.1	13280.0	128.7	4.0
	400	3392.7	2.1	3.7	1.7	536.8	45.2	20633.3	162.0	4.0
	500	2753.2 [0.2] <sup>†</sup>	1.4	2.5	1.5	671.9	150.6	23045.0	161.3	3.3
	600	[0.1] <sup>††</sup> *	-	-	-	878.0	156.2	26151.3	141.0	3.3
	800	***	-	-	-	1616.4	126.4	33607.7	115.7	3.0
	1000	***	-	-	-	1857.7	274.6	42229.3	140.3	3.3
	1500	***	-	-	-	3390.1 [1.4] <sup>††</sup>	867.2	52294.0	120.0	2.0
0.90	200	397.4	1.0	2.7	0.7	303.5	36.5	11781.0	180.0	4.3
	300	1128.5	0.6	3.0	0.7	352.5	48.7	13152.7	111.3	3.0
	400	2263.0	1.6	4.0	1.3	623.9	89.3	19482.3	130.3	3.3
	500	1660.2 [0.3] <sup>†</sup>	2.8	2.0	1.0	1392.1	228.6	25358.0	150.3	3.3
	600	2264.3 [0.1] <sup>††</sup>	2.2	2.0	1.0	857.5 [7.9] <sup>†</sup>	229.6	30051.0	153.5	3.0

†: Each dagger sign indicates one instance hitting the time limit with an integer feasible solution.

\*: Each asterisk sign indicates one instance hitting the time limit with no integer feasible solution.

**Acknowledgements** Singe Küçükyavuz and Merve Meraklı are supported, in part, by National Science Foundation Grant 1537317. Nilay Noyan acknowledges the support from The Scientific and Technological Research Council of Turkey under grant #115M560.

**Appendix A. Duality Results** In this section, we develop duality formulations and optimality conditions for the important special case of the multivariate CVaR-constrained two-stage model (1)-(2), where the mapping  $f$  is linear, the set  $\mathcal{X}$  is a polyhedral set, and the first-stage decisions are continuous. Let  $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}_+^{n_1} : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$  and  $f(\mathbf{x}) = \mathbf{f}^\top \mathbf{x}$  for some matrix  $A \in \mathbb{R}^{m_1 \times n_1}$ , and the vectors  $\mathbf{f} \in \mathbb{R}^{n_1}$  and  $\mathbf{b} \in \mathbb{R}^{m_1}$ . Then, the model of interest becomes

$$\text{(LinearP)} \quad \min \quad \mathbf{f}^\top \mathbf{x} + \sum_{s \in S} p_s \mathbf{q}_s^\top \mathbf{y}_s \quad (26a)$$

$$\text{s.t.} \quad \mathbf{A}\mathbf{x} \leq \mathbf{b}, \quad \mathbf{x} \in \mathbb{R}_+^{n_1}, \quad (26b)$$

$$(1b), (5f) - (5g). \quad (26c)$$

For a finite set  $\tilde{\mathcal{C}} = \{\tilde{\mathbf{c}}_{(1)}, \dots, \tilde{\mathbf{c}}_{(L)}\}$  we consider the following LP, referred to as (FiniteP( $\tilde{\mathcal{C}}$ )):

$$\min \{ \mathbf{f}^\top \mathbf{x} + \sum_{s \in S} p_s \mathbf{q}_s^\top \mathbf{y}_s : (5b) - (5d) \text{ (with } \bar{L} \text{ replaced by } L), (26b), \boldsymbol{\eta} \in \mathbb{R}^L, (5f) - (5g) \}.$$

The next lemma shows that (FiniteP( $\tilde{\mathcal{C}}$ )) is equivalent to (LinearP) for suitable choices of  $\tilde{\mathcal{C}}$ ; it is a simple consequence of Proposition 2.1.

LEMMA .1 Let  $\hat{\mathcal{C}} = \{\hat{\mathbf{c}}_{(1)}, \dots, \hat{\mathbf{c}}_{(L)}\}$  as defined in Proposition 2.1, and assume that the finite set  $\tilde{\mathcal{C}}$  satisfies  $\hat{\mathcal{C}} \subset \tilde{\mathcal{C}} \subset \mathcal{C}$ . Then a vector  $(\mathbf{x}, \mathbf{y})$  is a feasible (optimal) solution of (LinearP) if and only if  $(\mathbf{x}, \mathbf{y}, \boldsymbol{\eta}^{(\mathbf{x}, \mathbf{y})}, \mathbf{w}^{(\mathbf{x}, \mathbf{y})})$  is a feasible (optimal) solution of (FiniteP( $\tilde{\mathcal{C}}$ )), where  $\eta_{\ell}^{(\mathbf{x}, \mathbf{y})} = \text{VaR}_{\alpha}(\tilde{\mathbf{c}}_{(\ell)}^{\top} \hat{\mathbf{G}}(\mathbf{x}, \mathbf{y}))$  and  $w_{i\ell}^{(\mathbf{x}, \mathbf{y})} = [\eta_{\ell}^{(\mathbf{x}, \mathbf{y})} - \tilde{\mathbf{c}}_{(\ell)}^{\top} \hat{\mathbf{g}}_s(\mathbf{x}, \mathbf{y}_s)]_+$ .

The LP formulation (FiniteP( $\tilde{\mathcal{C}}$ )) does not immediately offer a tractable solution approach, however it leads to a natural delayed cut generation algorithm as discussed in Section 2. Moreover, (FiniteP( $\tilde{\mathcal{C}}$ )) provides an important direct way to derive strong duality results and optimality conditions, presented below. First, recall the following notation:  $\hat{\mathbf{g}}_s(\mathbf{x}, \mathbf{y}_s) = (\mathbf{g}_s^1(\mathbf{x}, \mathbf{y}_s), \mathbf{g}_s^2(\mathbf{x}, \mathbf{y}_s), \dots, \mathbf{g}_s^d(\mathbf{x}, \mathbf{y}_s))^{\top}$ ,  $s \in S$ , and  $\mathbf{g}_s^i(\mathbf{x}, \mathbf{y}_s) = \bar{\mathbf{g}}_s^i \mathbf{x} + \tilde{\mathbf{g}}_s^i \mathbf{y}_s$ ,  $i \in \{1, \dots, d\}$ ,  $s \in S$ . By abuse of notation, we introduce the random matrices  $\bar{\mathbf{g}} : \Omega \rightarrow \mathbb{R}^{d \times n_1}$  and  $\tilde{\mathbf{g}} : \Omega \rightarrow \mathbb{R}^{d \times n_2}$ , where  $\bar{\mathbf{g}}^{\top}(\omega_s) = \bar{\mathbf{g}}_s^{\top}$  and  $\tilde{\mathbf{g}}^{\top}(\omega_s) = \tilde{\mathbf{g}}_s^{\top}$  with the  $i$ th columns defined by  $\bar{\mathbf{g}}_s^i$  and  $\tilde{\mathbf{g}}_s^i$ , respectively. Denoting the set of all finitely supported finite non-negative measures on the scalarization polyhedron  $\mathcal{C}$  by  $\mathcal{M}_+^F(\mathcal{C})$ , we obtain the following dual problem to (LinearP):

$$\text{(LinearD)} \quad \max \quad - \int_{\mathcal{C}} \text{CVaR}_{\alpha}(\mathbf{c}^{\top} \mathbf{Z}) \mu(d\mathbf{c}) - \boldsymbol{\lambda}^{\top} \mathbf{b} + \mathbb{E}[\boldsymbol{\pi}^{\top} \mathbf{h}(\omega)] \quad (27a)$$

$$\text{s.t.} \quad \mathbb{E}(\nu) = \mu, \quad (27b)$$

$$\nu(\omega_s) \leq \frac{\mu}{1 - \alpha}, \quad \forall s \in S, \quad (27c)$$

$$\mathbb{E}[\boldsymbol{\pi}^{\top} T(\omega)] - \mathbb{E} \left( \int_{\mathcal{C}} \mathbf{c}^{\top} \bar{\mathbf{g}} \nu(d\mathbf{c}) \right) \leq \mathbf{f}^{\top} + \boldsymbol{\lambda}^{\top} A, \quad (27d)$$

$$\boldsymbol{\pi}_s^{\top} W_s - \int_{\mathcal{C}} \mathbf{c}^{\top} \tilde{\mathbf{g}}_s [\nu(\omega_s)](d\mathbf{c}) \leq \mathbf{q}_s^{\top}, \quad \forall s \in S, \quad (27e)$$

$$\boldsymbol{\lambda} \in \mathbb{R}_+^{m_1}, \quad \mu \in \mathcal{M}_+^F(\mathcal{C}), \quad \nu : \Omega \rightarrow \mathcal{M}_+^F(\mathcal{C}), \quad \boldsymbol{\pi}_s \in \mathbb{R}_+^{m_2}, \quad s \in S. \quad (27f)$$

The proof of the next duality theorem, which is similar to that of Theorem 3 in Noyan and Rudolf (2013), is omitted here for the sake of brevity. It mainly relies on Lemma .1 and the following facts for a finitely supported measure  $\mu \in \mathcal{M}_+^F(\mathcal{C})$ :  $\text{support}(\mu) = \{\mathbf{c} \in \mathcal{C} : \mu(\{\mathbf{c}\}) > 0\}$  and  $\int_{\mathcal{C}} u(\mathbf{c}) \mu(d\mathbf{c}) = \sum_{\mathbf{c} \in \text{support}(\mu)} u(\mathbf{c}) \mu(\{\mathbf{c}\})$  for a function  $u : \mathcal{C} \rightarrow \mathbb{R}$ .

THEOREM .1 The problem (LinearP) has a finite optimum value if and only if (LinearD) does, in which case the two optimum values coincide. In addition, a feasible solution  $(\mathbf{x}, \mathbf{y})$  of (LinearP) and a feasible solution  $(\boldsymbol{\lambda}, \mu, \nu, \boldsymbol{\pi})$  of (LinearD) are both optimal for their respective problems if and only if the following complementary slackness conditions hold:

- (i)  $\text{support}(\mu) \subset \left\{ \mathbf{c} : \text{CVaR}_{\alpha}(\mathbf{c}^{\top} \hat{\mathbf{G}}(\mathbf{x}, \mathbf{y})) = \text{CVaR}_{\alpha}(\mathbf{c}^{\top} \mathbf{Z}) \right\}$ ,
- (ii)  $\text{support}(\nu(\omega_s)) \subset \left\{ \mathbf{c} : \text{VaR}_{\alpha}(\mathbf{c}^{\top} \hat{\mathbf{G}}(\mathbf{x}, \mathbf{y})) \leq \mathbf{c}^{\top} \hat{\mathbf{g}}_s(\mathbf{x}, \mathbf{y}_s) \right\}, \quad s \in S$ ,
- (iii)  $\text{support}(\frac{\mu}{1 - \alpha} - \nu(\omega_s)) \subset \left\{ \mathbf{c} : \text{VaR}_{\alpha}(\mathbf{c}^{\top} \hat{\mathbf{G}}(\mathbf{x}, \mathbf{y})) \geq \mathbf{c}^{\top} \hat{\mathbf{g}}_s(\mathbf{x}, \mathbf{y}_s) \right\}, \quad s \in S$ ,
- (iv)  $\boldsymbol{\lambda}^{\top} (\mathbf{b} - A\mathbf{x}) = 0$ ,
- (v)  $\boldsymbol{\pi}_s^{\top} (T_s \mathbf{x} + W_s \mathbf{y}_s - \mathbf{h}_s) = 0, \quad s \in S$ ,
- (vi)  $(\mathbf{f}^{\top} + \boldsymbol{\lambda}^{\top} A - \mathbb{E}[\boldsymbol{\pi}^{\top} T(\omega)] + \mathbb{E}(\int_{\mathcal{C}} \mathbf{c}^{\top} \bar{\mathbf{g}} \nu(d\mathbf{c}))) \mathbf{x} = 0$ ,
- (vii)  $(\mathbf{q}_s^{\top} - \boldsymbol{\pi}_s^{\top} W_s + \int_{\mathcal{C}} \mathbf{c}^{\top} \tilde{\mathbf{g}}_s [\nu(\omega_s)](d\mathbf{c})) \mathbf{y}_s = 0, \quad s \in S$ .

We remark that our dual formulation is analogous to Haar's dual for semi-infinite linear programs (see, e.g., Bonnans and Shapiro, 2000). The finite representation of the multivariate CVaR relation (appearing in Proposition 2.1) proves to be useful to obtain strong duality results and optimality conditions directly from linear programming duality without the need for constraint qualifications.

**Lagrangian Duality.** The duality results established in Theorem .1 have a natural Lagrangian interpretation. In fact, measures on  $\mathcal{C}$  are a natural choice to use as Lagrange multipliers, since the CVaR

constraints in (1b) are indexed by the scalarization set  $\mathcal{C}$ . Accordingly, we define the Lagrangian function  $\mathcal{L} : (\mathcal{X}, \mathcal{Y}) \times \mathcal{M}_+^F(\mathcal{C}) \rightarrow \mathbb{R}$  as follows:

$$\mathcal{L}(\mathbf{x}, \mathbf{y}, \mu) = \mathbf{f}^\top \mathbf{x} + \sum_{s \in S} p_s \mathbf{q}_s^\top \mathbf{y}_s + \int_{\mathcal{C}} \text{CVaR}_\alpha(\mathbf{c}^\top \hat{\mathbf{G}}(\mathbf{x}, \mathbf{y})) \mu(d\mathbf{c}) - \int_{\mathcal{C}} \text{CVaR}_\alpha(\mathbf{c}^\top \mathbf{Z}) \mu(d\mathbf{c}). \quad (28)$$

Here, by abuse of notation,  $\mathbf{y} \in \mathbb{R}_+^{|S|n_2}$  represents the collection of decision vectors  $\mathbf{y}_s \in \mathbb{R}_+^{n_2}$ ,  $s \in S$ , and  $\mathbf{y} \in \mathcal{Y}$  refers to  $\mathbf{y}_s \in \mathcal{Y}(\mathbf{x}, \boldsymbol{\xi}(\omega_s))$ ,  $s \in S$ . For the ease of exposition, we assume that the feasible sets  $\mathcal{X}$  and  $\mathcal{Y}$  are compact; this is a very mild assumption due to the finiteness of the first-stage and second-stage objective function values. Then, (LinearP) is equivalent to  $\min_{(\mathbf{x}, \mathbf{y}) \in (\mathcal{X}, \mathcal{Y})} \max_{\mu \in \mathcal{M}_+^F(\mathcal{C})} \mathcal{L}(\mathbf{x}, \mathbf{y}, \mu)$ ,

while the corresponding Lagrangian dual problem is given by

$$(\text{LagrangianD}) \quad \max_{\mu \in \mathcal{M}_+^F(\mathcal{C})} \min_{(\mathbf{x}, \mathbf{y}) \in (\mathcal{X}, \mathcal{Y})} \mathcal{L}(\mathbf{x}, \mathbf{y}, \mu). \quad (29)$$

For any feasible solution of (LinearP) and any  $\mu \in \mathcal{M}_+^F(\mathcal{C})$ ,  $\mathcal{L}(\mathbf{x}, \mathbf{y}, \mu) \leq \mathbf{f}^\top \mathbf{x} + \sum_{s \in S} p_s \mathbf{q}_s^\top \mathbf{y}_s$ . The weak duality immediately follows:  $\text{OBF}_P \geq \text{OBF}_D$ , where  $\text{OBF}_P$  and  $\text{OBF}_D$ , respectively, denote the optimum objective values of (LinearP) and (LinearD). The next theorem, which is a consequence of Theorem .1, provides the strong duality result and optimality conditions.

**THEOREM .2** *If the primal problem (LinearP) has an optimal solution, then the dual problem (LinearD) also has an optimal solution, and the optimal objective values coincide. A primal feasible solution  $(\mathbf{x}^*, \mathbf{y}^*)$  and a dual feasible solution  $\mu^*$  are simultaneously optimal if and only if they satisfy the equations*

$$\int_{\mathcal{C}} \text{CVaR}_\alpha(\mathbf{c}^\top \hat{\mathbf{G}}(\mathbf{x}^*, \mathbf{y}^*)) \mu^*(d\mathbf{c}) = \int_{\mathcal{C}} \text{CVaR}_\alpha(\mathbf{c}^\top \mathbf{Z}) \mu^*(d\mathbf{c}), \quad (30)$$

$$\mathcal{L}(\mathbf{x}^*, \mathbf{y}^*, \mu^*) = \min_{(\mathbf{x}, \mathbf{y}) \in (\mathcal{X}, \mathcal{Y})} \mathcal{L}(\mathbf{x}, \mathbf{y}, \mu^*). \quad (31)$$

**PROOF.** Let us first show the following claim: *If the equations (30)-(31) hold for some primal feasible  $(\mathbf{x}^*, \mathbf{y}^*)$  and dual feasible  $\mu^*$ , then these solutions are simultaneously optimal with coinciding objective values.* According to (30) and (31), we have

$$\begin{aligned} \text{OBF}_P &\leq \mathbf{f}^\top \mathbf{x}^* + \sum_{s \in S} p_s \mathbf{q}_s^\top \mathbf{y}_s^* = \mathcal{L}(\mathbf{x}^*, \mathbf{y}^*, \mu^*) = \min_{(\mathbf{x}, \mathbf{y}) \in (\mathcal{X}, \mathcal{Y})} \mathcal{L}(\mathbf{x}, \mathbf{y}, \mu^*) \\ &\leq \max_{\mu \in \mathcal{M}_+^F(\mathcal{C})} \min_{(\mathbf{x}, \mathbf{y}) \in (\mathcal{X}, \mathcal{Y})} \mathcal{L}(\mathbf{x}, \mathbf{y}, \mu) = \text{OBF}_D. \end{aligned}$$

On the other hand, the weak duality implies that  $\text{OBF}_P \geq \text{OBF}_D$ , which proves the claim.

*We next prove the claim that if  $(\mathbf{x}^*, \mathbf{y}^*)$  is primal optimal and  $\mu^*$  is dual optimal, then they satisfy the equations (30)-(31). To this end, we first show that for any given primal optimal solution  $(\mathbf{x}^*, \mathbf{y}^*)$  there exists a corresponding dual feasible solution  $\bar{\mu}$  such that (30) and (31) are satisfied for the choice  $\mu^* = \bar{\mu}$ .* Let us consider an optimal solution  $(\mathbf{x}^*, \mathbf{y}^*)$  of (LinearP). By Theorem .1 there exists an optimal solution  $(\bar{\boldsymbol{\lambda}}, \bar{\mu}, \bar{\nu}, \bar{\boldsymbol{\pi}})$  of (LinearD) with the same optimal objective value. According to the complementary slackness condition (i), the equality  $\text{CVaR}_\alpha(\mathbf{c}^\top \hat{\mathbf{G}}(\mathbf{x}^*, \mathbf{y}^*)) = \text{CVaR}_\alpha(\mathbf{c}^\top \mathbf{Z})$  holds on the support of the measure  $\bar{\mu}$ , which implies (30). *To show that (31) also holds, we next prove that, substituting  $\mu^* = \bar{\mu}$ ,  $\mathcal{L}(\mathbf{x}, \mathbf{y}, \bar{\mu}) \geq \mathcal{L}(\mathbf{x}^*, \mathbf{y}^*, \bar{\mu})$  is valid for any primal feasible solution  $(\mathbf{x}, \mathbf{y})$  (and equality holds for  $(\mathbf{x}^*, \mathbf{y}^*)$ ).* For ease of exposition, let  $U(\mathbf{x}, \mathbf{y}, \bar{\mu}) = \int_{\mathcal{C}} \text{CVaR}_\alpha(\mathbf{c}^\top \hat{\mathbf{G}}(\mathbf{x}, \mathbf{y})) \bar{\mu}(d\mathbf{c})$  and  $U(\mathbf{Z}, \bar{\mu}) = \int_{\mathcal{C}} \text{CVaR}_\alpha(\mathbf{c}^\top \mathbf{Z}) \bar{\mu}(d\mathbf{c})$ . Then, using the definition of the Lagrangian function (28) and the dual feasibility conditions (27d)-(27e) we obtain the following relations:

$$\begin{aligned} \mathcal{L}(\mathbf{x}, \mathbf{y}, \bar{\mu}) + U(\mathbf{Z}, \bar{\mu}) &= \mathbf{f}^\top \mathbf{x} + \sum_{s \in S} p_s \mathbf{q}_s^\top \mathbf{y}_s + U(\mathbf{x}, \mathbf{y}, \bar{\mu}) \geq -\bar{\boldsymbol{\lambda}}^\top \mathbf{A} \mathbf{x} + \mathbb{E}[\bar{\boldsymbol{\pi}}^\top T(\omega) \mathbf{x}] \\ &\quad - \mathbb{E} \left( \int_{\mathcal{C}} \mathbf{c}^\top \bar{\mathbf{g}} \mathbf{x} \bar{\nu}(d\mathbf{c}) \right) + \sum_{s \in S} p_s \left( \bar{\boldsymbol{\pi}}_s^\top W_s \mathbf{y}_s - \int_{\mathcal{C}} \mathbf{c}^\top \bar{\mathbf{g}}_s \mathbf{y}_s [\bar{\nu}(\omega_s)](d\mathbf{c}) \right) + U(\mathbf{x}, \mathbf{y}, \bar{\mu}) = -\bar{\boldsymbol{\lambda}}^\top \mathbf{A} \mathbf{x} \end{aligned}$$

$$+ \sum_{s \in S} p_s \bar{\pi}_s^\top (T_s \mathbf{x} + W_s \mathbf{y}_s) - \mathbb{E} \left( \int_{\mathcal{C}} \mathbf{c}^\top \bar{\mathbf{g}} \mathbf{x} \bar{\nu}(\mathrm{d}\mathbf{c}) \right) - \sum_{s \in S} p_s \int_{\mathcal{C}} \mathbf{c}^\top \tilde{\mathbf{g}}_s \mathbf{y}_s [\bar{\nu}(\omega_s)](\mathrm{d}\mathbf{c}) + U(\mathbf{x}, \mathbf{y}, \bar{\mu}). \quad (32)$$

Now observe that the complementary slackness condition (iv) implies  $-\bar{\lambda}^\top \mathbf{A} \mathbf{x} \geq -\bar{\lambda}^\top \mathbf{b} = -\bar{\lambda}^\top \mathbf{A} \mathbf{x}^*$ , as  $\mathbf{A} \mathbf{x} \leq \mathbf{b}$  holds for all  $\mathbf{x} \in X$ , and the Lagrange multiplier  $\bar{\lambda}$  is nonnegative. Similarly, as  $T_s \mathbf{x} + W_s \mathbf{y}_s \geq \mathbf{h}_s$  holds for all  $\mathbf{y}_s \in \mathcal{Y}(\mathbf{x}, \xi(\omega_s))$ ,  $s \in S$ , and the Lagrange multiplier  $\bar{\pi}_s$  is nonnegative for all  $s \in S$ , the complementary slackness condition (v) implies that  $\bar{\pi}_s^\top (T_s \mathbf{x} + W_s \mathbf{y}_s) \geq \bar{\pi}_s^\top \mathbf{h}_s = \bar{\pi}_s^\top (T_s \mathbf{x}^* + W_s \mathbf{y}_s^*)$ . Therefore, the following relation is valid for any primal feasible solution  $(\mathbf{x}, \mathbf{y})$ :

$$-\bar{\lambda}^\top \mathbf{A} \mathbf{x} + \sum_{s \in S} p_s \bar{\pi}_s^\top (T_s \mathbf{x} + W_s \mathbf{y}_s) \geq -\bar{\lambda}^\top \mathbf{A} \mathbf{x}^* + \sum_{s \in S} p_s \bar{\pi}_s^\top (T_s \mathbf{x}^* + W_s \mathbf{y}_s^*). \quad (33)$$

Observing that  $U(\mathbf{Z}, \bar{\mu})$  is independent of the decision vectors  $\mathbf{x}$  and  $\mathbf{y}$ , by (32) and (33), it is sufficient to show that the inequality  $-\mathbb{E} \left( \int_{\mathcal{C}} \mathbf{c}^\top \bar{\mathbf{g}} \mathbf{x} \bar{\nu}(\mathrm{d}\mathbf{c}) \right) - \sum_{s \in S} p_s \int_{\mathcal{C}} \mathbf{c}^\top \tilde{\mathbf{g}}_s \mathbf{y}_s [\bar{\nu}(\omega_s)](\mathrm{d}\mathbf{c}) + U(\mathbf{x}, \mathbf{y}, \bar{\mu}) \geq -\mathbb{E} \left( \int_{\mathcal{C}} \mathbf{c}^\top \bar{\mathbf{g}} \mathbf{x}^* \bar{\nu}(\mathrm{d}\mathbf{c}) \right) - \sum_{s \in S} p_s \int_{\mathcal{C}} \mathbf{c}^\top \tilde{\mathbf{g}}_s \mathbf{y}_s^* [\bar{\nu}(\omega_s)](\mathrm{d}\mathbf{c}) + U(\mathbf{x}^*, \mathbf{y}^*, \bar{\mu})$  is valid for any primal feasible solution  $(\mathbf{x}, \mathbf{y})$  (and equality holds for  $(\mathbf{x}^*, \mathbf{y}^*)$ ). Accordingly, we prove that the following chain of inequalities holds for any  $(\mathbf{x}, \mathbf{y}) \in (\mathcal{X}, \mathcal{Y})$ , and that all inequalities hold with equality for  $(\mathbf{x}, \mathbf{y}) = (\mathbf{x}^*, \mathbf{y}^*)$ .

$$-\mathbb{E} \left( \int_{\mathcal{C}} \mathbf{c}^\top \bar{\mathbf{g}} \mathbf{x} \bar{\nu}(\mathrm{d}\mathbf{c}) \right) - \sum_{s \in S} p_s \int_{\mathcal{C}} \mathbf{c}^\top \tilde{\mathbf{g}}_s \mathbf{y}_s [\bar{\nu}(\omega_s)](\mathrm{d}\mathbf{c}) + U(\mathbf{x}, \mathbf{y}, \bar{\mu}) \quad (34)$$

$$= - \sum_{s \in S} p_s \int_{\mathcal{C}} \mathbf{c}^\top (\bar{\mathbf{g}}_s \mathbf{x} + \tilde{\mathbf{g}}_s \mathbf{y}_s) [\bar{\nu}(\omega_s)](\mathrm{d}\mathbf{c}) + \int_{\mathcal{C}} \left( \text{VaR}_\alpha(\mathbf{c}^\top \hat{\mathbf{G}}(\mathbf{x}, \mathbf{y})) + \frac{1}{1-\alpha} \sum_{s \in S} p_s [\mathbf{c}^\top \hat{\mathbf{g}}_s(\mathbf{x}, \mathbf{y}_s) - \text{VaR}_\alpha(\mathbf{c}^\top \hat{\mathbf{G}}(\mathbf{x}, \mathbf{y}))]_+ \right) \bar{\mu}(\mathrm{d}\mathbf{c}) \quad (35)$$

$$= - \sum_{s \in S} p_s \int_{\mathcal{C}} \left( \mathbf{c}^\top \hat{\mathbf{g}}_s(\mathbf{x}, \mathbf{y}_s) - \text{VaR}_\alpha(\mathbf{c}^\top \hat{\mathbf{G}}(\mathbf{x}, \mathbf{y})) \right) [\bar{\nu}(\omega_s)](\mathrm{d}\mathbf{c}) + \frac{1}{1-\alpha} \sum_{s \in S} p_s \int_{\mathcal{C}} [\mathbf{c}^\top \hat{\mathbf{g}}_s(\mathbf{x}, \mathbf{y}_s) - \text{VaR}_\alpha(\mathbf{c}^\top \hat{\mathbf{G}}(\mathbf{x}, \mathbf{y}))]_+ \bar{\mu}(\mathrm{d}\mathbf{c}) \quad (36)$$

$$\geq \sum_{s \in S} p_s \int_{\mathcal{C}} \left( [\mathbf{c}^\top \hat{\mathbf{g}}_s(\mathbf{x}, \mathbf{y}_s) - \text{VaR}_\alpha(\mathbf{c}^\top \hat{\mathbf{G}}(\mathbf{x}, \mathbf{y}))]_+ - (\mathbf{c}^\top \hat{\mathbf{g}}_s(\mathbf{x}, \mathbf{y}_s) - \text{VaR}_\alpha(\mathbf{c}^\top \hat{\mathbf{G}}(\mathbf{x}, \mathbf{y}))) \right) [\bar{\nu}(\omega_s)](\mathrm{d}\mathbf{c}) \quad (37)$$

$$\geq 0. \quad (38)$$

We note that an alternative definition of CVaR for a random variable  $V$  is given by  $\text{CVaR}_\alpha(V) = \text{VaR}_\alpha(V) + \frac{1}{1-\alpha} \mathbb{E}([V - \text{VaR}_\alpha(V)]_+)$ , where  $[z]_+ = \max(z, 0)$  (Rockafellar and Uryasev, 2000). Substituting this definition into the third term and expanding the expected value in the first term (34) becomes (35). Using the dual feasibility condition (27b) to replace  $\bar{\mu}$  in the second term of (35) provides (36). Then, (37) follows from (27c); equality for  $(\mathbf{x}, \mathbf{y}) = (\mathbf{x}^*, \mathbf{y}^*)$  is ensured by the complementary slackness condition (iii). Finally, (38) is a consequence of the trivial inequality  $[a]_+ \geq a$  and the non-negativity of the random measure  $\nu$ . Equality is again ensured for  $(\mathbf{x}, \mathbf{y}) = (\mathbf{x}^*, \mathbf{y}^*)$ , since the complementary slackness condition (ii) implies that  $\mathbf{c}^\top \hat{\mathbf{g}}_s(\mathbf{x}, \mathbf{y}_s) - \text{VaR}_\alpha(\mathbf{c}^\top \hat{\mathbf{G}}(\mathbf{x}, \mathbf{y})) \geq 0$  holds on the support of  $\bar{\nu}(\omega_s)$  for all  $s \in S$ .

Finally, let us consider a primal optimal solution  $(\mathbf{x}^*, \mathbf{y}^*)$  and a dual optimal solution  $\mu^*$ . We just showed above that there exists some dual feasible  $\bar{\mu}$  such that  $U(\mathbf{x}^*, \mathbf{y}^*, \bar{\mu}) = U(\mathbf{Z}, \bar{\mu})$  and  $\mathcal{L}(\mathbf{x}^*, \mathbf{y}^*, \bar{\mu}) = \min_{(\mathbf{x}, \mathbf{y}) \in (\mathcal{X}, \mathcal{Y})} \mathcal{L}(\mathbf{x}, \mathbf{y}, \bar{\mu}) = \mathbf{f}^\top \mathbf{x}^* + \sum_{s \in S} p_s \mathbf{q}_s^\top \mathbf{y}_s^*$ . As the set of feasible primal solutions is compact and the Lagrangian function  $\mathcal{L}$  is continuous, there also exists some  $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$  such that  $\mathcal{L}(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \mu^*) = \min_{(\mathbf{x}, \mathbf{y}) \in (\mathcal{X}, \mathcal{Y})} \mathcal{L}(\mathbf{x}, \mathbf{y}, \mu^*)$ . Then, by the optimality of  $\mu^*$ , we have  $\mathbf{f}^\top \mathbf{x}^* + \sum_{s \in S} p_s \mathbf{q}_s^\top \mathbf{y}_s^* = \min_{(\mathbf{x}, \mathbf{y}) \in (\mathcal{X}, \mathcal{Y})} \mathcal{L}(\mathbf{x}, \mathbf{y}, \bar{\mu}) \leq \min_{(\mathbf{x}, \mathbf{y}) \in (\mathcal{X}, \mathcal{Y})} \mathcal{L}(\mathbf{x}, \mathbf{y}, \mu^*) = \mathcal{L}(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \mu^*)$ , and consequently, the following relation holds:

$$\begin{aligned} U(\mathbf{x}^*, \mathbf{y}^*, \mu^*) - U(\mathbf{Z}, \mu^*) &= \mathcal{L}(\mathbf{x}^*, \mathbf{y}^*, \mu^*) - \mathbf{f}^\top \mathbf{x}^* - \sum_{s \in S} p_s \mathbf{q}_s^\top \mathbf{y}_s^* \\ &\geq \mathcal{L}(\mathbf{x}^*, \mathbf{y}^*, \mu^*) - \mathcal{L}(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \mu^*) = \mathcal{L}(\mathbf{x}^*, \mathbf{y}^*, \mu^*) - \min_{(\mathbf{x}, \mathbf{y}) \in (\mathcal{X}, \mathcal{Y})} \mathcal{L}(\mathbf{x}, \mathbf{y}, \mu^*) \geq 0. \end{aligned}$$

On the other hand, by the primal feasibility of  $(\mathbf{x}^*, \mathbf{y}^*)$  and the non-negativity of  $\mu^*$ ,  $U(\mathbf{x}^*, \mathbf{y}^*, \mu^*) - U(\mathbf{Z}, \mu^*) \leq 0$  holds, which immediately implies (30) and (31).  $\square$



## References

- Ahmed, S.: Convexity and decomposition of mean-risk stochastic programs **106**(3), 433–446 (2006). DOI <http://dx.doi.org/10.1007/s10107-005-0638-8>
- Artzner, P., Delbaen, F., Eber, J., Heath, D.: Coherent measures of risk. *Mathematical Finance* **9**(3), 203–228 (1999)
- Balcik, B., Beamon, B.: Facility location in humanitarian relief. *International Journal of Logistics: Research and Applications* **11**(2), 101–121 (2008)
- Bodur, M., Luedtke, J.R.: Mixed-integer rounding enhanced benders decomposition for multiclass service-system staffing and scheduling with arrival rate uncertainty. *Management Science*, in press (2016). DOI 10.1287/mnsc.2016.2455
- Bonnans, J.F., Shapiro, A.: Perturbation analysis of optimization problems. Springer Series in Operations Research. Springer-Verlag, New York (2000)
- Çelik, M., Ergun, O., Johnson, B., Keskinocak, P., Lorca, A., Pekgün, P., Swann, J.: Humanitarian logistics. In: Mirchandani, P. (ed.) *INFORMS Tutorials in Operations Research*, pp. 18–49. INFORMS, Maryland (2012)
- Dentcheva, D., Martinez, G.: Two-stage stochastic optimization problems with stochastic ordering constraints on the recourse. *European Journal of Operational Research* **219**(1), 1 – 8 (2012)
- Dentcheva, D., Ruszczyński, A.: Optimization with stochastic dominance constraints. *SIAM Journal on Optimization* **14**(2), 548–566 (2003)
- Dentcheva, D., Ruszczyński, A.: Optimization with multivariate stochastic dominance constraints. *Mathematical Programming* **117**(1), 111–127 (2009)
- Dentcheva, D., Wolfhagen, E.: Optimization with multivariate stochastic dominance constraints. *SIAM Journal on Optimization* **25**(1), 564–588 (2015)
- Dentcheva, D., Wolfhagen, E.: Two-stage optimization problems with multivariate stochastic order constraints. *Mathematics of Operations Research* **41**(1), 1–22 (2016)
- Döyen, A., Aras, N., Barbarosoğlu, G.: A two-echelon stochastic facility location model for humanitarian relief logistics. *Optimization Letters* **6**(6), 1123–1145 (2012)
- Dupáčová, J.: Risk objectives in two-stage stochastic programming models. *Kybernetika* **44**(2), 227–242 (2008). <http://dml.cz/dmlcz/135845>
- Elçi, O.: Chance-constrained stochastic programming models for humanitarian relief network design. MS Thesis, Sabanci University (2016)
- Fábián, C.I.: Handling CVaR objectives and constraints in two-stage stochastic models. *European Journal of Operational Research* **191**(3), 888 – 911 (2008)
- Gollmer, R., Gotzes, U., Schultz, R.: A note on second-order stochastic dominance constraints induced by mixed-integer linear recourse. *Mathematical Programming* **216**(1), 179–190 (2011)
- Gutjahr, W.J., Nolz, P.C.: Multicriteria optimization in humanitarian aid. *European Journal of Operational Research* **252**(2), 351 – 366 (2016)

- Homem-de-Mello, T., Mehrotra, S.: A cutting surface method for uncertain linear programs with linear stochastic dominance constraints. *SIAM Journal on Optimization* **20**(3), 1250–1273 (2009)
- Hong, X., Lejeune, M.A., Noyan, N.: Stochastic network design for disaster preparedness. *IIE Transactions* **47**(4), 329–357 (2015)
- Hu, J., Mehrotra, S.: Robust and stochastically weighted multiobjective optimization models and reformulations. *Operations Research* **60**(4), 936–953 (2012)
- Hu, J., Homem-de Mello, T., Mehrotra, S.: Sample average approximation of stochastic dominance constrained programs. *Mathematical Programming* **133**(1-2), 171–201 (2012)
- Huang, M., Smilowitz, K., Balcik, B.: Models for relief routing: Equity, efficiency and efficacy. *Transportation Research Part E* **48**(1), 2–18 (2012)
- Küçükayvuz, S., Noyan, N.: Cut generation for optimization problems with multivariate risk constraints. *Mathematical Programming* **159**(1), 165–199 (2016)
- Liberatore, F., Pizarro, C., Blas, C.S., Ortuño, M.T., Vitoriano, B.: Decision Aid Models for Disaster Management and Emergencies, chap. Uncertainty in Humanitarian Logistics for Disaster Management. A Review, pp. 45–74. Atlantis Press, Paris (2013)
- Liu, X., Küçükayvuz, S., Noyan, N.: Robust multicriteria risk-averse stochastic programming models (2015). [http://www.optimization-online.org/DB\\_FILE/2015/10/5157.pdf](http://www.optimization-online.org/DB_FILE/2015/10/5157.pdf)
- Noyan, N.: Risk-averse two-stage stochastic programming with an application to disaster management. *Computers and Operations Research* **39**(3), 541–559 (2012)
- Noyan, N., Rudolf, G.: Optimization with multivariate conditional value-at-risk-constraints. *Operations Research* **61**(4), 990–1013 (2013)
- Noyan, N., Rudolf, G.: Optimization with stochastic preferences based on a general class of scalarization functions (2016). [http://www.optimization-online.org/DB\\_HTML/2016/09/5636.html](http://www.optimization-online.org/DB_HTML/2016/09/5636.html)
- Rawls, C.G., Turnquist, M.A.: Pre-positioning of emergency supplies for disaster response. *Transportation Research Part B: Methodological* **44**(4), 521–534 (2010)
- Rawls, C.G., Turnquist, M.A.: Pre-positioning planning for emergency response with service quality constraints. *OR Spectrum* **33**(3), 481–198 (2011)
- Rockafellar, R., Uryasev, S.: Optimization of conditional value-at-risk. *The Journal of Risk* **2**(3), 21–41 (2000)
- Ruszczynski, A., Shapiro, A. (eds.): *Stochastic Programming, Handbooks in Operations Research and Management Science* 10. Elsevier, Amsterdam (2003)
- Salmerón, J., Apte, A.: Stochastic optimization for natural disaster asset prepositioning. *Production and Operations Management* **19**(5), 561–574 (2010)
- Takriti, S., Ahmed, S.: On robust optimization of two-stage systems. *Mathematical Programming* **99**(1), 109–126 (2004)

- Van Slyke, R., Wets, R.: L-shaped linear programs with applications to optimal control and stochastic programming. *SIAM Journal of Applied Mathematics* **17**(4), 638–663 (1969)
- Vitoriano, B., Ortuño, M.T., Tirado, G., Montero, J.: A multi-criteria optimization model for humanitarian aid distribution. *Journal of Global Optimization* **51**(2), 189–208 (2011)