manuscript

Sharing the Value-at-Risk under Distributional Ambiguity

Zhi Chen
Department of Management Sciences, College of Business, City University of Hong Kong, Kowloon Tong, Hong Kong
zhi.chen@cityu.edu.hk

Weijun Xie
Department of Industrial & Systems Engineering, Virginia Tech, Blacksburg, United States
wxie@vt.edu

This paper considers the problem of risk sharing, where a coalition of homogeneous agents, each bearing a random cost, aggregates their costs and shares the value-at-risk of such a risky position. Due to limited distributional information in practice, the joint distribution of agents' random costs is difficult to acquire. The coalition, being aware of the distributional ambiguity, thus evaluates the worst-case value-at-risk within a commonly agreed ambiguity set of the possible joint distributions. Through the lens of cooperative game theory, we show that this coalitional worst-case value-at-risk is subadditive for the popular ambiguity sets in the distributionally robust optimization literature that are based on *(i)* convex moments or *(ii)* Wasserstein distance to some reference distributions. In addition, we propose easy-to-compute core allocation schemes to share the worst-case value-at-risk. Our results can be readily extended to sharing the worst-case conditional value-at-risk under distributional ambiguity.

Key words: Risk sharing; value-at-risk; conditional value-at-risk; distributionally robust optimization.
History: The paper was submitted on July 13, 2019. This version: March 3, 2020.

---

1. Introduction

In a cooperative risk-sharing game, there is a set of agents $N=\{1,\ldots,n\}$, each bearing a random cost $\tilde{x}_i$ and associated with a homogeneous risk preference, considering aggregating their costs and sharing the value-at-risk (VaR) of such a risky position $\mathbb{P}\text{-VaR}_{1-\varepsilon}[\tilde{x}_N]$ with $\tilde{x}_N=\sum_{i\in N}\tilde{x}_i$. Here given a vector $\tilde{\boldsymbol{x}}=(\tilde{x}_i)_{i\in[N]}$ of random costs governed by a joint distribution $\mathbb{P}$ and a risk threshold $\varepsilon\in(0,1)$, the VaR of the aggregate random cost $\tilde{x}_N$ at level ε is defined as follows:

$$\mathbb{P}\text{-VaR}_{1-\varepsilon}[\tilde{x}_N]\triangleq\inf_{v\in\mathbb{R}}\left\{v \;\middle|\; \mathbb{P}[\tilde{x}_N>v]\leq\varepsilon\right\}.$$

Due to limited distributional information, the joint distribution $\mathbb{P}$ is usually difficult, if not impossible, to acquire in practice. Being ambiguity-aware, these agents evaluate instead the worst-case VaR within a commonly agreed family $\mathcal{F}$ of joint distributions:

$$\mathcal{F}\text{-VaR}_{1-\varepsilon}[\tilde{x}_N]=\sup_{\mathbb{P}\in\mathcal{F}}\mathbb{P}\text{-VaR}_{1-\varepsilon}[\tilde{x}_N]\triangleq\inf_{v\in\mathbb{R}}\left\{v \;\middle|\; \sup_{\mathbb{P}\in\mathcal{F}}\mathbb{P}[\tilde{x}_N>v]\leq\varepsilon\right\}.$$

Such a family $\mathcal{F}$, termed "ambiguity set", encodes partial distributional information about the unknown true joint distribution $\mathbb{P}$. The information is typically characterized by valid upper bounds on convex moments (*e.g.*, mean, covariance) as well as on statistical distances to some reference distributions—both types of distributional information are rigorously justifiable through finite-sample guarantees (see respectively, Delage and Ye 2010 and Mohajerin Esfahani and Kuhn 2018).

The fundamental incentive of sharing the worst-case VaR lays on the existence of the *core* of this cooperative game, *i.e.*, the non-emptiness of the following set

$$\text{Core} \triangleq \left\{ \boldsymbol{a} \;\middle|\; \sum_{i\in N} a_i = C(N),\ \sum_{i\in S} a_i \le C(S) \ \ \forall S \subseteq N \right\},$$

where the characteristic function $C(\cdot): 2^N \mapsto \mathbb{R}$ is defined through

$$C(S) = \mathcal{F}\text{-VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] \ \ \forall S \subseteq N \quad \text{and} \quad C(\emptyset) = 0.$$

Here for every $S \subseteq N$, the i^{th} entry of an indicator vector $\boldsymbol{e}_S \in \{0,1\}^N$ takes one if and only if $i \in S$.

In general, the core of the cooperative game $\langle N, C(\cdot)\rangle$ might be empty. Fortunately, since our characteristic function $C(\cdot)$ defined as the coalitional worst-case VaR is homogeneous with degree one (Shapiro et al. 2009), according to the result by Anily and Haviv (2014), the non-emptiness of the core can be guaranteed through the *subadditivity* of $C(\cdot)$. The subadditivity property, an indication of the economies of scale, can motivate two agents or two disjoint coalitions $S, S' \subseteq N$ of agents to cooperate, as they face a lower risk while cooperating compared to individual activities. However, the desired subadditivity *fails* to hold for a general ambiguity set $\mathcal{F}$.

This paper shows that the cooperative game of worst-case VaR sharing is subadditive under the popular ambiguity set based on convex moments and under the recently emerging Wasserstein ambiguity set based on the statistical distance, provided that the reference distribution, as well as the norm defining the distance, satisfies certain conditions. Our proof builds on interpreting the worst-case VaR using the (worst-case) conditional value-at-risk (CVaR). In particular, we show that the worst-case VaR always coincides with the worst-case CVaR under the convex moment ambiguity set; while under the Wasserstein ambiguity set, the worst-case VaR can be effectively bisected by querying a constraint on the CVaR of a *transformed* random variable under the reference distribution. For all mentioned subadditive cases, we propose easy-to-compute core allocation schemes to share the total risk of the aggregate risky position $\tilde{x}_N$.

1.1. Relevant Literature

This paper follows the stream of treating the risk sharing problem via the lens of cooperative game (see, for example, Borch 1962, Baton and Lemaire 1981, Denault 2001, Csóka et al. 2009). Originating from seminal works on the reinsurance market (Borch 1960a,b), the risk sharing problem has been of great interest in various industries, including actuarial science and finance, among others. In general, the risk sharing problem is concerned with how to divide the total risk of the aggregate risky position $\tilde{x}_N$ among the N participating agents such that each agent i would be endowed with a random cost $\tilde{y}_i$ and $\sum_{i\in N}\tilde{y}_i = \tilde{x}_N$. In our setting where agents have homogeneous risk preferences, it is indeed sufficient to analyze the cooperative game $\langle N, C(\cdot)\rangle$ with agents' transferable utilities; see more detailed discussion in Chen et al. (2017).

Our work is closely related to the important question of quantifying the worst-case VaR of $\tilde{x}_N$ under distributional ambiguity in risk management (see, *e.g.*, Embrechts and Puccetti 2010, Kaas et al. 2009). Given marginal distributions of random costs $\tilde{x}_i, i \in N$ and their partially specified dependence structure, bounds for the worst-case VaR of $\tilde{x}_N$ and conditions for their sharpness have been extensively explored in the literature of quantitative risk management (see Wang et al. 2013 and references therein). Our approach, without assuming known marginal distributions, mainly leverages techniques in distributionally robust optimization to assess the following *uncertainty quantification problem*

$$\sup_{\mathbb{P}\in\mathcal{F}} \mathbb{P}[\tilde{x}_N > v],$$

which plays a key role in the definition of worst-case VaR. Tractable deterministic reformulations for the uncertainty quantification problem under various ambiguity sets have been studied in the literature on distributionally robust optimization (see the survey paper by Hanasusanto et al. 2015 and references therein). While these deterministic reformulations naturally facilitate the assessment of our interested worst-case VaR of $\tilde{x}_N$, our analysis further reveals the intimate relationship between the worst-case VaR and its (worst-case) CVaR counterpart, which can be of independent interest in the distributionally robust optimization community.

In the following sections, we first introduce key ingredients in a cooperative game and CVaR, and we then discuss our results for the convex moment ambiguity set and the Wasserstein ambiguity set, respectively. Finally, we connect our results to the cooperative game of worst-case CVaR sharing. All proofs are relegated to the last section.

2. Preliminaries

This section introduces the basic concepts of cooperative game and the definition and properties of the worst-case CVaR.

2.1. Cooperative Game, Core and Subaddtivity

A cooperative game $\langle N, C(\cdot)\rangle$ with transferable utilities is characterized by two main components: *(i)* the set of agents $N = \{1,\ldots,n\}$ and *(ii)* the characteristic function $C(\cdot): 2^N \mapsto \mathbb{R}$. The characteristic function assigns the total cost (risk in our content) to each coalition $S \subseteq N$ of agents and defines the core and subbadditivity of the game.

DEFINITION 1 (CORE). The core for a cooperative game $\langle N, C(\cdot)\rangle$ is the set of all allocation vectors $\boldsymbol{a} = (a_i)_{i\in N}$ that satisfy *(i)* efficiency: $\sum_{i\in N} a_i = C(N)$, and *(ii)* coalitional rationality: $\sum_{i\in S} a_i \le C(S)$ for all $S \subseteq N$. The core is always well-defined, however, it can be empty.

DEFINITION 2 (SUBADDITIVITY). A cooperative game $\langle N, C(\cdot)\rangle$ is subadditive if its characteristic function is subadditive for any two disjoint subsets, that is,

$$C(S\cup T) \le C(S) + C(T) \quad \forall S,T \subseteq N : S\cap T = \emptyset.$$

The appealing subadditivity of the characteristic function, by iteself, does not necessarily imply the non-emptiness of the core.

EXAMPLE 1. Consider a subadditive cooperative game with the following characteristic function:

$$C(\{1\}) = C(\{2\}) = C(\{3\}) = 5,\ C(\{1,2\}) = C(\{2,3\}) = C(\{1,3\}) = 6,\ C(\{1,2,3\}) = 10.$$

The core satisfying $a_1 + a_2 \le 6$, $a_2 + a_3 \le 6$, and $a_1 + a_3 \le 6$ implies that $a_1 + a_2 + a_3 \le 9$, which contradicts to the efficiency constraint $a_1 + a_2 + a_3 = 10$. □

Fortunately it is known in the literature that if the characteristic function is homogeneous with degree one, then the subadditivity property is sufficient to guarantee the non-emptiness of the core (see theorem 1 in Anily and Haviv 2014). Note that the characteristic function is said to be homogeneous of degree one if for any integer m, the coalitional risk of cloning m times all agents in any coalition, leads to m times the original coalitional risk. In this paper, the characteristic function, defined as the coalitional worst-case VaR, is indeed homogeneous with degree one (Shapiro et al. 2009). Thus, to prove that the core is non-empty, it is sufficient to show that the coalitional worst-case VaR is subadditive.

2.2. Worst-Case CVaR

Given the joint distribution $\mathbb{P}$ of $\tilde{\boldsymbol{x}}$, the CVaR of the aggregate random cost $\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}$ (for any $S \subseteq N$) at the risk level ε is defined as

$$\mathbb{P}\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] \triangleq \inf_{\tau\in\mathbb{R}} \left\{ \tau + \mathbb{E}_{\mathbb{P}}\left[\frac{(\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}} - \tau)^+}{\varepsilon} \right] \right\};$$

and given an ambiguity set $\mathcal{F}$ of the possible joint distributions, the worst-case CVaR at level ε is defined as (see Shapiro and Kleywegt 2002)

$$\mathcal{F}\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] = \sup_{\mathbb{P}\in\mathcal{F}} \mathbb{P}\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}].$$

At the same risk level ε, the CVaR of the risky position $\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}$ governed by a known distribution $\mathbb{P}$ is larger than its VaR. Hence, the worst-case CVaR of $\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}$ is no smaller than the worst-case VaR of $\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}$ for any non-empty ambiguity set $\mathcal{F}$. For any disjoint sets $S, T \subseteq N$ and any non-empty ambiguity set $\mathcal{F}$, we have

$$\mathcal{F}\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_{S\cup T}^\top \tilde{\boldsymbol{x}}] = \mathcal{F}\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}} + \boldsymbol{e}_T^\top \tilde{\boldsymbol{x}}] \leq \mathcal{F}\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] + \mathcal{F}\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_T^\top \tilde{\boldsymbol{x}}],$$

where the inequality follows from the fact that the worst-case CVaR under any non-empty $\mathcal{F}$ remains coherent. By Theorem 1 in Anily and Haviv (2014) as well as the facts that any coherent risk measure is subaddtive and homogeneous with degree one, the core of a cooperative game where the characteristic function is defined as the coalitional worst-case CVaR is *always* non-empty. We refer to Delbaen (2000) for the general concept of coherent risk measures and their connection to convex game theory.

3. Convex Moment Ambiguity Set

In this section we consider the convex moment ambiguity set of the following canonial form

$$\mathcal{F}_{\text{M}} = \{\mathbb{P} \in \mathcal{P}_0(\mathbb{R}^n) \mid \mathbb{P}[\tilde{\boldsymbol{x}} \in \mathcal{X}] = 1,\ \mathbb{E}_{\mathbb{P}}[\tilde{\boldsymbol{x}}] = \boldsymbol{\mu},\ \mathbb{E}_{\mathbb{P}}[\boldsymbol{g}(\tilde{\boldsymbol{x}})] \leq \boldsymbol{\nu}\},$$

where the support set $\mathcal{X}$ is convex and compact[1], the mean value $\boldsymbol{\mu}$ is in the interior of the support $\mathcal{X}$ (*i.e.*, $\boldsymbol{\mu} \in \text{int}(\mathcal{X})$), and the convex function $\boldsymbol{g}(\cdot): \mathbb{R}^n \mapsto \mathbb{R}^m$ is closed as well as component-wise convex and satisfies $\boldsymbol{g}(\boldsymbol{\mu}) < \boldsymbol{\nu}$. The convex moment ambiguity set contains a wide range of ambiguity sets in its intuitive yet rich expression. Beside the support and mean information, the convex moment ambiguity set can further characterize the unknown true distribution through its dispersion specified by the convex function $\boldsymbol{g}(\cdot)$. Specifically, for every $j = 1, \ldots, m$, the expectation constraint $\mathbb{E}_{\mathbb{P}}[g_j(\tilde{\boldsymbol{x}})] \leq \nu_j$ bounds from above the expected dispersion specified by $g_j(\cdot)$ with a constant ν_j. Possible choices of the dispersion include the mean absolute deviations, the (co-)variances, higher order moments, among others (see, *e.g.*, Wiesemann et al. 2014, Hanasusanto et al. 2017, Xie and Ahmed 2018, Bertsimas et al. 2019, Li et al. 2019).

Quite notably, the worst-case VaR and CVaR coincide under the convex moment ambiguity set.

Theorem 1. *Given a convex moment ambiguity set $\mathcal{F}_{\text{M}}$, for any $S \subseteq N$, the coalitional worst-case VaR of the aggregate random cost $\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}$ coincides with its worst-case CVaR, that is,*

$$\mathcal{F}_{\text{M}}\text{-VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] = \mathcal{F}_{\text{M}}\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}].$$

With the established equivalence and the fact that the worst-case CVaR is subadditive, the coalitional worst-case VaR under the convex moment ambiguity set $\mathcal{F}_{\text{M}}$ is thus subadditive. Another byproduct of the established equivalence leads to an allocation scheme that belongs to the core.

[1] According to Xie and Ahmed (2018), Theorem 1 still holds if we relax the compactness assumption of the support set $\mathcal{X}$. However, this assumption is necessary to Theorem 2. Thus, for the sake of simplicity, we assume that $\mathcal{X}$ is compact throughout this section.

THEOREM 2. *Given a convex moment ambiguity set $\mathcal{F}_{\mathrm{M}}$, for any $S \subseteq N$, the coalitional worst-case VaR can be determined by the following convex optimization problem:*

$$\begin{aligned} C(S) = \mathcal{F}_{\mathrm{M}}\text{-}\mathrm{VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] = \sup_{\boldsymbol{x},\boldsymbol{y}} \ & \boldsymbol{e}_S^\top \boldsymbol{x} \\ \text{s.t.} \ & \varepsilon \boldsymbol{x} + (1-\varepsilon)\boldsymbol{y} = \boldsymbol{\mu} \\ & \varepsilon \boldsymbol{g}(\boldsymbol{x}) + (1-\varepsilon)\boldsymbol{g}(\boldsymbol{y}) \le \boldsymbol{\nu} \\ & \boldsymbol{x},\ \boldsymbol{y} \in \mathcal{X}. \end{aligned} \tag{1}$$

The allocation scheme $\boldsymbol{a} = (a_i)_{i\in N} = (x^\star_{N,i})_{i\in N}$, where $\boldsymbol{x}^\star_N$ is (part of) the optimal solution to problem (1) *with $S = N$, belongs to the core of the cooperative game $\langle N, C(\cdot)\rangle$.*

Following from Theorem 2, we can obtain a core allocation to the cooperative game $\langle N, C(\cdot)\rangle$ by solving problem (1) on one occasion of $S = N$. Note that the optimization problem (1) is convex and can be easily solved using general purpose solvers.

One notable special case of the convex moment ambiguity set is the marginal convex moment ambiguity set of the form

$$\mathcal{F}_{\mathrm{M}} = \{\mathbb{P} \in \mathcal{P}_0(\mathbb{R}^n) \mid \mathbb{P}[\tilde{\boldsymbol{x}} \in \mathcal{X}] = 1,\ \mathbb{E}_{\mathbb{P}}[\tilde{\boldsymbol{x}}] = \boldsymbol{\mu},\ \mathbb{E}_{\mathbb{P}}[\boldsymbol{g}_i(\tilde{x}_i)] \le \boldsymbol{\nu}_i\ \forall i \in N\}, \tag{2}$$

where the support set $\mathcal{X} = [\underline{\boldsymbol{x}}, \overline{\boldsymbol{x}}]$ is a box and where in stark contrast to the generic convex moment ambiguity set, only marginal moment information about an individual random cost $\tilde{x}_i$ is specified and no dependency between different random costs is included. Over marginal convex moment ambiguity sets, the coalitional worst-case VaR is in fact *additive*, that is,

$$C(S \cup T) = C(S) + C(T) \quad \forall S, T \subseteq N : S \cap T = \emptyset.$$

This is due to the fact that for any $S \subseteq N$, it corresponds to a *comonotonic* worst-case distribution, *i.e.*, there exists one comonotonic worst-case distribution for the coalitional worst-case VaR.

DEFINITION 3 (COMONOTONICITY). A random vector $\tilde{\boldsymbol{x}} = (\tilde{x}_1, \ldots, \tilde{x}_n)$ is comonotonic if there are a random variable $\tilde{z}$ and non-decreasing functions $f_1, \ldots, f_n$ such that

$$(\tilde{x}_1, \ldots, \tilde{x}_n) \overset{\text{dist}}{=} (f_1(\tilde{z}), \ldots, f_n(\tilde{z})).$$

THEOREM 3. *Given a marginal convex moment ambiguity set $\mathcal{F}_{\mathrm{M}}$ in* (2)*, for any $S \subseteq N$, the coalitional worst-case VaR is additive, that is,*

$$C(S) = \mathcal{F}_{\mathrm{M}}\text{-}\mathrm{VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] = \sum_{i\in S} \mathcal{F}_{\mathrm{M}}\text{-}\mathrm{VaR}_{1-\varepsilon}[\tilde{x}_i] = \sum_{i\in S} C(\{i\}).$$

Remark 1. (Known Marginal Normal Distributions and Comonotonicity) In fact, it is not uncommon for a coalition to perceive an additive coalitional worst-case VaR associated with a comonotonic worst-case distribution. For another example, suppose that each agent bears a normal random cost $\tilde{x}_i \sim \mathbb{N}(\mu_i, \sigma_i^2)$ and the coalition $S \subseteq N$ only knows that the aggregate random cost $\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}$ is normally distributed but has no clue about the dependency among these marginal normal components. In such cases, the aggregate random cost takes $\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}} \sim \mathbb{N}(\boldsymbol{e}_S^\top \boldsymbol{\mu}, \boldsymbol{e}_S^\top \boldsymbol{\Sigma} \boldsymbol{e}_S)$, where it is only known that the covariance matrix $\boldsymbol{\Sigma} \succeq \boldsymbol{0}$ satisfies $\boldsymbol{\Sigma}_{ii} = \sigma_i^2$ for every $i \in N$. The coalitional worst-case VaR at the risk level $\varepsilon \in (0, 1/2]$ of this normal distribution amounts to

$$C(S) = \max_{\boldsymbol{\Sigma} \succeq \boldsymbol{0},\ \boldsymbol{\Sigma}_{ii} = \sigma_i^2\ \forall i \in N} \left\{ \boldsymbol{e}_S^\top \boldsymbol{\mu} + \Phi^{-1}(\varepsilon) \sqrt{\boldsymbol{e}_S^\top \boldsymbol{\Sigma} \boldsymbol{e}_S} \right\} = \boldsymbol{e}_S^\top (\boldsymbol{\mu} + \Phi^{-1}(\varepsilon) \boldsymbol{\sigma}),$$

where the optimal covariance matrix with random costs being perfectly positive correlated gives the second equality. That is, the worst-case distribution is *comonotonic* while the coalitional worst-case VaR is *additive*. This result naturally extends to the setting of known marginal elliptical distributions[2] but ambiguous dependency. The argument is similar and is omitted for brevity.

Remark 2. In a recent paper, Ghosal and Wiesemann (2018) also proved the subadditivity (additivity) of worst-case VaR under the convex (marginal) moment ambiguity set, based on an elegant analysis of the optimization problem for determining the worst-case VaR (a reformulation of problem (9) in our paper), whereas our proof follows straightforwardly from the notable coincidence of worst-case VaR and CVaR under the convex moment ambiguity set.

4. Wasserstein Ambiguity Set

In this section we focus on the Wasserstein ambiguity set which has recently attracted a great deal of research interest in both stochastic programming and distributionally robust optimization communities. The Wasserstein ambiguity set can be viewed as a ball in the space of distributions with respect to the celebrated Wasserstein distance. Equipped with a general norm $\|\cdot\|$ on $\mathbb{R}^n$, the Wasserstein distance $d(\mathbb{P}_1, \mathbb{P}_2)$ between two distributions $\mathbb{P}_1$ and $\mathbb{P}_2$ on $\mathbb{R}^n$ is defined as the minimal transportation cost of moving $\mathbb{P}_1$ to $\mathbb{P}_2$, under the premise that the cost of moving a unit mass from $\boldsymbol{x}_1$ to $\boldsymbol{x}_2$ amounts to $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|$. Mathematically, this gives that

$$d(\mathbb{P}_1, \mathbb{P}_2) = \inf_{\mathbb{P} \in \mathcal{Q}(\mathbb{P}_1, \mathbb{P}_2)} \mathbb{E}_{\mathbb{P}}[\|\tilde{\boldsymbol{x}}_1 - \tilde{\boldsymbol{x}}_2\|],$$

where $\tilde{\boldsymbol{x}}_1$ (respectively, $\tilde{\boldsymbol{x}}_2$) follows the distribution $\mathbb{P}_1$ (respectively, $\mathbb{P}_2$) and $\mathcal{Q}(\mathbb{P}_1, \mathbb{P}_2)$ represents the set of all distributions on $\mathbb{R}^n \times \mathbb{R}^n$ with marginals $\mathbb{P}_1$ and $\mathbb{P}_2$. The Wasserstein ambiguity set

[2] We will formally introduce elliptical distributions subsequently and we refer interested readers to Gupta et al. (2013) for more details and properties about elliptical distributions.

$\mathcal{F}_{\mathrm{W}}(\theta)$ is then defined as a ball of radius $\theta \geq 0$ with respect to the Wasserstein distance, around a prescribed reference distribution $\mathbb{P}_{\mathrm{R}}$:

$$\mathcal{F}_{\mathrm{W}}(\theta) = \{\mathbb{P} \in \mathcal{P}(\mathbb{R}^n) \mid d(\mathbb{P}, \mathbb{P}_{\mathrm{R}}) \leq \theta\}.$$

Compared with the convex moment ambiguity set, the Wasserstein ambiguity set provides an attractive option that is especially favorable in data-driven settings. If a finite dataset $\{\hat{\boldsymbol{x}}_k\}_{k\in[K]}$ of historical realizations of random costs is available, a natural choice for $\mathbb{P}_{\mathrm{R}}$ is the empirical distribution $\hat{\mathbb{P}} = \frac{1}{K}\sum_{k=1}^{K} \delta_{\hat{\boldsymbol{x}}_k}$, where $\delta_{\hat{\boldsymbol{x}}}$ denotes the Dirac function that places unit mass on the realization $\tilde{\boldsymbol{x}} = \hat{\boldsymbol{x}}$. If confident estimations about mean and covariance are to be incorporated with shape information, it is also helpful to choose $\mathbb{P}_{\mathrm{R}}$ from the family of elliptical distributions that generalizes multivariate normal distribution and multivariate t-distribution, among others. We refer to Kuhn et al. (2019) on a tutorial of Wasserstein distributionally robust optimization models.

Quite interestingly, the worst-case VaR under the Wasserstein ambiguity set can be effectively bisected by querying a constraint involving the CVaR of a *transformed* random variable under the reference distribution.

Theorem 4. *For any reference distribution $\mathbb{P}_{\mathrm{R}}$ and any Wasserstein radius $\theta \geq 0$, we have*

$$\begin{aligned} C(S) = \mathcal{F}_{\mathrm{W}}(\theta)\text{-VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] = \inf_{v}\ & v \\ \text{s.t.}\ & -\mathbb{P}_{\mathrm{R}}\text{-CVaR}_{1-\varepsilon}[-(v - \boldsymbol{e}_S^\top \tilde{\boldsymbol{x}})^+] \geq \frac{\theta}{\varepsilon}\|\boldsymbol{e}_S\|_* \\ & v \geq \mathbb{P}_{\mathrm{R}}\text{-VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}], \end{aligned} \tag{3}$$

where $\|\cdot\|_*$ is the dual norm of $\|\cdot\|$ and $(x)^+ = \max\{x, 0\}$.

In the subsequent sections, we investigate the case where the reference distribution is the discrete empirical distribution and the case where the reference distribution is the continuous elliptical distribution, respectively.

4.1. Empirical Reference Distribution

In this subsection, we consider the empirical reference distribution $\mathbb{P}_{\mathrm{R}} = \frac{1}{K}\sum_{k\in[K]} \delta_{\hat{\boldsymbol{x}}_k}$, which is a discrete distribution uniformly supported on empirical realizations $\hat{\boldsymbol{x}}_1, \ldots, \hat{\boldsymbol{x}}_K$. In this case, we can further simplify the coalitional worst-case VaR determined in Theorem 4.

Proposition 1. *Given the empirical reference distribution $\mathbb{P}_{\mathrm{R}} = \frac{1}{K}\sum_{k\in[K]} \delta_{\hat{\boldsymbol{x}}_k}$ and the Wasserstein radius $\theta \geq 0$, for any $S \subseteq N$, the coalitional worst-case VaR is determined by the following optimization problem:*

$$\begin{aligned} C(S) = \mathcal{F}_{\mathrm{W}}(\theta)\text{-VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] = \min_{v}\ & v \\ \text{s.t.}\ & \sum_{j\in[\kappa]} (v - \boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_j(S)})^+ + (\varepsilon K - \kappa)(v - \boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_\kappa(S)})^+ \geq \theta K \|\boldsymbol{e}_S\|_* \\ & v \geq \boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_\kappa(S)}, \end{aligned} \tag{4}$$

where $\kappa = \lfloor \varepsilon K \rfloor + 1$ *and* $\boldsymbol{\pi}(S)$ *is a (coalition-dependent) permutation of* $[K]$ *that orders the empirical realizations in increasing coalitional sum, i.e.,* $\boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_1(S)} \geq \boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_2(S)} \geq \cdots \geq \boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_K(S)}$.

The constraints in problem (4) essentially provide a deterministic reformulation of the distributionally robust chance constraint $\sup_{\mathbb{P}\in\mathcal{F}_{\mathrm{W}}(\theta)} \mathbb{P}[\boldsymbol{e}_S^\top \boldsymbol{x} > v] \leq \varepsilon$. That is,

$$\sup_{\mathbb{P}\in\mathcal{F}_{\mathrm{W}}(\theta)} \mathbb{P}[\boldsymbol{e}_S^\top \boldsymbol{x} > v] \leq \varepsilon \Longleftrightarrow \begin{cases} f_S(v) \triangleq \sum_{j\in[\kappa]} (v - \boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_j(S)})^+ + (\varepsilon K - \kappa)(v - \boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_\kappa(S)})^+ \geq \theta K \|\boldsymbol{e}_S\|_* \\ v \geq \boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_\kappa(S)}. \end{cases}$$

The function $f_S(v)$ defined above is non-negative, non-decreasing, and convex piecewise affine with break points $\boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_\kappa(S)}, \ldots, \boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_1(S)}$. In fact, it can be represented as

$$f_S(v) = \begin{cases} 0 & \text{if } v < \boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_\kappa(S)} \\ (\varepsilon K - \kappa)(v - \boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_\kappa(S)}) + jv - \sum_{i\in[j]} \boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_{\kappa+1-i}(S)} & \text{if } \begin{array}{l} \boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_{\kappa+1-j}(S)} \leq v < \boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_{\kappa-j}(S)} \\ \text{for some } j \in [\kappa - 1] \end{array} \\ \varepsilon K v - \sum_{i\in[\kappa]} \boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_i(S)} - (1 + \varepsilon K - \kappa) \boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_\kappa(S)} & \text{if } v \geq \boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_1(S)}. \end{cases}$$

Let $v^\star$ be the optimal solution to problem (4) and $\boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_0 \triangleq \infty$. Clearly, $v^* \geq \boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_\kappa(S)}$. Hence solving problem (4) is equivalent to finding the smallest index $j^\star \in [\kappa]$ such that $v^\star$ of problem (4) lies in the range $[\boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\kappa+1-j^\star}, \boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\kappa-j^\star})$. We can then express the coalitional worst-case VaR as

$$C(S) = v^\star = \frac{(\varepsilon K - \kappa)\boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_\kappa(S)} + \sum_{i\in[j^\star]} \boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_{\kappa+1-i}(S)} + \theta K \|\boldsymbol{e}_S\|_*}{j^\star + \varepsilon K - \kappa}. \tag{5}$$

Recall that when $\theta = 0$, the Wasserstein ambiguity set shrinks to a singleton set containing only the reference distribution. In such cases, the worst-case VaR degenerates to the empirical VaR (*i.e.*, in (5), we have $j^\star = 1$ and $C(S) = v^\star = \boldsymbol{e}_S^\top \hat{\boldsymbol{x}}_{\pi_\kappa(S)}$), which, purely depends on the observed empirical realizations. However, the coalitional empirical VaR is *not* subadditive in general.

EXAMPLE 2. Consider $K = 100$ empirical realizations $\hat{\boldsymbol{x}}_1, \ldots, \hat{\boldsymbol{x}}_{100}$ such that

$$\hat{\boldsymbol{x}}_k = \begin{cases} (1,0) & \text{if } k = 1, \ldots, 9 \\ (0,1) & \text{if } k = 10, \ldots, 18 \\ (0,0) & \text{if } k = 19, \ldots, 100. \end{cases}$$

For the empirical reference distribution $\mathbb{P}_{\mathrm{R}} = \frac{1}{K}\sum_{k\in[K]} \delta_{\hat{\boldsymbol{x}}_k}$, we have $C(\{1\}) = \mathbb{P}_{\mathrm{R}}\text{-VaR}_{0.1}[\tilde{x}_1] = 0$, $C(\{2\}) = \mathbb{P}_{\mathrm{R}}\text{-VaR}_{0.1}[\tilde{x}_2] = 0$, $C(\{1,2\}) = \mathbb{P}_{\mathrm{R}}\text{-VaR}_{0.1}[\tilde{x}_1 + \tilde{x}_2] = 1$. The inequality $C(\{1\}) + C(\{2\}) < C(\{1,2\})$ reveals that the coalitional empirical VaR at the risk level 0.1 is not subadditive. □

The above example reveals that the subadditivity of the coalitional worst-case VaR critically depends on both the observed empirical realizations and the chosen risk level. To ensure the coalitional empirical VaR to be subadditive, one possibility is that its random components are perfectly positive dependent. That is, if the empirical distribution is comonotonic, the coalitional empirical VaR becomes (sub)additive (recall the aforementioned Proposition 6.15 in McNeil et al. 2005). Therefore, the coalitional worst-case VaR under a singleton Wasserstein ambiguity set with the comonotonic empirical reference distribution (*i.e.*, $\mathcal{F}_{\mathrm{W}}(0)=\{\mathbb{P}_{\mathrm{R}}\}$ and $\mathbb{P}_{\mathrm{R}}$ is comonotonic), $C(S)=\mathcal{F}_{\mathrm{W}}(0)\text{-}\mathrm{VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}]$, is indeed subadditive. Our next result shows that when observing comonotonic empirical distribution, even if the coalition has ambiguity about the observation (*i.e.*, $\theta>0$), the coalitional worst-case VaR may still be subadditive, as long as the ambiguity (measured by the radius θ) is not very large.

THEOREM 5. *Suppose that the empirical reference distribution is comonotonic and*

$$0\le\theta\le\min_{S\subseteq N}\left\{\frac{(1+\varepsilon K-\kappa)(\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_{\kappa-1}-\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_{\kappa})}{K\|\boldsymbol{e}_S\|_*}\right\}$$

where for all $S\subseteq N$, $\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_1\ge\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_2\ge\cdots\ge\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_K$.

(i) The coalitional worst-case VaR

$$C(S)=\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_\kappa+\frac{\theta K\|\boldsymbol{e}_S\|_*}{1+\varepsilon K-\kappa}\quad\forall S\subseteq N \tag{6}$$

is subadditive, where by default, we let $\frac{0}{0}=0$.

(ii) Let $\boldsymbol{x}_N^\star$ *be the optimal solution to the optimization problem* $\|\boldsymbol{e}_N\|_*\triangleq\max\{\boldsymbol{e}_N^\top\boldsymbol{x}\mid\|\boldsymbol{x}\|\le1\}$. *The allocation scheme* $\boldsymbol{a}=(a_i)_{i\in N}$ *with*

$$a_i=\hat{x}_{\kappa,i}+\frac{\theta K}{1+\varepsilon K-\kappa}x_{N,i}^\star\quad\forall i\in[N]$$

belongs to the core of the cooperative game $\langle N,C(\cdot)\rangle$, *where* $C(\cdot)$ *is defined in* (6).

REMARK 3. When εK is an integer, the results in Theorem 5 can be strengthened as follows: the upper bound on the radius θ becomes

$$\theta\le\min_{S\subseteq N}\left\{\frac{\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_{\varepsilon K-1}-\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_{\varepsilon K}}{K\|\boldsymbol{e}_S\|_*}\right\},$$

whereas the conditional worst-case VaR amounts to $C(S)=\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_{\varepsilon K}+\theta K\|\boldsymbol{e}_S\|_*$ and the allocation scheme follows correspondingly.

Quite interestingly, the proof of Theorem 5 inspires us that the coalition may also perceive a subadditive coalitional worst-case VaR if it only accesses to limited empirical realizations or if it has a large ambiguity about the empirical realizations, regardless of whether the empirical distribution is comonotonic or not.

THEOREM 6. *Suppose that the empirical reference distribution satisfies one of the conditions:*

(i) if $K<1/\varepsilon$, *that is, the empirical realizations are scarce; or*

(ii) if $\theta \geq \max_{S\subseteq N}\left\{\frac{\varepsilon K\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_{\pi_1(S)}-\sum_{i\in[\kappa]}\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_{\pi_i(S)}-(\varepsilon K+1-\kappa)\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_{\pi_\kappa(S)}}{K\|\boldsymbol{e}_S\|_*}\right\}$*, that is, the ambiguity is beyond certain degree.*

Then under either condition, the coalitional worst-case VaR is subadditive and is given by

$$C(S)=\mathbb{P}_{\mathrm{R}}\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}]+\frac{\theta\|\boldsymbol{e}_S\|_*}{\varepsilon}\quad\forall S\subseteq N.$$

Using a similar approach as in Theorem 5, we can propose a core allocation scheme based on the triangle inequality of the dual norm. The proof follows from similar lines as in that of Theorem 5 and is thus omitted for brevity.

Corollary 1. *Let* $\boldsymbol{x}_N^\star$ *be the optimal solution to determine* $\|\boldsymbol{e}_N\|_*=\max\{\boldsymbol{e}_N^\top\boldsymbol{x}\mid\|\boldsymbol{x}\|\leq 1\}$*.*

(i) The allocation scheme $\boldsymbol{a}=(a_i)_{i\in N}$ *with*

$$a_i=\hat{x}_{\pi_1(N),i}+\frac{\theta}{\varepsilon}x_{N,i}^\star\quad\forall i\in N$$

belongs to the core of the cooperative game $\langle N,C(\cdot)\rangle$ *under condition (i) in Theorem 5.*

(ii) The allocation scheme $\boldsymbol{a}=(a_j)_{j\in N}$ *with*

$$a_j=(\varepsilon K-\kappa)\hat{x}_{\pi_\kappa(N),j}+\sum_{i\in[\kappa]}\hat{x}_{\pi_{\kappa+1-i}(N),j}+\frac{\theta}{\varepsilon}x_{N,j}^\star\quad\forall j\in N$$

belongs to the core of the cooperative game $\langle N,C(\cdot)\rangle$ *under condition (ii) in Theorem 5.*

4.2. Elliptical Reference Distribution

In this subsection, we consider the case when the central reference distribution $\mathbb{P}_{\mathrm{R}}$ is an elliptical distribution, a generalization of multivariate normal distributions. Elliptical distributions have been widely used to quantify the risk in the literature (see, e.g., Landsman and Valdez 2003, Embrechts 2000, Pérignon and Smith 2010, Embrechts et al. 2002, Kamdem 2005). We denote an elliptical distribution as $\mathbb{P}_{\mathrm{E}}(\boldsymbol{\mu},\boldsymbol{\Sigma},g)$ with a location parameter $\boldsymbol{\mu}$, a positive definite matrix $\boldsymbol{\Sigma}$, and a generating function g, where its probability density function f is

$$f(\boldsymbol{x})=k\cdot g\left(\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)$$

with a positive normalization scalar k.

A random vector $\tilde{\boldsymbol{x}}\sim\mathbb{P}_{\mathrm{E}}(\boldsymbol{\mu},\boldsymbol{\Sigma},g)$ following an elliptical distribution can be written into $\tilde{\boldsymbol{x}}=\boldsymbol{\mu}+\boldsymbol{\Sigma}^{1/2}\tilde{\boldsymbol{\xi}}$, where $\tilde{\boldsymbol{\xi}}\sim\mathbb{P}_{\mathrm{E}}(\boldsymbol{0},\boldsymbol{I},g)$ denotes the standard elliptical random vector. In addition, any linear combination of components of $\tilde{\boldsymbol{x}}$ is elliptically distributed, which implies that the random scalar $\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}=\boldsymbol{e}_S^\top\boldsymbol{\mu}+\boldsymbol{e}_S^\top\boldsymbol{\Sigma}^{1/2}\tilde{\boldsymbol{\xi}}$ follows an elliptical distribution $\mathbb{P}_{\mathrm{E}}(\mu_S,\sigma_S^2,g)$ with $\mu_S=\boldsymbol{e}_S^T\boldsymbol{\mu}$ and $\sigma_S^2=\boldsymbol{e}_S^T\boldsymbol{\Sigma}\boldsymbol{e}_S$. For other interesting properties of elliptical distributions, we refer the readers to Gupta et al. (2013). Using the aforementioned properties, we can rewrite the elliptical random scalar into $\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}} = \tilde{x}_S = \mu_S + \sigma_S \tilde{z}$, where $\tilde{z}$ is referred to as the standard elliptical random variable following the *standard elliptical distribution* $\mathbb{P}_{\mathrm{E}}^0(0,1,g)$ generated by g. The probability density function of the standard elliptical distribution is given by $\phi(z) = kg(z^2/2)$ and the corresponding cumulative distribution function is

$$\Phi(a) = \int_{-\infty}^{a} kg\left(\frac{z^2}{2}\right) dz,$$

where its inverse is denoted by $\Phi^{-1}(\cdot)$.

The following result shows that the worst-case VaR under the Wasserstein ball around an elliptical distribution $\mathbb{P}_{\mathrm{E}}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$ is subadditive, provided that the norm $\|\cdot\|$ is properly chosen; moreover, we derive an allocation scheme in the core.

Theorem 7. *Suppose that the reference distribution $\mathbb{P}_{\mathrm{R}}$ is an elliptical distribution $\mathbb{P}_{\mathrm{E}}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$, the Wasserstein radius $\theta \geq 0$, and the Wasserstein distance is equipped with a Mahalanobis norm associated with the positive definite matrix $\boldsymbol{\Sigma}$, defined as $\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{x}}$.*

(i) The coalitional worst-case VaR at the risk level $\varepsilon \in (0, 1/2]$ is subadditive and is given by

$$C(S) = \mathcal{F}_{\mathrm{W}}(\theta)\text{-VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] = \boldsymbol{e}_S^\top \boldsymbol{\mu} + \eta^\star \sqrt{\boldsymbol{e}_S^T \boldsymbol{\Sigma} \boldsymbol{e}_S} \quad \forall S \subseteq N,$$

where $\eta^\star$ the minimal $\eta \geq \Phi^{-1}(1-\varepsilon)$ that satisfies

$$\eta(\Phi(\eta) - (1-\varepsilon)) - \int_{\frac{1}{2}(\Phi^{-1}(1-\varepsilon))^2}^{\frac{\eta^2}{2}} kg(z) dz \geq \theta.$$

(ii) The allocation scheme $\boldsymbol{a} = (a_i)_{i \in N}$ with

$$a_i = \mu_i + \eta^\star \cdot \frac{\boldsymbol{e}_{\{i\}}^\top \boldsymbol{\Sigma} \boldsymbol{e}_N}{\sqrt{\boldsymbol{e}_N^\top \boldsymbol{\Sigma} \boldsymbol{e}_N}} \quad \forall i \in N$$

belongs to the core of the cooperative game $\langle N, C(\cdot) \rangle$.

Quite notably, from the proof of Theorem 7, we note that the nominal elliptical VaR is given by

$$\mathbb{P}_{\mathrm{E}}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)\text{-VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] = \boldsymbol{e}_S^\top \boldsymbol{\mu} + \Phi^{-1}(1-\varepsilon) \sqrt{\boldsymbol{e}_S^T \boldsymbol{\Sigma} \boldsymbol{e}_S} \quad \forall S \subseteq N.$$

Since $\eta^\star \geq \Phi^{-1}(1-\varepsilon)$, if we define a new risk level $\varepsilon^\star \triangleq 1 - \Phi(\eta^\star) \leq \varepsilon$, we can then see that the coalitional worst-case VaR is equivalent to the elliptical VaR at a slightly more risky level (*i.e.*, $1 - \varepsilon^\star \geq 1 - \varepsilon$). This observation is summarized below.

Corollary 2. *Under the conditions in Theorem 7, the coalitional worst-case VaR at the risk level $\varepsilon \in (0, 1/2]$ is equivalent to the nominal elliptical VaR at the risk level $\varepsilon^\star = 1 - \Phi(\eta^\star)$, that is,*

$$C(S) = \mathcal{F}_{\mathrm{W}}(\theta)\text{-VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] = \mathbb{P}_{\mathrm{E}}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)\text{-VaR}_{1-\varepsilon^\star}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] \quad \forall S \subseteq N.$$

5. Sharing the Worst-Case CVaR

Our results in the previous two sections can be readily applied to the risk sharing problem where the coalition of agents consider sharing the CVaR of the aggregate risky position under distributional ambiguity. In such cases, the characteristic function is instead defined as the coalitional worst-case CVaR as follows:

$$\overline{C}(S) = \mathcal{F}\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] \quad \forall S \subseteq N \quad \text{and} \quad \overline{C}(\emptyset) = 0.$$

As mentioned in Section 2, the cooperative game $\langle N, \overline{C}(\cdot) \rangle$ is always subadditive under any non-empty ambiguity set $\mathcal{F}$ and has the non-empty core. For the convex moment ambiguity set, following from Theorem 1, the allocation scheme proposed in Theorem 2 lies in the core of the cooperative game $\langle N, \overline{C}(\cdot) \rangle$. Therefore, in this section, we will mainly focus on the coalitional worst-case CVaR under Wasserstein Ambiguity Set.

We first establish the following closed-form relationship between the worst-case CVaR and the nominal CVaR under the reference distribution. Exploring this relationship, we can still leverage established ingredients to propose a core allocation scheme.

PROPOSITION 2. *For any reference distribution $\mathbb{P}_{\text{R}}$ and any Wasserstein radius $\theta \geq 0$, we have*

$$\mathcal{F}_{\text{W}}(\theta)\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] = \mathbb{P}_{\text{R}}\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] + \frac{\theta \|\boldsymbol{e}_S\|_*}{\varepsilon}.$$

Proposition 2 says that the coalitional worst-case CVaR amounts to the nominal coalitional CVaR at the same risk level under the reference distribution plus a product term involving *(i)* the Wasserstein radius θ, *(ii)* the dual norm $\|\boldsymbol{e}_S\|_*$ and *(iii)* the risk level ε. In a recent work, Bartl et al. (2020) studied computational aspects of the worst-case optimized certainty equivalents of a single dimensional uncertain loss and arrived at the same closed-form relationship for CVaR as in Proposition 2 when n, *i.e.*, the number of agents, is one. Another interesting result is that the coalitional worst-case CVaR in Proposition 2 has exactly the same expression as the coalitional worst-case VaR in Theorem 6. Therefore, the coalitional worst-case VaR and CVaR coincide under the Wasserstein ambiguity set $\mathcal{F}_{\text{W}}(\theta)$ around the empirical distribution satisfying conditions in Theorem 6. We summarize this interesting observation as below.

COROLLARY 3. *Under either condition in Theorem 6, for any $S \subseteq N$, the coalitional worst-case VaR of the aggregate random cost $\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}$ coincides with its worst-case CVaR.*

For the coalitional worst-case CVaR under Wasserstein ambiguity set with an empirical reference distribution, we propose the following core allocation scheme.

Theorem 8. *Given a Wasserstein ambiguity set $\mathcal{F}_{\rm W}(\theta)$ around the empirical distribution, for any $S \subseteq N$, the coalitional worst-case CVaR can be determined by the convex optimization problem*

$$\begin{aligned} \overline{C}(S) = \mathcal{F}_{\rm W}(\theta)\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] = \sup_{\boldsymbol{\lambda},\boldsymbol{\eta},\boldsymbol{y},\boldsymbol{z}} \;& \frac{1}{\varepsilon}\sum_{k\in[K]} \boldsymbol{e}_S^\top \boldsymbol{y}_k \\ \text{s.t.}\;\; & \sum_{k\in[K]} \lambda_k = \varepsilon \\ & \sum_{k\in[K]} \left(\|\boldsymbol{y}_k - \lambda_k \hat{\boldsymbol{x}}_k\| + \|\boldsymbol{z}_k - \eta_k \hat{\boldsymbol{x}}_k\| \right) \le \theta \\ & \lambda_k + \eta_k = \frac{1}{K} \qquad \forall k\in[K] \\ & \lambda_k, \eta_k \ge 0,\;\; \boldsymbol{y}_k, \boldsymbol{z}_k \in \mathbb{R}^K \qquad \forall k \in [K]. \end{aligned} \tag{7}$$

The allocation scheme $\boldsymbol{a} = (a_i)_{i\in N} = (\sum_{k\in[K]} y^\star_{Nk,i})_{i\in N}$, where $(\boldsymbol{y}^\star_{N1},\dots,\boldsymbol{y}^\star_{NK})$ is (part of) the optimal solution to problem (7) with $S = N$, belongs to the core of the cooperative game $\langle N, \bar{C}(\cdot)\rangle$.

Following from Theorem 7, for the coalitional worst-case CVaR under Wasserstein ambiguity set with an elliptical reference distribution, we propose the following core allocation scheme.

Theorem 9. *Suppose that the conditions in Theorem 7 hold.*

(i) The coalitional worst-case CVaR at the risk level $\varepsilon \in (0, 1/2]$ is subadditive and is given by

$$\overline{C}(S) = \mathcal{F}_{\rm W}(\theta)\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] = \boldsymbol{e}_S^\top \boldsymbol{\mu} + \bar{\eta}^\star \sqrt{\boldsymbol{e}_S^T \boldsymbol{\Sigma} \boldsymbol{e}_S} + \frac{\theta \|\boldsymbol{e}_S\|_*}{\varepsilon} \quad \forall S \subseteq N,$$

where $\bar{\eta}^\star$ is defined as

$$\bar{\eta}^\star = \frac{1}{\varepsilon} \int_{\frac{1}{2}(\Phi^{-1}(1-\varepsilon))^2}^{\infty} k g(z) dz.$$

(ii) Let $\boldsymbol{x}^\star_N$ be the optimal solution to determine $\|\boldsymbol{e}_N\|_* = \max\{\boldsymbol{e}_N^\top \boldsymbol{x} \mid \|\boldsymbol{x}\| \le 1\}$. Then the allocation scheme $\boldsymbol{a} = (a_i)_{i\in N}$ with

$$a_i = \mu_i + \bar{\eta}^\star \cdot \frac{\boldsymbol{e}_{\{i\}}^\top \boldsymbol{\Sigma} \boldsymbol{e}_N}{\sqrt{\boldsymbol{e}_N^\top \boldsymbol{\Sigma} \boldsymbol{e}_N}} + \frac{\theta}{\varepsilon} \cdot x^\star_{N,i} \quad \forall i \in N$$

belongs to the core of the cooperative game $\langle N, \bar{C}(\cdot)\rangle$.

6. Conclusion

Using the cooperative game framework, we study the problem of sharing the worst-case value-at-risk of an aggregate risky position of the sum of some random costs, for which the true joint distribution is ambiguous and is only partially characterized by certain distributional information. We show that the game is subadditive under several different assumptions on the available distributional information and for all these assumptions, we propose accordingly easy-to-compute core allocation schemes. Our results that reveal the intimate relation between the worst-case VaR and the worst-case CVaR shed light on analyzing the problem of sharing the worst-case CVaR under distributional ambiguity and they may be found of independent interest. For future research, it would be interesting to explore the extension of sharing general risk measures under ambiguity.

7. Proofs

7.1. Proof for Section 3

Proof of Theorem 1. According to the minimax theory of semi-infinite programs (see Proposition 2.3 in Shapiro and Kleywegt 2002), the worst-case CVaR is equivalent to

$$\mathcal{F}\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] = \sup_{\mathbb{P}\in\mathcal{F}} \mathbb{P}\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] = \inf_{\tau\in\mathbb{R}} \left\{ \tau + \sup_{\mathbb{P}\in\mathcal{F}} \mathbb{E}_{\mathbb{P}}\left[\frac{(\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}} - \tau)^+}{\varepsilon} \right] \right\}.$$

We use standard techniques in distributionally robust optimization (see, *e.g.*, Wiesemann et al. 2014, Xie and Ahmed 2018) to reformulate the inner maximization problem over $\mathbb{P}$ subject to the convex moment ambiguity set $\mathcal{F}_{\mathrm{M}}$:

$$\sup_{\mathbb{P}\in\mathcal{F}_{\mathrm{M}}} \mathbb{E}_{\mathbb{P}}[(\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}} - \tau)^+],$$

which can be expressed as the *moment problem*

$$\begin{aligned}
\sup\ & \int_{\boldsymbol{x}\in\mathcal{X}} (\boldsymbol{e}_S^\top \boldsymbol{x} - \tau)^+ \mathbb{P}(\mathrm{d}\boldsymbol{x}) \\
\text{s.t.}\ & \int_{\boldsymbol{x}\in\mathcal{X}} \mathbb{P}(\mathrm{d}\boldsymbol{x}) = 1 \\
& \int_{\boldsymbol{x}\in\mathcal{X}} \boldsymbol{x}\mathbb{P}(\mathrm{d}\boldsymbol{x}) = \boldsymbol{\mu} \\
& \int_{\boldsymbol{x}\in\mathcal{X}} \boldsymbol{g}(\boldsymbol{x})\mathbb{P}(\mathrm{d}\boldsymbol{x}) \le \boldsymbol{\nu} \\
& \mathbb{P} \in \mathcal{P}_0(\mathbb{R}^n).
\end{aligned}$$

This is a semi-infinite optimizaiton problem: finitely many (moment) constraints, but infinitely many decision variables (optimization over the probability measure $\mathbb{P}$). When $\boldsymbol{\mu} \in \operatorname{int}\mathcal{X}$, the moment problem attains the same optimal value as its dual problem:

$$\begin{aligned}
\inf\ & \alpha + \boldsymbol{\mu}^\top\boldsymbol{\beta} + \boldsymbol{\nu}^\top\boldsymbol{\gamma} \\
\text{s.t.}\ & \alpha + \boldsymbol{x}^\top\boldsymbol{\beta} + \boldsymbol{g}(\boldsymbol{x})^\top\boldsymbol{\gamma} \ge (\boldsymbol{e}_S^\top\boldsymbol{x} - \tau)^+ \quad \forall \boldsymbol{x}\in\mathcal{X} \\
& \alpha\in\mathbb{R},\ \boldsymbol{\beta}\in\mathbb{R}^n,\ \boldsymbol{\gamma}\in\mathbb{R}^m_+.
\end{aligned}$$

Note that the semi-infinite constraint can be seperated into two:

$$\alpha + \boldsymbol{x}^\top\boldsymbol{\beta} + \boldsymbol{g}(\boldsymbol{x})^\top\boldsymbol{\gamma} \ge \boldsymbol{e}_S^\top\boldsymbol{x} - \tau \quad \forall \boldsymbol{x}\in\mathcal{X} \quad \text{and} \quad \alpha + \boldsymbol{x}^\top\boldsymbol{\beta} + \boldsymbol{g}(\boldsymbol{x})^\top\boldsymbol{\gamma} \ge 0 \quad \forall \boldsymbol{x}\in\mathcal{X}.$$

Combining with the outer minimization problem, the worst-case CVaR, $\mathcal{F}_{\mathrm{M}}\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}]$ can now be reformulated as an optimization problem

$$\begin{aligned}
\inf_{\alpha,\boldsymbol{\beta},\boldsymbol{\gamma},\tau}\ & \tau + \frac{1}{\varepsilon}(\alpha + \boldsymbol{\mu}^\top\boldsymbol{\beta} + \boldsymbol{\nu}^\top\boldsymbol{\gamma}) \\
\text{s.t.}\ & \alpha + \boldsymbol{x}^\top\boldsymbol{\beta} + \boldsymbol{g}(\boldsymbol{x})^\top\boldsymbol{\gamma} \ge \boldsymbol{e}_S^\top\boldsymbol{x} - \tau \quad \forall \boldsymbol{x}\in\mathcal{X} \\
& \alpha + \boldsymbol{x}^\top\boldsymbol{\beta} + \boldsymbol{g}(\boldsymbol{x})^\top\boldsymbol{\gamma} \ge 0 \quad \forall \boldsymbol{x}\in\mathcal{X} \\
& \alpha\in\mathbb{R},\ \boldsymbol{\beta}\in\mathbb{R}^n,\ \boldsymbol{\gamma}\in\mathbb{R}^m_+,\ \tau\in\mathbb{R}.
\end{aligned} \tag{8}$$

The worst-case VaR, $\mathcal{F}_{\mathrm{M}}\text{-VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}]$, by its definition, amounts to the optimal value of the optimization problem

$$\begin{aligned} \inf_{v} \ & v \\ \text{s.t.} \ & \sup_{\mathbb{P}\in\mathcal{F}_{\mathrm{M}}} \mathbb{P}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}} > v] \le \varepsilon \\ & v \in \mathbb{R}. \end{aligned}$$

Likewise, the maximization problem in the constraint, can be expressed as a moment problem

$$\begin{aligned} \sup \ & \int_{\boldsymbol{x}\in\mathcal{X}} \mathbb{I}[\boldsymbol{e}_S^\top \boldsymbol{x} > v]\mathbb{P}(\mathrm{d}\boldsymbol{x}) \\ \text{s.t.} \ & \int_{\boldsymbol{x}\in\mathcal{X}} \mathbb{P}(\mathrm{d}\boldsymbol{x}) = 1 \\ & \int_{\boldsymbol{x}\in\mathcal{X}} \boldsymbol{x}\mathbb{P}(\mathrm{d}\boldsymbol{x}) = \boldsymbol{\mu} \\ & \int_{\boldsymbol{x}\in\mathcal{X}} \boldsymbol{g}(\boldsymbol{x})\mathbb{P}(\mathrm{d}\boldsymbol{x}) \le \boldsymbol{\nu} \\ & \mathbb{P} \in \mathcal{P}_0(\mathbb{R}^n) \end{aligned}$$

over an indicator function $\mathbb{I}[\cdot]$. The dual problem is a minimization problem that takes

$$\begin{aligned} \inf \ & \alpha + \boldsymbol{\mu}^\top \boldsymbol{\beta} + \boldsymbol{\nu}^\top \boldsymbol{\gamma} \\ \text{s.t.} \ & \alpha + \boldsymbol{x}^\top \boldsymbol{\beta} + \boldsymbol{g}(\boldsymbol{x})^\top \boldsymbol{\gamma} \ge \mathbb{I}[\boldsymbol{e}_S^\top \boldsymbol{x} > v] \quad \forall \boldsymbol{x} \in \mathcal{X} \\ & \alpha \in \mathbb{R},\ \boldsymbol{\beta} \in \mathbb{R}^n,\ \boldsymbol{\gamma} \in \mathbb{R}^m_+, \end{aligned}$$

for which the semi-infinite constraint can be expanded into two:

$$\alpha + \boldsymbol{x}^\top \boldsymbol{\beta} + \boldsymbol{g}(\boldsymbol{x})^\top \boldsymbol{\gamma} \ge 1 \quad \forall \boldsymbol{x} \in \mathcal{X} : \boldsymbol{e}_S^\top \boldsymbol{x} - v \quad \text{and} \quad \alpha + \boldsymbol{x}^\top \boldsymbol{\beta} + \boldsymbol{g}(\boldsymbol{x})^\top \boldsymbol{\gamma} \ge 0 \quad \forall \boldsymbol{x} \in \mathcal{X}.$$

Substituting the dual problem back into the constraint, the optimization problem that determines the worst-case VaR can be reformulated as the following optimization problem

$$\begin{aligned} \inf_{\alpha,\boldsymbol{\beta},\boldsymbol{\gamma},v} \ & v \\ \text{s.t.} \ & \alpha + \boldsymbol{\mu}^\top \boldsymbol{\beta} + \boldsymbol{\nu}^\top \boldsymbol{\gamma} \le \varepsilon \\ & \alpha + \boldsymbol{x}^\top \boldsymbol{\beta} + \boldsymbol{g}(\boldsymbol{x})^\top \boldsymbol{\gamma} \ge 1 && \forall \boldsymbol{x} \in \mathcal{X} : \boldsymbol{e}_S^\top \boldsymbol{x} > v \\ & \alpha + \boldsymbol{x}^\top \boldsymbol{\beta} + \boldsymbol{g}(\boldsymbol{x})^\top \boldsymbol{\gamma} \ge 0 && \forall \boldsymbol{x} \in \mathcal{X} \\ & \alpha \in \mathbb{R},\ \boldsymbol{\beta} \in \mathbb{R}^n,\ \boldsymbol{\gamma} \in \mathbb{R}^m_+,\ v \in \mathbb{R}. \end{aligned}$$

The second constraint is satisfied if and only if the following constraint system is infeasible:

$$\begin{cases} \alpha + \boldsymbol{x}^\top \boldsymbol{\beta} + \boldsymbol{g}(\boldsymbol{x})^\top \boldsymbol{\gamma} - 1 < 0 \\ -\boldsymbol{e}_S^\top \boldsymbol{x} + v < 0 \\ \boldsymbol{x} \in \mathcal{X}. \end{cases}$$

Following from the alternative theorem for convex programs (Ben-Tal and Nemirovski 2001), there exist $\lambda, \theta \geq 0$ such that *(i)* λ, θ are not all zero and *(ii)*

$$\lambda(\alpha + \boldsymbol{x}^\top \boldsymbol{\beta} + \boldsymbol{g}(\boldsymbol{x})^\top \boldsymbol{\gamma} - 1) + \theta(-\boldsymbol{e}_S^\top \boldsymbol{x} + v) \geq 0 \quad \forall \boldsymbol{x} \in \mathcal{X}.$$

Since $\mathcal{X}$ is compact, there exists $\boldsymbol{x} \in \mathcal{X}$ such that $-\boldsymbol{e}_S^\top \boldsymbol{x} + v < 0$, and we must have $\lambda > 0$. This implies the above constraint (*i.e.*, the second constraint in the worst-case VaR reformulation) is essentially equivalent to an semi-infinite constraint

$$\alpha + \boldsymbol{x}^\top \boldsymbol{\beta} + \boldsymbol{g}(\boldsymbol{x})^\top \boldsymbol{\gamma} - 1 + \theta(-\boldsymbol{e}_S^\top \boldsymbol{x} + v) \geq 0 \quad \forall \boldsymbol{x} \in \mathcal{X}.$$

In fact, $\theta > 0$ because otherwise the above problem would be unbounded and would contradict to the boundedness of the worst-case VaR, which follows from $\mathcal{X}$ being compact. We can thus conduct the substitutions $\alpha \leftarrow \alpha/\theta$, $\boldsymbol{\beta} \leftarrow \boldsymbol{\beta}/\theta$, $\boldsymbol{\gamma} \leftarrow \boldsymbol{\gamma}/\theta$, and $\theta \leftarrow 1/\theta$ to obtain an equivalent reformulation:

$$\begin{aligned} \inf_{\alpha,\boldsymbol{\beta},\boldsymbol{\gamma},v,\theta} \quad & v \\ \text{s.t.} \quad & \alpha + \boldsymbol{\mu}^\top \boldsymbol{\beta} + \boldsymbol{\nu}^\top \boldsymbol{\gamma} \leq \theta\varepsilon \\ & \alpha + \boldsymbol{x}^\top \boldsymbol{\beta} + \boldsymbol{g}(\boldsymbol{x})^\top \boldsymbol{\gamma} \geq \theta + \boldsymbol{e}_S^\top \boldsymbol{x} - v && \forall \boldsymbol{x} \in \mathcal{X} \\ & \alpha + \boldsymbol{x}^\top \boldsymbol{\beta} + \boldsymbol{g}(\boldsymbol{x})^\top \boldsymbol{\gamma} \geq 0 && \forall \boldsymbol{x} \in \mathcal{X} \\ & \alpha \in \mathbb{R},\ \boldsymbol{\beta} \in \mathbb{R}^n,\ \boldsymbol{\gamma} \in \mathbb{R}^m_+,\ v \in \mathbb{R},\ \theta \in \mathbb{R}_+. \end{aligned}$$

Observe that for any $\mathbb{P} \in \mathcal{F}_{\mathrm{M}}$, $\alpha + \boldsymbol{\mu}^\top \boldsymbol{\beta} + \boldsymbol{\nu}^\top \boldsymbol{\gamma} \geq \alpha + \mathbb{E}_{\mathbb{P}}[\tilde{\boldsymbol{x}}]^\top \boldsymbol{\beta} + \mathbb{E}_{\mathbb{P}}[\boldsymbol{g}(\tilde{\boldsymbol{x}})]^\top \boldsymbol{\gamma} \geq 0$. Hence we can define a new variable $\tau = v - \theta$ and take the optimal choice $\theta = (\alpha + \boldsymbol{\mu}^\top \boldsymbol{\beta} + \boldsymbol{\nu}^\top \boldsymbol{\gamma})/\varepsilon$. Then the above reformulation of the worst-case VaR is exactly the same as that of the worst-case CVaR. □

Proof of Theorem 2. By Theorem 1, for a convex moment ambiguity set $\mathcal{F}_{\mathrm{M}}$ and for any $S \subseteq N$, the coalitional worst-case VaR, $C(S)$ equals the optimal value of the following reformulation of worst-case CVaR:

$$\begin{aligned} C(S) = \inf_{\alpha,\boldsymbol{\beta},\boldsymbol{\gamma},\tau} \quad & \tau + \frac{1}{\varepsilon}\left(\alpha + \boldsymbol{\mu}^\top \boldsymbol{\beta} + \boldsymbol{\nu}^\top \boldsymbol{\gamma}\right) \\ \text{s.t.} \quad & \alpha + \boldsymbol{x}^\top \boldsymbol{\beta} + \boldsymbol{g}(\boldsymbol{x})^\top \boldsymbol{\gamma} \geq \boldsymbol{e}_S^\top \boldsymbol{x} - \tau && \forall \boldsymbol{x} \in \mathcal{X} \\ & \alpha + \boldsymbol{y}^\top \boldsymbol{\beta} + \boldsymbol{g}(\boldsymbol{y})^\top \boldsymbol{\gamma} \geq 0 && \forall \boldsymbol{y} \in \mathcal{X} \\ & \alpha \in \mathbb{R},\ \boldsymbol{\beta} \in \mathbb{R}^n,\ \boldsymbol{\gamma} \in \mathbb{R}^m_+,\ \tau \in \mathbb{R}. \end{aligned} \tag{9}$$

The above representation can be interpreted as the robust counterpart of an uncertain convex program with constraint-wise uncertainty, and can be solved by a decision maker choosing α, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ and τ in the face of the worst possible outcomes $\boldsymbol{x}$, $\boldsymbol{y} \in \mathcal{X}$. According to Theorem 4.1. in Beck and

Ben-Tal (2009) and the facts that $\mathcal{X}$ is convex and compact and $\boldsymbol{g}(\cdot)$ is convex, the optimal value of problem (9) coincides with its dual problem:

$$\begin{aligned} C(S) = \sup_{\lambda_1,\lambda_2,\boldsymbol{x},\boldsymbol{y}} \inf_{\alpha,\boldsymbol{\beta},\boldsymbol{\gamma},\tau} \ & D(\alpha,\boldsymbol{\beta},\boldsymbol{\gamma},\tau,\lambda_1,\lambda_2,\boldsymbol{x},\boldsymbol{y}) \\ \text{s.t.} \ & \lambda_1,\lambda_2 \in \mathbb{R}_+,\ \boldsymbol{x},\boldsymbol{y} \in \mathcal{X} \\ & \alpha \in \mathbb{R},\ \boldsymbol{\beta} \in \mathbb{R}^n,\ \boldsymbol{\gamma} \in \mathbb{R}^m_+,\ \tau \in \mathbb{R}, \end{aligned} \tag{10}$$

where λ_1, λ_2 are dual variables associated with the two constraints of problem (9), respectively, while the objective function takes

$$\begin{aligned} & D(\alpha,\boldsymbol{\beta},\boldsymbol{\gamma},\tau,\lambda_1,\lambda_2,\boldsymbol{x},\boldsymbol{y}) \\ = {} & \tau + \frac{1}{\varepsilon}\big(\alpha + \boldsymbol{\mu}^\top\boldsymbol{\beta} + \boldsymbol{\nu}^\top\boldsymbol{\gamma}\big) - \lambda_1(\alpha + \boldsymbol{x}^\top\boldsymbol{\beta} + \boldsymbol{g}(\boldsymbol{x})^\top\boldsymbol{\gamma} - \boldsymbol{e}_S^\top\boldsymbol{x} + \tau) - \lambda_2(\alpha + \boldsymbol{y}^\top\boldsymbol{\beta} + \boldsymbol{g}(\boldsymbol{y})^\top\boldsymbol{\gamma}). \end{aligned}$$

Optimizing the inner infimum problem in problem (10) leads to

$$\begin{aligned} C(S) = \sup_{\lambda_1,\lambda_2,\boldsymbol{x},\boldsymbol{y}} \ & \lambda_1 \boldsymbol{e}_S^\top \boldsymbol{x} \\ \text{s.t.} \ & \lambda_1\boldsymbol{x} + \lambda_2\boldsymbol{y} = \frac{1}{\varepsilon}\boldsymbol{\mu} \\ & \lambda_1\boldsymbol{g}(\boldsymbol{x}) + \lambda_2\boldsymbol{g}(\boldsymbol{x}) \le \frac{1}{\varepsilon}\boldsymbol{\nu} \\ & \lambda_1 + \lambda_2 = \frac{1}{\varepsilon},\ \lambda_1 = 1 \\ & \lambda_1,\lambda_2 \in \mathbb{R}_+,\ \boldsymbol{x},\boldsymbol{y} \in \mathcal{X}. \end{aligned}$$

The above problem is exactly equal to problem (1), which is understood as the dual of another uncertain convex program where the decision maker benefits from the best possible outcomes. The equivalence between the minimization reformulation (9) and the maximization reformulation (1) is actually a nice duality property called “primal worst equals dual best” in robust optimization; see its origin in Beck and Ben-Tal (2009).

Let $(\boldsymbol{x}_N^\star, \boldsymbol{y}_N^\star)$ be an optimal solution to problem (1) in correspondence to the grand coalition $S = N$. The defined allocation scheme $\boldsymbol{a} = \boldsymbol{x}_N^\star$ satisfies

$$\sum_{i\in N} a_i = \sum_{i\in N} x_{N,i}^\star = C(N).$$

In addition, since the feasible region of problem (1) is independent of S, the solution $(\boldsymbol{x}_N^\star, \boldsymbol{y}_N^\star)$ is feasible to problem (1) for any coalition $S \subseteq N$. We thus have

$$\sum_{i\in S} a_i = \sum_{i\in S} x_{N,i}^\star \le C(S) \quad \forall S \subseteq N.$$

Therefore, the allocation scheme belongs to the core. □

Proof of Theorem 3. Since the marginal convex moment ambiguity set is a special convex moment ambiguity set, it holds that $\mathcal{F}_{\mathrm{M}}\text{-VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}] \le \sum_{i\in S}\mathcal{F}_{\mathrm{M}}\text{-VaR}_{1-\varepsilon}[\tilde{x}_i]$. It is then sufficient to show

$\mathcal{F}_{\rm M}$-$\text{VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}] \geq \sum_{i\in S}\mathcal{F}_{\rm M}$-$\text{VaR}_{1-\varepsilon}[\tilde{x}_i]$. To this end, we consider a subset of the marginal convex moment ambiguity set defined by $\underline{\mathcal{F}}_{\rm M} = \mathcal{F}_{\rm M}\cap\mathcal{F}_{\rm C}$ with $\mathcal{F}_{\rm C} = \{\mathbb{P}\mid\mathbb{P}\text{ is comonotonic}\}$. Clearly we have $\mathcal{F}_{\rm M}$-$\text{VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}] \geq \underline{\mathcal{F}}_{\rm M}$-$\text{VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}]$. Note that we can represent the marginal convex moment ambiguity set as the product of individual ambiguity sets, $\mathcal{F}_{\rm M} = \times_{i\in N}\mathcal{F}_i$, where each individual set

$$\mathcal{F}_i = \{\mathbb{P}\in\mathcal{P}_0(\mathbb{R}^n)\mid\mathbb{P}[\underline{x}_i\leq\tilde{x}_i\leq\overline{x}_i]=1,\ \mathbb{E}_{\mathbb{P}}[\tilde{x}_i]=\mu_i,\ \mathbb{E}_{\mathbb{P}}[\boldsymbol{g}_i(\tilde{x}_i)]\leq\boldsymbol{\nu}_i\}$$

specifies only distributional information of the $i^{\rm th}$ agent's random cost $\tilde{x}_i$. Hence, we can represent the smaller ambiguity set $\underline{\mathcal{F}}_{\rm M}$ as $(\times_{i\in N}\mathcal{F}_i)\cap\mathcal{F}_{\rm C}$.

Let $\mathcal{Q}(\mathbb{P}_1,\ldots,\mathbb{P}_n)$ be set of all joint distributions generated by marginal ones $\mathbb{P}_1,\ldots,\mathbb{P}_n$ with comonotonicity property. We can express the coalitional worst-case VaR over $\underline{\mathcal{F}}_{\rm M}$ as follows:

$$\begin{aligned}
\underline{\mathcal{F}}_{\rm M}\text{-VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}] &= \sup_{\mathbb{P}\in\underline{\mathcal{F}}_{\rm M}}\mathbb{P}\text{-VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}]\\
&= \sup_{\mathbb{P}_i\in\mathcal{F}_i\ \ \forall i\in N}\ \sup_{\mathbb{P}\in\mathcal{Q}(\mathbb{P}_1,\ldots,\mathbb{P}_n)}\mathbb{P}\text{-VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}]\\
&= \sup_{\mathbb{P}_i\in\mathcal{F}_i\ \ \forall i\in N}\sum_{i\in S}\mathbb{P}_i\text{-VaR}_{1-\varepsilon}[\tilde{x}_i]\\
&= \sum_{i\in S}\sup_{\mathbb{P}_i\in\mathcal{F}_i}\mathbb{P}_i\text{-VaR}_{1-\varepsilon}[\tilde{x}_i]\\
&= \sum_{i\in S}\sup_{\mathbb{P}_i\in\mathcal{F}_{\rm M}}\mathbb{P}_i\text{-VaR}_{1-\varepsilon}[\tilde{x}_i]\\
&= \sum_{i\in S}\mathcal{F}_{\rm M}\text{-VaR}_{1-\varepsilon}[\tilde{x}_i].
\end{aligned}$$

The third equality follows from the additivity of VaR under a comonotonic distribution (see Proposition 6.15 in McNeil et al. 2005). The fourth one is because that ambiguity sets $\mathcal{F}_i$, $i\in N$ are separable. The fifth one is because that for the random cost $\tilde{x}_i$, the marginal distributional information of other random costs are redundant. The last one is due to definition of the worst-case VaR. Finally, we can establish

$$\mathcal{F}_{\rm M}\text{-VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}]\geq\underline{\mathcal{F}}_{\rm M}\text{-VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}]=\sum_{i\in S}\mathcal{F}_{\rm M}\text{-VaR}_{1-\varepsilon}[\tilde{x}_i]\geq\mathcal{F}_{\rm M}\text{-VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}].$$

Thus both inequalities in the above relation are indeed equalities. This concludes the proof. □

7.2. Proof for Section 4

Proof of Theorem 4. If $\theta=0$, the first constraint in (3) is redundant and $v^\star=\mathbb{P}_{\rm R}$-$\text{VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}]$. Moreover, the Wasserstein ambiguity set is a singleton set containing only the reference distribution $\mathbb{P}_{\rm R}$. Hence $C(S)=\mathcal{F}_{\rm W}(0)$-$\text{VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}]=\mathbb{P}_{\rm R}$-$\text{VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}]$. Thus, the statement follows.

We next focus on $\theta>0$. Note that $\mathcal{F}_{\rm W}(\theta)$-$\text{VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}]$ is determined by the optimization problem

$$\inf_{v\in\mathbb{R}}\left\{v\;\middle|\;\sup_{\mathbb{P}\in\mathcal{F}_{\rm W}(\theta)}\mathbb{P}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}>v]\leq\varepsilon\right\}.$$

Using Theorem 1 in Gao and Kleywegt (2016) or Theorem 1 in Blanchet and Murthy (2019), the distributionally robust chance constraint $\sup_{\mathbb{P}\in\mathcal{F}_{\mathrm{W}}(\theta)}\mathbb{P}[\boldsymbol{e}_S^\top\boldsymbol{x}>v]\le\varepsilon$ can be reformulated as

$$\min_{\lambda\ge 0}\left\{\lambda\theta-\int_{\boldsymbol{x}}\inf_{\boldsymbol{y}}\left\{\lambda\|\boldsymbol{y}-\boldsymbol{x}\|-\mathbb{I}[\boldsymbol{e}_S^\top\boldsymbol{y}>v]\right\}\mathbb{P}_{\mathrm{R}}(\mathrm{d}\boldsymbol{x})\right\}\le\varepsilon. \tag{11}$$

We will implement Theorem 1 and Corollary 1 in Xie (2018) to analyze inequality (11). It holds that $\lambda>0$ because otherwise the objective of the right-hand side problem amounts to 1 and contradicts to $\varepsilon\in(0,1)$. Consider the second term in the right-hand side problem, we have

$$\inf_{\boldsymbol{y}}\left\{\lambda\|\boldsymbol{y}-\boldsymbol{x}\|-\mathbb{I}[\boldsymbol{e}_S^\top\boldsymbol{y}>v]\right\}=-\left(\lambda\cdot\frac{(v-\boldsymbol{e}_S^\top\boldsymbol{x})^+}{\|\boldsymbol{e}_S\|_*}-1\right)^-.$$

Indeed, if $\boldsymbol{e}_S^\top\boldsymbol{x}>v$, we can let $\boldsymbol{y}=\boldsymbol{x}$ so that

$$\inf_{\boldsymbol{y}}\left\{\lambda\|\boldsymbol{y}-\boldsymbol{x}\|-\mathbb{I}[\boldsymbol{e}_S^\top\boldsymbol{y}>v]\right\}=-1=-\left(\lambda\cdot\frac{(v-\boldsymbol{e}_S^\top\boldsymbol{x})^+}{\|\boldsymbol{e}_S\|_*}-1\right)^-;$$

whereas if $\boldsymbol{e}_S^\top\boldsymbol{x}\le v$, we have

$$\begin{aligned}\inf_{\boldsymbol{y}}\left\{\lambda\|\boldsymbol{y}-\boldsymbol{x}\|-\mathbb{I}[\boldsymbol{e}_S^\top\boldsymbol{y}>v]\right\}&=\min\left\{\inf_{\boldsymbol{y}:\boldsymbol{e}_S^\top\boldsymbol{y}>v}\left\{\lambda\|\boldsymbol{y}-\boldsymbol{x}\|-1\right\},\inf_{\boldsymbol{y}:\boldsymbol{e}_S^\top\boldsymbol{y}\le v}\lambda\|\boldsymbol{y}-\boldsymbol{x}\|\right\}\\&=\min\left\{\inf_{\boldsymbol{y}:\boldsymbol{e}_S^\top\boldsymbol{y}>v}\left\{\lambda\|\boldsymbol{y}-\boldsymbol{x}\|-1\right\},0\right\}\\&=-\left(\inf_{\boldsymbol{y}:\boldsymbol{e}_S^\top\boldsymbol{y}>v}\left\{\lambda\|\boldsymbol{y}-\boldsymbol{x}\|-1\right\}\right)^-\\&=-\left(\lambda\cdot\frac{(v-\boldsymbol{e}_S^\top\boldsymbol{x})^+}{\|\boldsymbol{e}_S\|_*}-1\right)^-.\end{aligned}$$

Here in the last equality, for any norm $\|\cdot\|$ and $\lambda>0$ we note the following closed-form expression of the optimization problem (see, *e.g.*, Lemma 2 in Chen et al. 2018):

$$\inf_{\boldsymbol{y}:\boldsymbol{e}_S^\top\boldsymbol{y}>v}\lambda\|\boldsymbol{y}-\boldsymbol{x}\|=\lambda\cdot\frac{(v-\boldsymbol{e}_S^\top\boldsymbol{x})^+}{\|\boldsymbol{e}_S\|_*}.$$

Hence, we can rewrite inequality (11) based on the above expression. Accordingly, the worst-case probability constraint in the definition of the worst-case VaR becomes

$$\min_{\lambda>0}\left\{\lambda\theta+\int_{\boldsymbol{x}}\left(\lambda\cdot\frac{(v-\boldsymbol{e}_S^\top\boldsymbol{x})^+}{\|\boldsymbol{e}_S\|_*}-1\right)^-\mathrm{d}\mathbb{P}_{\mathrm{R}}(\boldsymbol{x})\right\}\le\varepsilon.$$

Multiplying both sides by $(\lambda\varepsilon)^{-1}>0$ and letting $\tau=-1/\lambda$, we arrive at

$$\frac{\theta}{\varepsilon}+\min_{\tau\in\mathbb{R}}\left\{\frac{1}{\varepsilon}\int_{\boldsymbol{x}}\left(-\frac{(v-\boldsymbol{e}_S^\top\boldsymbol{x})^+}{\|\boldsymbol{e}_S\|_*}-\tau\right)^+\mathrm{d}\mathbb{P}_{\mathrm{R}}(\boldsymbol{x})+\tau\right\}\le 0$$

because no $\tau\ge 0$ is feasible. Multiplying both sides by $\|\boldsymbol{e}_S\|_*$ and replacing $\tau\|\boldsymbol{e}_S\|_*$ by τ, we have

$$\frac{\theta}{\varepsilon}\|\boldsymbol{e}_S\|_*+\min_{\tau\in\mathbb{R}}\left\{\frac{1}{\varepsilon}\int_{\boldsymbol{x}}\left(-(v-\boldsymbol{e}_S^\top\boldsymbol{x})^+-\tau\right)^+\mathbb{P}_{\mathrm{R}}(\mathrm{d}\boldsymbol{x})+\tau\right\}\le 0,$$

which is essentially

$$\frac{\theta}{\varepsilon}\|\boldsymbol{e}_S\|_* + \mathbb{P}_{\mathrm{R}}\text{-CVaR}_{1-\varepsilon}[-(v-\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}})^+] \le 0.$$

In fact, any v satisfying this constraint must be no smaller than $\mathbb{P}_{\mathrm{R}}\text{-VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}]$. □

Proof of Proposition 1. The statement trivially holds when $\theta = 0$, thus we will focus on $\theta > 0$. For the empirical reference distribution $\mathbb{P}_{\mathrm{R}} = \frac{1}{K}\sum_{k\in[K]}\delta_{\hat{\boldsymbol{x}}_k}$, according to Theorem 4, we have

$$\begin{aligned} C(S) &= \inf_{v\in\mathbb{R}}\left\{ v \;\middle|\; -\mathbb{P}_{\mathrm{R}}\text{-CVaR}_{1-\varepsilon}[-(v-\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}})^+] \ge \frac{\theta}{\varepsilon}\|\boldsymbol{e}_S\|_* \right\} \\ &= \inf_{v\in\mathbb{R}}\left\{ v \;\middle|\; -\min_{\tau\in\mathbb{R}}\left\{ \frac{1}{\varepsilon}\sum_{k\in[K]}\frac{1}{K}\big(-(v-\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_k)^+ - \tau\big)^+ + \tau \right\} \ge \frac{\theta}{\varepsilon}\|\boldsymbol{e}_S\|_* \right\}. \end{aligned}$$

Because $\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_{\pi_1(S)} \ge \boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_{\pi_2(S)} \ge \cdots \ge \boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_{\pi_K(S)}$, for any $v\in\mathbb{R}$ we have $-(v-\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_{\pi_1(S)})^+ \ge -(v-\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_{\pi_2(S)})^+ \ge \cdots \ge -(v-\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_{\pi_K(S)})^+$. By Theorem 10 in Rockafellar and Uryasev (2002), an optimal $\tau^\star$ to the inner minimization problem is $\tau^\star = -(v-\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_{\pi_\kappa(S)})^+$ with $\kappa = \lfloor\varepsilon K\rfloor + 1$. The statement then follows from substituting such an optimal $\tau^\star$ into that inner minimization problem. □

Proof of Theorem 5. When the empirical reference distribution is comonotonic, the permutation $\boldsymbol{\pi}(S)$ (defined in Proposition 1) reordering the empirical realizations is indeed independent of the coalition S. That is to say, we can assume $\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_1 \ge \boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_2 \ge \cdots \ge \boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_K$ without loss of generality.

When θ is small as stated in the theorem, following from (5), we have $j^\star = 1$ and thus the expression (6) holds. To show the subadditivity of $C(S)$, it is sufficient to show that $\|\boldsymbol{e}_S\|_*$ is subadditive because $(\theta K)/(1+\varepsilon K-\kappa) \ge 0$ and $\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_\kappa$ is additive. Indeed, for any two disjoint sets $S, T\subseteq N$, since $\boldsymbol{e}_{S\cup T} = \boldsymbol{e}_S + \boldsymbol{e}_T$, by triangle inequality we have $\|\boldsymbol{e}_{S\cup T}\|_* \le \|\boldsymbol{e}_S\|_* + \|\boldsymbol{e}_T\|_*$.

The defined allocation scheme $\boldsymbol{a}$ satisfies $\sum_{i\in N} a_i = C(N)$. Observe that for any $S\subseteq N$, $\boldsymbol{x}_N^\star$ is also feasible to the optimization problem $\|\boldsymbol{e}_S\|_* \triangleq \max\{\boldsymbol{e}_S^\top\boldsymbol{x} \mid \|\boldsymbol{x}\| \le 1\}$. This implies that for all $S\subseteq N$, $\sum_{i\in S} x_{N,i}^\star \le \|\boldsymbol{e}_S\|_*$. Thus, we have

$$\sum_{i\in S} a_i = \sum_{i\in S}\left(\hat{x}_{\kappa,i} + \frac{\theta K}{1+\varepsilon K-\kappa}x_{N,i}^\star\right) \le \boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_\kappa + \frac{\theta K}{1+\varepsilon K-\kappa}\|\boldsymbol{e}_S\|_* = C(S) \quad \forall S\subseteq N,$$

where the inequality is due to $(\theta K)/(1+\varepsilon K-\kappa)\ge 0$. This concludes that the proposed allocation scheme belongs to the core. □

Proof of Theorem 6. We prove the statement under condition *(i)* and condition *(ii)* separately.

Suppose that $K < 1/\varepsilon$. In this case, $\kappa = \lfloor\varepsilon K\rfloor + 1 = 1$. Since there is only one interval that $v^\star$ can belong to, we must have $j^\star = 1$ in (5). That is, we have

$$C(S) = \boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_{\pi_1(S)} + \frac{\theta\|\boldsymbol{e}_S\|_*}{\varepsilon} = \mathbb{P}_{\mathrm{R}}\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}] + \frac{\theta\|\boldsymbol{e}_S\|_*}{\varepsilon} \quad \forall S\subseteq N,$$

where the last equality is due to the alternative definition of CVaR under the empirical reference distribution $\mathbb{P}_{\mathrm{R}}$ (see, e.g., Theorem 10 in Rockafellar and Uryasev 2002). Since both $\|\boldsymbol{e}_S\|_*$ and $\mathbb{P}_{\mathrm{R}}$-$\mathrm{CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}]$ are subadditive, so is $C(S)$.

We next look at the second condition where θ is large as stated in the theorem. In this case, $v^\star$ must belong to the largest interval. That is, in (5), for all $S \subseteq N$ we have $j^\star = \kappa$ and

$$C(S) = \frac{1}{\varepsilon K}\left[(\varepsilon K - \kappa)\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_{\pi_\kappa(S)} + \sum_{i\in[\kappa]}\boldsymbol{e}_S^\top\hat{\boldsymbol{x}}_{\pi_{\kappa+1-i}(S)}\right] + \frac{\theta\|\boldsymbol{e}_S\|_*}{\varepsilon} = \mathbb{P}_{\mathrm{R}}\text{-}\mathrm{CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}] + \frac{\theta\|\boldsymbol{e}_S\|_*}{\varepsilon},$$

where the last equality also follows from the alternative definition of CVaR under the empirical reference distribution $\mathbb{P}_{\mathrm{R}}$. Similarly, we can conclude that $C(S)$ is subadditive. □

Proof of Theorem 7. According to Theorem 4 and properties of elliptical distributions, the constraint involving CVaR with $\mathbb{P}_{\mathrm{R}} = \mathbb{P}_{\mathrm{E}}(\boldsymbol{\mu},\boldsymbol{\Sigma},g)$ now becomes

$$-\mathbb{P}_{\mathrm{E}}(\mu_S,\sigma_S,g)\text{-}\mathrm{CVaR}_{1-\varepsilon}[-(v-\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}})^+] \geq \frac{\theta}{\varepsilon}\|\boldsymbol{e}_S\|_*, \tag{12}$$

where $\mu_S = \boldsymbol{e}_S^T\boldsymbol{\mu}$ and $\sigma_S^2 = \boldsymbol{e}_S^T\boldsymbol{\Sigma}\boldsymbol{e}_S$.

For ease of notation, we denote the elliptical distribution $\mathbb{P}_{\mathrm{E}}(\mu_S,\sigma_S,g)$ by $\mathbb{P}_S$ and denote its probability density function as

$$h(y) = \frac{k}{\sigma_S}\cdot g\left(\frac{(y-\mu_S)^2}{2\sigma_S^2}\right).$$

Using the alternative definition of CVaR, the CVaR on the left-hand side of (12) is given by

$$\begin{aligned}
-\mathbb{P}_S\text{-}\mathrm{CVaR}_{1-\varepsilon}[-(v-\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}})^+] &= -\mathbb{E}_{\mathbb{P}_S}[-(v-\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}})^+ \mid -(v-\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}})^+ \geq \mathbb{P}_S\text{-}\mathrm{VaR}_{1-\varepsilon}[-(v-\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}})^+]] \\
&= -\frac{1}{\varepsilon}\int_{\inf\{y \mid -(v-y)^+ \geq \mathbb{P}_S\text{-}\mathrm{VaR}_{1-\varepsilon}[-(v-\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}})^+]\}}^{+\infty} -(v-y)^+ h(y)dy \\
&= \frac{1}{\varepsilon}\int_{\inf\{y \mid -(v-y)^+ \geq \mathbb{P}_S\text{-}\mathrm{VaR}_{1-\varepsilon}[-(v-\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}})^+]\}}^{v} (v-y)h(y)dy \\
&= \frac{1}{\varepsilon}\int_{\mathbb{P}_S\text{-}\mathrm{VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}]}^{v} (v-y)h(y)dy.
\end{aligned} \tag{13}$$

Here, the last equality follows from:

$$\begin{aligned}
&\inf\{y \mid -(v-y)^+ \geq \mathbb{P}_S\text{-}\mathrm{VaR}_{1-\varepsilon}[-(v-\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}})^+]\} \\
\Longleftrightarrow\ &\inf\{y \mid \min\{y-v,0\} \geq \mathbb{P}_S\text{-}\mathrm{VaR}_{1-\varepsilon}[\min\{\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}-v,0\}]\} \\
\Longleftrightarrow\ &\inf\{y \mid \min\{y,v\}-v \geq \mathbb{P}_S\text{-}\mathrm{VaR}_{1-\varepsilon}[\min\{\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}},v\}-v]\} \\
\Longleftrightarrow\ &\inf\{y \mid \min\{y,v\} \geq \mathbb{P}_S\text{-}\mathrm{VaR}_{1-\varepsilon}[\min\{\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}},v\}]\} \\
\Longleftrightarrow\ &\inf\{y \mid y \geq \mathbb{P}_S\text{-}\mathrm{VaR}_{1-\varepsilon}[\min\{\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}},v\}]\} \\
\Longleftrightarrow\ &\inf\{y \mid y \geq \mathbb{P}_S\text{-}\mathrm{VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}]\},
\end{aligned}$$

where the third equivalence follows from the translation invariance of VaR, the fourth one is due to $\mathbb{P}_S$-$\mathrm{VaR}_{1-\varepsilon}[\min\{\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}, v\}]\} \leq v$, and the last one is because that for any risk threshold ε, the worst-case VaR satisfies

$$v = \sup_{\mathbb{P}\in\mathcal{F}_{\mathrm{W}}(\theta)} \mathbb{P}\text{-}\mathrm{VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}] \geq \mathbb{P}_S\text{-}\mathrm{VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}], \tag{14}$$

while the $(1-\varepsilon)$ quantiles of $\mathbb{P}_S$-$\mathrm{VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}]$ and $\mathbb{P}_S$-$\mathrm{VaR}_{1-\varepsilon}[\min\{\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}, v\}]$ coincide.

For notational convenience, let us denote $q_{1-\varepsilon} = \mathbb{P}_S$-$\mathrm{VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}]$, which by its definition, satisfies

$$\frac{q_{1-\varepsilon}-\mu_S}{\sigma_S} = \frac{\mathbb{P}_S\text{-}\mathrm{VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}-\mu_S]}{\sigma_S} = \mathbb{P}_S\text{-}\mathrm{VaR}_{1-\varepsilon}\left[\frac{\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}-\mu_S}{\sigma_S}\right] = \mathbb{P}_{\mathrm{E}}^0\text{-}\mathrm{VaR}_{1-\varepsilon}[\tilde{z}] = \Phi^{-1}(1-\varepsilon),$$

where the first equality follows from the translation invariance of VaR, the second one follows from the positive homogeneity of VaR, and the last one follows from the definition of VaR under the standard elliptical distribution $\mathbb{P}_{\mathrm{E}}^0$.

Following the last expression of CVaR in (13), we further have

$$\frac{1}{\varepsilon}\int_{q_{1-\varepsilon}}^{v}(v-y)h(y)dy = \frac{1}{\varepsilon}\int_{q_{1-\varepsilon}}^{v} v\cdot\frac{k}{\sigma_S}\cdot g\left(\frac{(y-\mu_S)^2}{2\sigma_S^2}\right)dy - \frac{1}{\varepsilon}\int_{q_{1-\varepsilon}}^{v} y\cdot\frac{k}{\sigma_S}\cdot g\left(\frac{(y-\mu_S)^2}{2\sigma_S^2}\right)dy.$$

The first component

$$\int_{q_{1-\varepsilon}}^{v}\frac{v}{\varepsilon}\cdot\frac{k}{\sigma_S}\cdot g\left(\frac{(y-\mu_S)^2}{2\sigma_S^2}\right)dy = \frac{v}{\varepsilon}\int_{\frac{q_{1-\varepsilon}-\mu_S}{\sigma_S}}^{\frac{v-\mu_S}{\sigma_S}} kg\left(\frac{z^2}{2}\right)dz = \frac{v}{\varepsilon}\left(\Phi\left(\frac{v-\mu_S}{\sigma_S}\right)-\Phi\left(\frac{q_{1-\varepsilon}-\mu_S}{\sigma_S}\right)\right),$$

while the second component

$$\begin{aligned}
&\frac{1}{\varepsilon}\int_{q_{1-\varepsilon}}^{v} y\cdot\frac{k}{\sigma_S}\cdot g\left(\frac{(y-\mu_S)^2}{2\sigma_S^2}\right)dy\\
&= \frac{1}{\varepsilon}\int_{q_{1-\varepsilon}}^{v}\frac{(y-\mu_S)}{\sigma_S}\cdot k\cdot g\left(\frac{(y-\mu_S)^2}{2\sigma_S^2}\right)dy + \frac{1}{\varepsilon}int_{q_{1-\varepsilon}}^{v}\mu_S\cdot\frac{k}{\sigma_S}\cdot g\left(\frac{(y-\mu_S)^2}{2\sigma_S^2}\right)dy\\
&= \frac{1}{\varepsilon}\sigma_S\int_{q_{1-\varepsilon}}^{v}\frac{(y-\mu_S)}{\sigma_S}\cdot k\cdot g\left(\frac{(y-\mu_S)^2}{2\sigma_S^2}\right)d\left(\frac{y-\mu_S}{\sigma_S}\right) + \frac{1}{\varepsilon}\mu_S\left(\Phi\left(\frac{v-\mu_S}{\sigma_S}\right)-\Phi\left(\frac{q_{1-\varepsilon}-\mu_S}{\sigma_S}\right)\right)\\
&= \frac{1}{\varepsilon}\sigma_S\int_{\frac{q_{1-\varepsilon}-\mu_S}{\sigma_S}}^{\frac{v-\mu_S}{\sigma_S}} t\cdot k\cdot g\left(\frac{t^2}{2}\right)dt + \frac{1}{\varepsilon}\mu_S\left(\Phi\left(\frac{v-\mu_S}{\sigma_S}\right)-\Phi\left(\frac{q_{1-\varepsilon}-\mu_S}{\sigma_S}\right)\right)\\
&= \frac{1}{\varepsilon}\sigma_S\int_{\frac{(q_{1-\varepsilon}-\mu_S)^2}{\sigma_S^2}}^{\frac{(v-\mu_S)^2}{\sigma_S^2}} kg(z)dz + \frac{1}{\varepsilon}\mu_S\left(\Phi\left(\frac{v-\mu_S}{\sigma_S}\right)-\Phi\left(\frac{q_{1-\varepsilon}-\mu_S}{\sigma_S}\right)\right).
\end{aligned}$$

Hence, we arrive at

$$\frac{1}{\varepsilon}\int_{q_{1-\varepsilon}}^{v}(v-y)h(y)dy = \frac{\sigma_S}{\varepsilon}\left[\frac{(v-\mu_S)}{\sigma_S}\left(\Phi\left(\frac{v-\mu_S}{\sigma_S}\right)-\Phi\left(\frac{q_{1-\varepsilon}-\mu_S}{\sigma_S}\right)\right) - \int_{\frac{(q_{1-\varepsilon}-\mu_S)^2}{2\sigma_S^2}}^{\frac{(v-\mu_S)^2}{2\sigma_S^2}} kg(z)dz\right].$$

Together with constraint (14), the constraint (12) is equivalent to

$$v \geq q_{1-\varepsilon}, \quad \frac{(v-\mu_S)}{\sigma_S}\left(\Phi\left(\frac{v-\mu_S}{\sigma_S}\right) - \Phi\left(\frac{q_{1-\varepsilon}-\mu_S}{\sigma_S}\right)\right) - \int_{\frac{(q_{1-\varepsilon}-\mu_S)^2}{2\sigma_S^2}}^{\frac{(v-\mu_S)^2}{2\sigma_S^2}} kg(z)dz \geq \frac{\varepsilon\theta\|\boldsymbol{e}_S\|_*}{\sigma_S} = \varepsilon\theta, \quad (15)$$

where the equality follows from the fact that the dual norm of the Mahalanobis norm is $\|\boldsymbol{x}\|_* = \sqrt{\boldsymbol{x}^\top\boldsymbol{\Sigma}\boldsymbol{x}}$. Let $\eta = (v-\mu_S)/\sigma_S$. Using (15), we can rewrite the worst-case VaR as

$$\begin{aligned} C(S) = \mathcal{F}_{\mathrm{W}}(\theta)\text{-VaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}] = \min_{\eta}\ & \mu_S + \sigma_S\eta \\ \text{s.t. }\ & \eta(\Phi(\eta)-(1-\varepsilon)) - \int_{\frac{(\Phi^{-1}(1-\varepsilon))^2}{2}}^{\frac{\eta^2}{2}} kg(z)dz \geq \varepsilon\theta \\ & \eta \geq \Phi^{-1}(1-\varepsilon). \end{aligned} \quad (16)$$

The function

$$V(\eta) \triangleq \eta(\Phi(\eta)-(1-\varepsilon)) - \int_{\frac{(\Phi^{-1}(1-\varepsilon))^2}{2}}^{\frac{\eta^2}{2}} kg(z)dz$$

is monotonically increasing on $\eta \geq \Phi^{-1}(1-\varepsilon)$ since for any $\eta > \Phi^{-1}(1-\varepsilon)$, we have

$$V'(\eta) = \Phi(\eta) - (1-\varepsilon) + \eta\phi(\eta) - \eta kg\left(\frac{\eta^2}{2}\right) = \Phi(\eta) - (1-\varepsilon) > 0.$$

Therefore, we can solve the one-dimensional optimization problem (16) effectively via a bisection search, and we have $C(S) = \mu_S + \eta^\star\sigma_S$ as claimed. Following from the subadditivity of standard deviation, for any positive definite matrix $\boldsymbol{\Sigma}$ and nonnegative $\eta^\star$, $C(S) = \mu_S + \sigma_S\eta^\star$ is subadditive.

Finally, we check that the proposed $\boldsymbol{a}$ lies in the core. On the one hand, it holds that

$$\sum_{i\in N} a_i = \sum_{i\in N}\left(\mu_i + \eta^\star \cdot \frac{\boldsymbol{e}_{\{i\}}^\top\boldsymbol{\Sigma}\boldsymbol{e}_N}{\sqrt{\boldsymbol{e}_N^\top\boldsymbol{\Sigma}\boldsymbol{e}_N}}\right) = \boldsymbol{e}_N^\top\boldsymbol{\mu} + \eta^\star\sqrt{\boldsymbol{e}_N^\top\boldsymbol{\Sigma}\boldsymbol{e}_N} = C(N).$$

On the other hand, for any $S \subseteq N$, we have

$$\sum_{i\in S} a_i = \sum_{i\in S}\left(\mu_i + \eta^\star \cdot \frac{\boldsymbol{e}_{\{i\}}^\top\boldsymbol{\Sigma}\boldsymbol{e}_N}{\sqrt{\boldsymbol{e}_N^\top\boldsymbol{\Sigma}\boldsymbol{e}_N}}\right) = \boldsymbol{e}_S^\top\boldsymbol{\mu} + \eta^\star\frac{\boldsymbol{e}_S^\top\boldsymbol{\Sigma}\boldsymbol{e}_N}{\sqrt{\boldsymbol{e}_N^\top\boldsymbol{\Sigma}\boldsymbol{e}_N}} \leq \boldsymbol{e}_S^\top\boldsymbol{\mu} + \eta^\star\sqrt{\boldsymbol{e}_S^\top\boldsymbol{\Sigma}\boldsymbol{e}_S} = C(S),$$

where the inequality follows from a Cauchy-Schwartz inequality $\boldsymbol{e}_S^\top\boldsymbol{\Sigma}\boldsymbol{e}_N \leq \sqrt{\boldsymbol{e}_S^\top\boldsymbol{\Sigma}\boldsymbol{e}_S}\sqrt{\boldsymbol{e}_N^\top\boldsymbol{\Sigma}\boldsymbol{e}_N}$. □

7.3. Proof for Section 5

Proof of Proposition 2. When $\theta = 0$, the statement trivially holds. Thus, it remains to prove the case when $\theta > 0$.

Firstly, we rewrite the coalitional worst-case CVaR as follows:

$$\mathcal{F}_{\mathrm{W}}(\theta)\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}}] = \inf_{\tau\in\mathbb{R}}\left\{\tau + \frac{1}{\varepsilon}\sup_{\mathbb{P}\in\mathcal{F}_{\mathrm{W}}(\theta)} \mathbb{E}_{\mathbb{P}}[(\boldsymbol{e}_S^\top\tilde{\boldsymbol{x}} - \tau)^+]\right\},$$

where the first equality is due to the fact that the Wasserstein ambiguity set $\mathcal{F}_{\mathrm{W}}(\theta)$ is weakly compact for any $\theta > 0$ (Boissard et al. 2011).

Secondly, We show that for any $\tau \in \mathbb{R}$, the following relation holds:

$$\sup_{\mathbb{P}\in\mathcal{F}_{\mathrm{W}}(\theta)} \mathbb{E}_{\mathbb{P}}[(\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}} - \tau)^+] = \mathbb{E}_{\mathbb{P}_{\mathrm{R}}}[(\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}} - \tau)^+] + \theta \|\boldsymbol{e}_S\|_*. \tag{17}$$

By Theorem 1 in Gao and Kleywegt (2016) or Theorem 1 in Blanchet and Murthy (2019), we have

$$\sup_{\mathbb{P}\in\mathcal{F}_{\mathrm{W}}(\theta)} \mathbb{E}_{\mathbb{P}}[(\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}} - \tau)^+] = \min_{\lambda\geq 0} \left\{ \lambda\theta + \int_{\boldsymbol{x}} \sup_{\boldsymbol{y}} \left\{ (\boldsymbol{e}_S^\top \boldsymbol{y} - \tau)^+ - \lambda\|\boldsymbol{y}-\boldsymbol{x}\| \right\} \mathbb{P}_{\mathrm{R}}(\mathrm{d}\boldsymbol{x}) \right\}. \tag{18}$$

For any fixed $\boldsymbol{x}$, it holds that

$$\begin{aligned} \sup_{\boldsymbol{y}} \left\{ (\boldsymbol{e}_S^\top \boldsymbol{y} - \tau)^+ - \lambda\|\boldsymbol{y}-\boldsymbol{x}\| \right\} &= \sup_{\boldsymbol{y}} \left\{ \max\left\{ \boldsymbol{e}_S^\top \boldsymbol{y} - \tau - \lambda\|\boldsymbol{y}-\boldsymbol{x}\|, -\lambda\|\boldsymbol{y}-\boldsymbol{x}\| \right\} \right\} \\ &= \max\left\{ \sup_{\boldsymbol{y}} \left\{ \boldsymbol{e}_S^\top \boldsymbol{y} - \tau - \lambda\|\boldsymbol{y}-\boldsymbol{x}\| \right\}, \sup_{\boldsymbol{y}} \left\{ -\lambda\|\boldsymbol{y}-\boldsymbol{x}\| \right\} \right\} \\ &= \max\left\{ \sup_{\boldsymbol{y}} \left\{ \boldsymbol{e}_S^\top \boldsymbol{y} - \tau - \lambda\|\boldsymbol{y}-\boldsymbol{x}\| \right\}, 0 \right\}, \end{aligned} \tag{19}$$

where in the last equality, we choose the optimal $\boldsymbol{y}^\star = \boldsymbol{x}$ for the problem $\sup_{\boldsymbol{y}}\{-\lambda\|\boldsymbol{y}-\boldsymbol{x}\|\}$ by noting $\lambda \geq 0$. Using conic duality, we have

$$\sup_{\boldsymbol{y}} \left\{ \boldsymbol{e}_S^\top \boldsymbol{y} - \lambda\|\boldsymbol{y}-\boldsymbol{x}\| \right\} = \begin{cases} \boldsymbol{e}_S^\top \boldsymbol{x} & \text{if } \lambda \geq \|\boldsymbol{e}_S\|_* \\ +\infty & \text{if } \lambda \in [0, \|\boldsymbol{e}_S\|_*). \end{cases} \tag{20}$$

Combining (19) with (20), we obtain

$$\sup_{\boldsymbol{y}} \left\{ (\boldsymbol{e}_S^\top \boldsymbol{y} - \tau)^+ - \lambda\|\boldsymbol{y}-\boldsymbol{x}\| \right\} = \begin{cases} (\boldsymbol{e}^\top \boldsymbol{x} - \tau)^+ & \text{if } \lambda \geq \|\boldsymbol{e}_S\|_* \\ +\infty & \text{if } \lambda \in [0, \|\boldsymbol{e}_S\|_*). \end{cases}$$

We plug this expression into (18). Observe that it is sufficient to focus on $\lambda \geq \|\boldsymbol{e}_S\|_*$, thus we have

$$\sup_{\mathbb{P}\in\mathcal{F}_{\mathrm{W}}(\theta)} \mathbb{E}_{\mathbb{P}}[(\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}} - \tau)^+] = \min_{\lambda\geq\|\boldsymbol{e}_S\|_*} \left\{ \lambda\theta + \int_{\boldsymbol{x}} (\boldsymbol{e}^\top \boldsymbol{x} - \tau)^+ \mathbb{P}_{\mathrm{R}}(\mathrm{d}\boldsymbol{x}) \right\},$$

which by choosing the optimal $\lambda^\star = \|\boldsymbol{e}_S\|_*$, we verify the desired relation (17).

Finally, applying the relation (17), we arrive at

$$\begin{aligned} \mathcal{F}\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] &= \inf_{\tau\in\mathbb{R}} \left\{ \tau + \frac{1}{\varepsilon} \sup_{\mathbb{P}\in\mathcal{F}_{\mathrm{W}}(\theta)} \mathbb{E}_{\mathbb{P}}[(\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}} - \tau)^+] \right\} \\ &= \inf_{\tau\in\mathbb{R}} \left\{ \tau + \frac{1}{\varepsilon} \mathbb{E}_{\mathbb{P}_{\mathrm{R}}}[(\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}} - \tau)^+] + \frac{\theta\|\boldsymbol{e}_S\|_*}{\varepsilon} \right\} \\ &= \inf_{\tau\in\mathbb{R}} \left\{ \tau + \frac{1}{\varepsilon} \mathbb{E}_{\mathbb{P}_{\mathrm{R}}}[(\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}} - \tau)^+] \right\} + \frac{\theta\|\boldsymbol{e}_S\|_*}{\varepsilon} \\ &= \mathbb{P}_{\mathrm{R}}\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] + \frac{\theta\|\boldsymbol{e}_S\|_*}{\varepsilon}. \end{aligned}$$

This concludes the proof. $\square$

Proof of Theorem 8. Using now standard techniques in distributionally robust optimization, the worst-case CVaR, $\mathcal{F}_{\mathrm{W}}(\theta)$-CVaR$_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}]$ can be formulated as the following optimization problem:

$$\begin{array}{lll} \inf\limits_{\boldsymbol{\alpha},\beta,\tau} & \tau + \dfrac{1}{K}\sum\limits_{k\in[K]} \alpha_k + \theta\beta & \\ \text{s.t.} & \alpha_k + \|\boldsymbol{y} - \hat{\boldsymbol{x}}_k\|\beta \geq \dfrac{1}{\varepsilon}(\boldsymbol{e}_S^\top \boldsymbol{y} - \tau) & \forall \boldsymbol{y} \in \mathbb{R}^K,\ k \in [K] \\ & \alpha_k + \|\boldsymbol{z} - \hat{\boldsymbol{x}}_k\|\beta \geq 0 & \forall \boldsymbol{z} \in \mathbb{R}^K,\ k \in [K] \\ & \boldsymbol{\alpha} \in \mathbb{R}^K,\ \beta \geq 0,\ \tau \in \mathbb{R}, & \end{array}$$

see Mohajerin Esfahani and Kuhn (2018, § 5.1 and § 7.1) for a detailed derivation or Proposition 6 in Chen et al. (2018). By the technique "primal worst equals dual best" as explored in the proof of Theorem 2, the optimal value of the above problem coincides with that of problem (7), where the feasible region is independent of the coalition S. Following similar lines as in the proof of Theorem 2, the proposed allocation scheme belongs to the core. □

Proof of Theorem 9. According to Proposition 2, we have

$$\mathcal{F}_{\mathrm{W}}(\theta)\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] = \mathbb{P}_{\mathrm{E}}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)\text{-CVaR}_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}] + \frac{\theta\|\boldsymbol{e}_S\|_*}{\varepsilon}.$$

The derivation of $\mathbb{P}_{\mathrm{E}}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$-CVaR$_{1-\varepsilon}[\boldsymbol{e}_S^\top \tilde{\boldsymbol{x}}]$ is similar to that in Theorem 7 and is thus omitted.

Let $\boldsymbol{x}_N^\star$ be the optimal solution to the problem that determines $\|\boldsymbol{e}_N\|_*$. A hybrid of results in Theorem 5 and Theorem 7 concludes that the proposed allocation scheme belongs to the core. □

Acknowledgments

The research of Zhi Chen is funded by the start-up grant for new faculty (project number: 7200646) from the City University of Hong Kong. The authors are grateful to Viet Anh Nguyen and Soroosh Shafieezadeh-Abadeh for the helpful discussions on distributionally robust optimization with Wasserstein distance. Valuable comments from the editors and two anonymous reviewers are also gratefully acknowledged.

References

Anily, Shoshana, Moshe Haviv. 2014. Subadditive and homogeneous of degree one games are totally balanced. *Operations Research* **62**(4) 788–793.

Bartl, Daniel, Samuel Drapeau, Ludovic Tangpi. 2020. Computational aspects of robust optimized certainty equivalents and option pricing. *Mathematical Finance* **30**(1) 287–309.

Baton, Bernard, Jean Lemaire. 1981. The core of a reinsurance market. *ASTIN Bulletin: The Journal of the IAA* **12**(1) 57–71.

Beck, Amir, Aharon Ben-Tal. 2009. Duality in robust optimization: primal worst equals dual best. *Operations Research Letters* **37**(1) 1–6.

Ben-Tal, Aharon, Arkadi Nemirovski. 2001. *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*. SIAM.

Bertsimas, Dimitris, Melvyn Sim, Meilin Zhang. 2019. Adaptive distributionally robust optimization. *Management Science* **65**(2) 604–618.

Blanchet, Jose, Karthyek Murthy. 2019. Quantifying distributional model risk via optimal transport. *Mathematics of Operations Research* **44**(2) 565–600.

Boissard, Emmanuel, et al. 2011. Simple bounds for the convergence of empirical and occupation measures in 1-wasserstein distance. *Electronic Journal of Probability* **16** 2296–2333.

Borch, Karl. 1960a. Reciprocal reinsurance treaties seen as a two-person co-operative game. *Scandinavian Actuarial Journal* **1960**(1-2) 29–58.

Borch, Karl. 1960b. The safety loading of reinsurance premiums. *Scandinavian Actuarial Journal* **1960**(3-4) 163–184.

Borch, Karl. 1962. Equilibrium in a reinsurance market. *Econometrica* **30**(3) 424–444.

Chen, Xin, Zhenyu Hu, Shuanglong Wang. 2017. Stable risk sharing and its monotonicity. *Available at SSRN*.

Chen, Zhi, Daniel Kuhn, Wolfram Wiesemann. 2018. Data-driven chance constrained programs over Wasserstein balls. *Available at Optimization Online*.

Csóka, Péter, P Jean-Jacques Herings, László Á Kóczy. 2009. Stable allocations of risk. *Games and Economic Behavior* **67**(1) 266–276.

Delage, Erick, Yinyu Ye. 2010. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research* **58**(3) 595–612.

Delbaen, Freddy. 2000. Coherent risk measures Lecture notes, Pisa.

Denault, Michel. 2001. Coherent allocation of risk capital. *Journal of risk* **4** 1–34.

Embrechts, Paul. 2000. Extreme value theory: potential and limitations as an integrated risk management tool. *Derivatives Use, Trading & Regulation* **6**(1) 449–456.

Embrechts, Paul, Alexander McNeil, Daniel Straumann. 2002. Correlation and dependence in risk management: properties and pitfalls. *Risk management: value at risk and beyond* **1** 176–223.

Embrechts, Paul, Giovanni Puccetti. 2010. Risk aggregation. *Copula Theory and Its Applications*. Springer, 111–126.

Gao, Rui, Anton J Kleywegt. 2016. Distributionally robust stochastic optimization with Wasserstein distance. *arXiv preprint arXiv:1604.02199*.

Ghosal, Shubhechyya, Wolfram Wiesemann. 2018. The distributionally robust chance constrained vehicle routing problem *Available on Optimization Online*.

Gupta, Arjun K, Tamas Varga, Taras Bodnar. 2013. *Elliptically contoured models in statistics and portfolio theory*. Springer.

Hanasusanto, Grani A, Vladimir Roitch, Daniel Kuhn, Wolfram Wiesemann. 2015. A distributionally robust perspective on uncertainty quantification and chance constrained programming. *Mathematical Programming* **151**(1) 35–62.

Hanasusanto, Grani A, Vladimir Roitch, Daniel Kuhn, Wolfram Wiesemann. 2017. Ambiguous joint chance constraints under mean and dispersion information. *Operations Research* **65**(3) 751–767.

Kaas, Rob, Roger JA Laeven, Roger B Nelsen. 2009. Worst VaR scenarios with given marginals and measures of association. *Insurance: Mathematics and Economics* **44**(2) 146–158.

Kamdem, Jules Sadefo. 2005. Value-at-risk and expected shortfall for linear portfolios with elliptically distributed risk factors. *International Journal of Theoretical and Applied Finance* **8**(05) 537–551.

Kuhn, Daniel, Peyman Mohajerin Esfahani, Viet Anh Nguyen, Soroosh Shafieezadeh-Abadeh. 2019. Wasserstein distributionally robust optimization: Theory and applications in machine learning. *Forthcoming in INFORMS TutORials in Operations Research.*

Landsman, Zinoviy M, Emiliano A Valdez. 2003. Tail conditional expectations for elliptical distributions. *North American Actuarial Journal* **7**(4) 55–71.

Li, Bowen, Ruiwei Jiang, Johanna L Mathieu. 2019. Ambiguous risk constraints with moment and unimodality information. *Mathematical Programming* **173**(1-2) 151–192.

McNeil, Alexander J, Rüdiger Frey, Paul Embrechts, et al. 2005. *Quantitative risk management: concepts, techniques and tools*, vol. 3. Princeton university press Princeton.

Mohajerin Esfahani, Peyman, Daniel Kuhn. 2018. Data-driven distributionally robust optimization using the Wasserstein metric: performance guarantees and tractable reformulations. *Mathematical Programming* **171**(1-2) 1–52.

Pérignon, Christophe, Daniel R Smith. 2010. Diversification and value-at-risk. *Journal of Banking & Finance* **34**(1) 55–66.

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

Shapiro, Alexander, Darinka Dentcheva, Andrzej Ruszczyński. 2009. *Lectures on stochastic programming: modeling and theory*. SIAM.

Shapiro, Alexander, Anton Kleywegt. 2002. Minimax analysis of stochastic problems. *Optimization Methods and Software* **17**(3) 523–542.

Wang, Ruodu, Liang Peng, Jingping Yang. 2013. Bounds for the sum of dependent risks and worst value-at-risk with monotone marginal densities. *Finance and Stochastics* **17**(2) 395–417.

Wiesemann, Wolfram, Daniel Kuhn, Melvyn Sim. 2014. Distributionally robust convex optimization. *Operations Research* **62**(6) 1358–1376.

Xie, Weijun. 2018. On distributionally robust chance constrained program with Wasserstein distance. *Available at Optimization Online.*

Xie, Weijun, Shabbir Ahmed. 2018. On deterministic reformulations of distributionally robust joint chance constrained optimization problems. *SIAM Journal on Optimization* **28**(2) 1151–1182.