Math. Program., Ser. A (2019) 173:393–430
https://doi.org/10.1007/s10107-017-1224-6

FULL LENGTH PAPER

Identifying effective scenarios in distributionally robust stochastic programs with total variation distance

Hamed Rahimian[1] · **Güzin Bayraksan**[1] · **Tito Homem-de-Mello**[2]

Received: 2 March 2016 / Accepted: 23 December 2017 / Published online: 22 January 2018

Abstract Traditional stochastic programs assume that the probability distribution of uncertainty is known. However, in practice, the probability distribution oftentimes is not known or cannot be accurately approximated. One way to address such distributional ambiguity is to work with distributionally robust convex stochastic programs (DRSPs), which minimize the worst-case expected cost with respect to a set of probability distributions. In this paper we analyze the case where there is a finite number of possible scenarios and study the question of how to identify the critical scenarios resulting from solving a DRSP. We illustrate that not all, but only some scenarios might have "effect" on the optimal value, and we formally define this notion for our general class of problems. In particular, we examine problems where the distributional ambiguity is modeled by the so-called total variation distance. We propose easy-to-check conditions to identify *effective* and *ineffective* scenarios for that class of problems. Computational results show that identifying effective scenarios provides useful insight on the underlying uncertainties of the problem.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10107-017-1224-6) contains supplementary material, which is available to authorized users.

✉ Güzin Bayraksan
bayraksan.1@osu.edu

Hamed Rahimian
rahimian.1@osu.edu

Tito Homem-de-Mello
tito.hmello@uai.cl

[1] Department of Integrated Systems Engineering, The Ohio State University, Columbus, OH 43210, USA

[2] School of Business, Universidad Adolfo Ibañez, Santiago, Chile

Keywords Stochastic programming · Distributionally robust optimization · Risk measures · Scenario analysis

Mathematics Subject Classification 90C15 · 90C47

1 Introduction

A crucial task when building stochastic optimization models is finding an underlying probability distribution to represent the uncertainty. Most often, partial information about the uncertainty is available through a series of historical data. In such circumstances, traditional stochastic programs rely on approximating the underlying probability distribution. However, in many real-world applications, the underlying probability distribution cannot be accurately determined, even when historical data are available. For example, even though we may have historical demand data, demands might shift as a result of economic development, technological innovations, and so forth. This distributional "ambiguity" might lead to highly suboptimal decisions. An alternative modeling approach to handle such an issue is to use "distributionally robust stochastic programs" or "ambiguous stochastic programs," which assume the underlying probability distribution lies in an ambiguous set of distributions and hedge against the worst-case probability distribution in the ambiguity set.

To formalize these notions, consider a convex stochastic optimization problem of the form

$$\min_{x \in \mathbb{X}} \sum_{\omega \in \Omega} q_\omega h_\omega(x), \tag{1}$$

where $\mathbb{X} \subseteq \mathbb{R}^m$ represents the deterministic feasibility set for the decision vector x, assumed to be a non-empty convex compact set. We assume that there exists a finite set of possible scenarios that can occur and denote that set by $\Omega = \{\omega_1, \dots, \omega_n\}$. A generic scenario in that set is denoted by ω. Throughout the paper we shall use bold notation to denote scenario-dependent vectors; so, in (1), $\mathbf{q}$ denotes a known probability vector to represent the uncertainty. We assume that the cost function $h_\omega(\cdot) : \mathbb{X} \mapsto \mathbb{R}$ is convex and real-valued in an open set containing $\mathbb{X}$, for all $\omega \in \Omega$. For $x \in \mathbb{X}$, we denote by $\mathbf{h}(x)$ the random variable that takes values $h_\omega(x)$, $\omega \in \Omega$. Problem (1) contains a wide range of problems, including two-stage stochastic linear programs with recourse (SLP-2 for short), provided it satisfies the aforementioned assumptions. For instance, in SLP-2 we have $\mathbb{X} := \{x \in \mathbb{R}^m : Ax = b,\ x \geq 0\}$ (assumed to be non-empty and compact) and $h_\omega(x) = cx + \mathcal{H}_\omega(x)$, where $\mathcal{H}_\omega(x)$ represents the second-stage cost function given decision x and scenario ω: $\mathcal{H}_\omega(x) = \min_{y_\omega}\{k_\omega y_\omega : D_\omega y_\omega = B_\omega x + d_\omega,\ y_\omega \geq 0\}$ (assumed to be real-valued in an open set containing $\mathbb{X}$ for all $\omega \in \Omega$).

The distributionally robust version of stochastic program (1), called DRSP for short, can be formulated as

$$\min_{x \in \mathbb{X}} \left\{ f(x) := \max_{\mathbf{p} \in \mathcal{P}} \sum_{\omega \in \Omega} p_\omega h_\omega(x) \right\}, \tag{DRSP}$$

where $\mathcal{P}$ is the *ambiguity set of distributions* and is a subset of all probability distributions on Ω. The inner maximization problem in (DRSP) hedges against the worst-case probability distribution in the ambiguity set $\mathcal{P}$. For a given $x \in \mathbb{X}$, we refer to this inner problem as the *worst-case expected problem* at x.

Several approaches to model the distributional ambiguity have been studied in the literature. One common approach is moment based, in which the ambiguity set contains all probability distributions whose moments (typically the first two moments) satisfy certain properties; see, e.g., [12,16,38,44]. In another widely studied approach, the ambiguity set is constructed by considering all probability distributions whose distances to a *nominal* probability distribution are sufficiently small. Many distances have been used in the literature to construct the ambiguity set; e.g., Prohorov metric [13], Kantorovich metric [30,32,45], Wasserstein metric [14,15,29], and ζ-structure probability metric [48]. Recently, a famous class of distances in information theory—ϕ-divergences—has attracted increasing attention. The authors in the pioneering work [5], as well as [4,24], study general ϕ-divergences. Some papers have studied a particular ϕ-divergence including Kullback-Leibler divergence [8,20,21,43], χ^2 distance [26], and variation distance [25]. More recently, DRSPs in a multistage setting has been studied; see, e.g., [2,31,39,41,46]. For a detailed review of different approaches to model the distributional ambiguity, we refer to [17] and references therein.

The solution of (DRSP) problem yields an optimal decision vector x^* as well as an optimal probability distribution $\mathbf{p}^* := \mathbf{p}^*(x^*)$. A natural question that arises is: what does that distribution look like? By construction, $\mathbf{p}^*$ is the worst-case distribution corresponding to an optimal decision x^*. So, presumably, the support of $\mathbf{p}^*$ (i.e., the subset of scenarios for which $p^*_\omega > 0$) would indicate the *critical scenarios* we are concerned about. Similarly, the scenarios in the complement of that set appear to have no importance for our optimal decision. Such study of scenarios may help the decision maker understand their problem better and encourage them to collect more accurate information surrounding certain scenarios. Consider, for example, a traditional newsvendor model. In this case, it is intuitive that if we are very conservative—i.e., if we take the ambiguity set to be very large so we protect against the worst possible distribution—the critical scenario is the one in which demand is the lowest. But suppose we are not very conservative. Then, what is the region of demand values that are critical? How far in the tail of the demand distribution are these scenarios? The answer to these questions may lead the decision maker to do a more careful study to better estimate that portion of the tail of the distribution, which often gets neglected by standard goodness-of-fit procedures. Further examples of practical situations are discussed in Sect. 5.

In this paper we study the issue of critical scenarios in the context of DRSP. One major contribution of our work is a precise definition of what is meant by a critical scenario. Roughly speaking, critical scenarios according to our definition are those scenarios whose removal from the problem (defined in a precise way in Sect. 2) leads to changes in the optimal value; we call those *effective scenarios*. Likewise, *ineffective scenarios* are those that can be removed from the problem without causing any changes. As it turns out—and perhaps somewhat surprisingly—we cannot determine whether a scenario is effective simply by looking at the value of p^*_ω for that scenario. Indeed, it is possible to have effective scenarios for which $p^*_\omega = 0$, and ineffective scenarios for which $p^*_\omega > 0$.

How to determine then the set of effective/ineffective scenarios? One way, of course, would be to re-solve the problem multiple times, i.e., without each scenario, and observe the cases in which there were any changes. Clearly, this is not practical, especially for large problems. Our goal is then to develop criteria for effectiveness/ineffectiveness, but in order to do so, it is necessary to exploit specific structures of the ambiguity set $\mathcal{P}$. Following the literature on DRSPs with ϕ-divergences, we define $\mathcal{P}$ as the set of distributions $\mathbf{p}$ that are not "too far" from a "nominal" distribution $\mathbf{q}$ that is pre-established by the user or inferred from existing data. Here, "not too far" refers to the value of the *total variation distance* $V(\mathbf{p}, \mathbf{q}) := \frac{1}{2}\sum_{\omega\in\Omega} |p_\omega - q_\omega|$, which is a particular case of ϕ-divergence. By exploiting the linear structure of this function, we are able to obtain *easy-to-check conditions* based on the relationship between $\mathbf{p}^*$ and $\mathbf{q}$ that allow us to classify many of the scenarios. While some scenarios may remain unclassified, in our numerical experiments such unclassified scenarios represented a small fraction of the total number of scenarios.

Our choice to work with the total variation distance has other benefits. The first one is conceptual: as shown in [4], under the total variation distance it is possible to have $p^*_\omega = 0$ for a scenario for which $q_\omega > 0$, and we may also have $p^*_\omega > 0$ for a scenario for which $q_\omega = 0$ (these situations are called respectively "suppressed" and "popped" scenarios in that paper). These possibilities broaden the set of scenarios for which our easy-to-check conditions apply. The second benefit is computational: when the total variation distance is used to define $\mathcal{P}$ in (DRSP), the result is a computationally tractable optimization model which can be solved by a decomposition algorithm. Finally, the third benefit is interpretative: when using the total variation distance, the resulting DRSP has an attractive interpretation in terms of risk, as it can be expressed as a linear combination of the worst-case and the Conditional Value-at-Risk (CVaR) of the cost function.

To the best of our knowledge, this is the first work that defines the notion of effective scenarios for DRSP. There are papers that introduce "support constraints"—which is somewhat similar to our notion of effective scenarios—in the context of optimization problems subject to convex constraints parameterized by an uncertainty parameter [9–11]. They define a support constraint as any constraint whose removal changes the optimal value of a constraint-sampled problem. By determining the maximum number of support constraints, the authors in [9,10] probabilistically describe the feasibility of a randomized sample-based solution if a new realization of uncertainty is observed. Similarly, the authors in [11] use this notion to study the optimality of a randomized solution to a minimization of the worst-case cost problem.

Our work is also related to the work in [4]. As mentioned earlier, the authors of that paper define the notions of "suppressing" and "popping" a scenario and provide a categorization of ϕ-divergence functions based on these two notions. As it turns out, not all ϕ-divergence functions allow for both suppressing and popping of scenarios—but the total variation distance does. There is, however, no classification of scenarios according to their effect on the problem.

It is important to observe that our study should not be interpreted as either typical stability analysis of stochastic programs [35] or typical perturbation analysis of optimization problems [6]. The former studies the change in the optimal value of a stochastic program with respect to changes in the nominal distribution $\mathbf{q}$, and the latter studies the change in the optimal value of an optimization problem with respect to changes in the input parameters. In contrast, our work assesses whether the optimal value of (DRSP) changes when a collection of scenarios is excluded. This requires examining a problem with additional constraints (see Sect. 2 for details). Recognizing these differences, our study requires new definitions and analyses.

To summarize, the contributions of this article are as follows:

(i) We introduce the notions of effective and ineffective scenarios for a general class of distributionally robust convex stochastic optimization problems with finite support. We give alternative definitions for these notions and present some general properties of effective/ineffective scenarios.
(ii) We provide easy-to-check sufficient conditions to identify the effectiveness of scenarios for a specific class of (DRSP), namely (DRSP) formed using the total variation distance (DRSP-V).
(iii) As the ambiguity set enlarges, it is perhaps natural to conjecture that the sets of effective/ineffective scenarios are nested and the number of effective/ineffective scenarios are monotone. We examine monotonicity in the number of effective/ineffective scenarios and nestedness of the corresponding sets in more detail. We illustrate through counterexamples that these properties do not hold in general.
(iv) We numerically illustrate the concepts introduced in this paper. Computational results indicate that easy-to-check conditions work well by identifying a relatively large set of scenarios immediately after solution of the problem.
(v) Finally, we investigate how effective and ineffective scenarios can be used to gain insight into the underlying uncertainties of a problem.

The rest of this paper is outlined as follows. In Sect. 2, we formally introduce the notions of effective and ineffective scenarios for a generic DRSP and present some properties of such scenarios. In Sect. 3, we review the total variation distance and its risk interpretation for DRSP. In Sect. 4, we focus on the DRSP-V and provide easy-to-check conditions to identify effective/ineffective scenarios in that setting. We then present numerical experiments to test the efficacy of easy-to-check conditions and analyze the insights obtained from the study of effectiveness of scenarios in Sect. 5. In that section, we also discuss the monotonicity properties of the sets of effective/ineffective scenarios. We end with conclusions and discussion of future work in Sect. 6. Some results and proofs are presented in the Appendix; references to these results are preceded by "A-" in the text.

2 Identifying effectiveness of scenarios for DRSP

As mentioned earlier, the idea behind identifying effective scenarios is to verify whether the optimal value of (DRSP) changes when a scenario (or, more generally, a set of scenarios) is removed from the problem. The first task is then is to define what is meant by "removing" a set of scenarios $\mathcal{F} \subset \Omega$ from the problem. We do that by restricting the ambiguity set $\mathcal{P}$ to those probabilities $\mathbf{p}$ for which $p_\omega = 0,\ \omega \in \mathcal{F}$. This ensures that the scenarios in $\mathcal{F}$ are not in the support of any optimal worst-case probability $\mathbf{p}^*$. We shall call the resulting problem the *assessment problem of scenarios in* $\mathcal{F}$. More formally, this problem can be formulated as

$$\min_{x\in\mathbb{X}} \left\{ f^{\mathrm{A}}(x;\mathcal{F}) := \max_{\mathbf{p}\in\mathcal{P}^{\mathrm{A}}(\mathcal{F})} \sum_{\omega\in\mathcal{F}^{\mathsf{c}}} p_\omega h_\omega(x) \right\}, \tag{2}$$

where $\mathcal{P}^{\mathrm{A}}(\mathcal{F}) := \mathcal{P} \cap \{p_\omega = 0,\ \omega \in \mathcal{F}\}$ is the ambiguity set of distributions for the assessment problem of scenarios in $\mathcal{F}$. In (2), $\mathcal{F}^{\mathsf{c}}$ denotes the complement of set $\mathcal{F}$, i.e., $\Omega \setminus \mathcal{F}$. Notice that the assessment problem does not eliminate scenarios from the sample space; it only enforces $p_\omega = 0$ for all $\omega \in \mathcal{F}$. In Sect. 4.2, we further investigate the assessment problem when the total variation distance is used to model the distributional ambiguity; see (9)–(10)

Note that the assessment problem of scenarios in $\mathcal{F}$ might not be well defined. For instance, if too many scenarios are restricted to have a zero worst-case probability, the inner maximization problem in (2) might become infeasible, in which case (2) is not well defined. By convention, we set $f^{\mathrm{A}}(x;\mathcal{F}) = -\infty$ in such cases.

Based on the definition of the assessment problem in (2), one might think that it suffices to look at the value of p^*_ω to decide about the effectiveness of a scenario ω. After all, if p^*_ω is already zero then it seems that nothing will change if we re-solve the problem with the extra constraint $p_\omega = 0$. As it turns out, however, this is not necessarily true, as demonstrated by a simple example in Sect. 2.1. We will then present the formal definition and properties in Sect. 2.2.

2.1 Motivating example

Example 1 Let us consider a newsvendor problem to decide on how much inventory, x, should be ordered to minimize cost, z. Suppose that the unit cost is 2 and unit revenue is 3. The uncertain demand takes values $\mathbf{d} = (2, 5, 1)$, with nominal probabilities $\mathbf{q} = (0.30, 0.70, 0)$. Then, we can write $h_\omega(x) = 2x - 3\min\{x, d_\omega\}$. Suppose that the ambiguity set contains all possible probability distributions on the support of $\mathbf{d}$. Then, the worst-case expected problem reduces to the largest cost (i.e., $2x - 3\min\{x, 1\}$). The optimal x^* and z^* are 1 and -1, respectively. Also, $\mathbf{p}^* = (0.50, 0.50, 0)$ is an optimal solution for the worst-case probabilities corresponding to x^*.

Now, let us look at some restricted versions of the ambiguity set, by removing one scenario at a time. It can be seen that if either the first or second scenario (a scenario with a positive worst-case probability, $p^*_1 = p^*_2 = 0.5$) is removed, the largest cost would still be $2x - 3\min\{x, 1\}$, and hence the optimal solution and optimal value do not change. On the other hand, if the third scenario (the scenario with a zero worst-case probability, $p^*_3 = 0$) is removed, the optimal value changes. Since the worst-case probability of the third scenario is zero, by looking at the objective function of the restricted primal problem (i.e., $\max_{\mathbf{p}\in\mathcal{P}\cap\{p_3=0\}} \sum_{\omega\in\Omega} p_\omega h_\omega(x)$), one might expect that the optimal solution and optimal value would still be 1 and -1, respectively. However, the objective function of the restricted problem (i.e., $2x - 3\min\{x, 2\}$) reveals that the optimal solution and optimal value would be 2 and -2, respectively. □

The points learned from the above example are as follows:

- The worst-case probability of a scenario—possibly observed after a numerical solution procedure—is not necessarily an indication of its effect on the optimal value. Hence, a decision maker interested in identifying the effectiveness of scenarios cannot rely only on the values of the worst-case probabilities.
- For the sake of illustration, we assumed $\mathcal{P}$ is the set of all probability distributions on Ω. Nevertheless, the ambiguity set $\mathcal{P}$ can be any subset of probability distributions on Ω, and a similar behavior might happen.
- Generally speaking, x^* is not necessarily unique. On the other hand, given x^*, the worst-case probability is not necessarily unique as well. In Example 1, worst-case probabilities $\mathbf{p}^* = (0, 0, 1)$ are also optimal at optimal solution $x^* = 1$. This implies there might be a connection between these observations and the existence of multiple optimal probabilities. For DRSP-V, we prove that the existence of multiple worst-case probabilities is a necessary condition to observe this behavior. In such cases, there exists an alternative worst-case probability distribution for which the worst-case probabilities of scenarios are more indicative of their effects on the optimal value (see Sect. 4.3 for more details).
- In line with the above comment about existence of multiple solutions, one might expect that the use of a modified model that ensures a unique optimal solution may resolve this issue. For instance, one can augment the objective function of the worst-case expected problem with a quadratic function of $\mathbf{p}$. While this "trick" ensures uniqueness of solutions, as a by-product; such a modified model may cause scenarios that do not have effect on the optimal value to have a positive worst-case probability.

As illustrated above, we must examine worst-case probabilities in more detail. Below we will provide a precise definition for effectiveness of scenarios. Later on, based on this definition and subsequent properties, we will present easy-to-check conditions to identify the effectiveness of scenarios for DRSP-V. These easy-to-check conditions require low computational cost and are used in a post-optimality fashion.

2.2 Definition and properties

Consider a subset $\mathcal{F} \subset \Omega$. Let $S := \arg\min_{x \in \mathbb{X}} f(x)$ and $S^{\mathrm{A}}(\mathcal{F}) := \arg\min_{x \in \mathbb{X}} f^{\mathrm{A}}(x; \mathcal{F})$ denote the set of optimal solutions to (DRSP) and the assessment problem of scenarios in $\mathcal{F}$ given in (2), respectively. Suppose that $x^* \in S$ and $\bar{x} \in S^{\mathrm{A}}(\mathcal{F})$. For any $x \in \mathbb{X}$, the corresponding inner maximization problem in (2) is more restrictive than the worst-case expected problem in (DRSP). Therefore, $f^{\mathrm{A}}(x; \mathcal{F}) \leq f(x)$ for all $x \in \mathbb{X}$. Combining this argument with suboptimality conditions, we obtain

$$f^{\mathrm{A}}(\bar{x}; \mathcal{F}) \leq f^{\mathrm{A}}(x^*; \mathcal{F}) \leq f(x^*) \leq f(\bar{x}). \tag{3}$$

The above inequalities lie at the heart of the definition we present below. Essentially, Definition 1 says a scenario is effective if its removal causes the optimal value of the problem to change.

Definition 1 A subset $\mathcal{F} \subset \Omega$ is called *effective* if $\min_{x\in\mathbb{X}} f^{\mathrm{A}}(x;\mathcal{F}) < \min_{x\in\mathbb{X}} f(x)$. A subset $\mathcal{F} \subset \Omega$ is called *ineffective* if it is not effective.

Remark 1 If the assessment problem is not well defined, the corresponding subset of scenarios is effective by definition.

Proposition 1 below relates effectiveness to the sets of optimal solutions. The first part of the proposition shows a sufficient condition for effectiveness in terms of the corresponding sets of optimal solutions. The second part provides another characterization of effectiveness by focusing on all optimal solutions to the assessment problem.

Proposition 1 *A subset $\mathcal{F} \subset \Omega$ is effective if $S \not\subset S^{A}(\mathcal{F})$. Moreover, a subset $\mathcal{F} \subset \Omega$ is effective if and only if $f^{\mathrm{A}}(\bar{x};\mathcal{F}) < f(\bar{x})$ for all $\bar{x} \in S^{\mathrm{A}}(\mathcal{F})$.*

Proof We first prove the first part. Suppose $S \not\subset S^{\mathrm{A}}(\mathcal{F})$, i.e., there exists some $x^* \in S$ such that $x^* \notin S^{\mathrm{A}}(\mathcal{F})$. By (3), we have that $f(x^*) \geq f^{\mathrm{A}}(x^*;\mathcal{F})$, and since $x^* \notin S^{\mathrm{A}}(\mathcal{F})$ it follows that $f^{\mathrm{A}}(x^*;\mathcal{F}) > f^{\mathrm{A}}(\bar{x};\mathcal{F})$. Thus, $\mathcal{F}$ is effective.

Now, we prove the second part of the proposition.

"$\Longrightarrow$": By Definition 1, we have $f^{\mathrm{A}}(\bar{x};\mathcal{F}) < f(x^*)$, for all $x^* \in S$ and all $\bar{x} \in S^{\mathrm{A}}(\mathcal{F})$. By suboptimality of $\bar{x}$ to (DRSP), we have $f(x^*) \leq f(\bar{x}),\ \forall \bar{x} \in S^{\mathrm{A}}(\mathcal{F})$. As a result, $f^{\mathrm{A}}(\bar{x};\mathcal{F}) < f(\bar{x}),\ \forall \bar{x} \in S^{\mathrm{A}}(\mathcal{F})$.

"$\Longleftarrow$": Suppose $f^{\mathrm{A}}(\bar{x};\mathcal{F}) < f(\bar{x}),\ \forall \bar{x} \in S^{\mathrm{A}}(\mathcal{F})$. Consider $x^* \in S$, and an arbitrary $\bar{x} \in S^{\mathrm{A}}(\mathcal{F})$. If $\bar{x} \in S \cap S^{\mathrm{A}}(\mathcal{F})$, we have $f(x^*) = f(\bar{x})$. So, $f^{\mathrm{A}}(\bar{x};\mathcal{F}) < f(x^*)$ for $x^* \in S$ and such $\bar{x}$. Otherwise (i.e., if $\bar{x} \in S^{\mathrm{A}}(\mathcal{F})$ but $\bar{x} \notin S$), suppose by contradiction $f^{\mathrm{A}}(\bar{x};\mathcal{F}) = f(x^*)$. So, $f^{\mathrm{A}}(\bar{x};\mathcal{F}) = f^{\mathrm{A}}(x^*;\mathcal{F})$ by (3) and therefore $x^* \in S^{\mathrm{A}}(\mathcal{F})$. As a result, we have found an element x^* of $S^{\mathrm{A}}(\mathcal{F})$ such that $f^{\mathrm{A}}(x^*;\mathcal{F}) = f(x^*)$, which is a contradiction. □

Remark 2 By the second part of Proposition 1, for a subset $\mathcal{F}$ to be ineffective, there must exist at least one solution $\bar{x} \in S^{\mathrm{A}}(\mathcal{F})$, for which $f^{\mathrm{A}}(\bar{x};\mathcal{F}) = f(\bar{x})$. We can further classify an ineffective subset of scenarios according to whether $f^{\mathrm{A}}(\bar{x};\mathcal{F}) = f(\bar{x})$ for all $\bar{x} \in S^{\mathrm{A}}(\mathcal{F})$ or not. If this equality holds for all $\bar{x} \in S^{\mathrm{A}}(\mathcal{F})$, we call subset $\mathcal{F}$ *fully* ineffective. Otherwise, we call it *partially* ineffective. This subclassification may be suitable for a decision maker more interested in optimal solutions than the optimal value. However, finding *all* optimal solutions $\bar{x} \in S^{\mathrm{A}}(\mathcal{F})$ may require a large computational cost. On the other hand, we are interested in easy-to-check conditions that identify the effectiveness of scenarios without solving the assessment problem. There is clearly a trade-off between obtaining easy-to-check conditions and having a more refined classification. We opt for deriving the simple conditions, as these can already yield useful information, without having to solve the assessment problem.

One can conjecture that the effectiveness of a subset of scenarios might be affected in interaction with other subsets of scenarios. The following proposition addresses the effectiveness of union of an effective subset and intersection of an ineffective subset with any other subset of Ω.

Proposition 2 *(i) The union of an effective subset with any other subset of Ω is effective. (ii) The intersection of an ineffective subset with any other subset of Ω is ineffective.*

Proof (i) Suppose $\mathcal{F}_1$ is an effective subset and $\mathcal{F}_2$ is an arbitrary subset of Ω. First, because $\mathcal{F}_1 \cup \mathcal{F}_2 \supseteq \mathcal{F}_1$, we have $\min_{x\in\mathbb{X}} f^{\mathrm{A}}(x; \mathcal{F}_1 \cup \mathcal{F}_2) \le \min_{x\in\mathbb{X}} f^{\mathrm{A}}(x; \mathcal{F}_1)$ by a similar argument as (3). On the other hand, $\min_{x\in\mathbb{X}} f^{\mathrm{A}}(x; \mathcal{F}_1) < \min_{x\in\mathbb{X}} f(x)$ because $\mathcal{F}_1$ is effective. These imply $\min_{x\in\mathbb{X}} f^{\mathrm{A}}(x; \mathcal{F}_1 \cup \mathcal{F}_2) < \min_{x\in\mathbb{X}} f(x)$, and hence $\mathcal{F}_1 \cup \mathcal{F}_2$ is effective by Definition 1.

(ii) Suppose $\mathcal{F}_1$ is an ineffective subset and $\mathcal{F}_2$ is an arbitrary subset of Ω. First, because $\mathcal{F}_1 \supseteq \mathcal{F}_1 \cap \mathcal{F}_2 \supseteq \emptyset$, we have $\min_{x\in\mathbb{X}} f^{\mathrm{A}}(x; \mathcal{F}_1) \le \min_{x\in\mathbb{X}} f^{\mathrm{A}}(x; \mathcal{F}_1 \cap \mathcal{F}_2) \le \min_{x\in\mathbb{X}} f(x)$ by a similar argument as (3). On the other hand, $\min_{x\in\mathbb{X}} f^{\mathrm{A}}(x; \mathcal{F}_1) = \min_{x\in\mathbb{X}} f(x)$ because $\mathcal{F}_1$ is ineffective. These imply $\min_{x\in\mathbb{X}} f^{\mathrm{A}}(x; \mathcal{F}_1 \cap \mathcal{F}_2) = \min_{x\in\mathbb{X}} f(x)$, and hence $\mathcal{F}_1 \cap \mathcal{F}_2$ is ineffective by Definition 1. □

Corollary 1 *A subset of an ineffective subset is ineffective.*

Proof The proof is immediate from Proposition 2(ii). □

One can also ask whether the union of two ineffective subsets is ineffective or not. This question is more delicate because two individual ineffective subsets may cause a change in the objective function when removed together. We will give such an example in Sect. 4.3 (Example 2). Furthermore, we will provide conditions under which union of specific ineffective subsets remains ineffective for DRSP-V (Theorem 4).

3 DRSP with variation-type distances

We now narrow our focus to DRSP with total variation distance, considering also two variants that use one-sided deviations (Sect. 3.1). Then, we present notation that will be used in the rest of the paper to identify the effectiveness of scenarios for this class of DRSP (Sect. 3.2).

DRSP-V—(DRSP) formed via the total variation distance—can be formulated as

$$\min_{x\in\mathbb{X}} \left\{ f_\gamma(x) = \max_{\mathbf{p}\in\mathcal{P}_\gamma} \sum_{\omega\in\Omega} p_\omega h_\omega(x) \right\}, \tag{DRSP-V}$$

where

$$\mathcal{P}_\gamma := \left\{ \mathbf{p} \,:\, \mathrm{V}(\mathbf{p}, \mathbf{q}) \le \gamma, \quad \sum_{\omega\in\Omega} p_\omega = 1,\ \mathbf{p} \ge \mathbf{0} \right\} \tag{4}$$

and $\mathrm{V}(\mathbf{p}, \mathbf{q}) := \frac{1}{2}\sum_{\omega\in\Omega} |p_\omega - q_\omega|$ denotes the total variation distance between $\mathbf{p}$ and $\mathbf{q}$. The ambiguity set $\mathcal{P}_\gamma$ contains all probability vectors $\mathbf{p}$ whose total variation distance to the nominal probability vector $\mathbf{q}$ is limited above by *the level of robustness* γ. For notational simplicity, we drop the dependence of the objective function $f_\gamma(x)$ and the ambiguity set $\mathcal{P}_\gamma$ on the total variation distance.

In total variation distance, both upper and lower deviations from the nominal probability distribution contribute to the distance. However, one might be interested in two other variation-type distances, in which only a single-sided deviation is considered. That is, one might consider $\mathrm{V}^{\mathrm{R}}(\mathbf{p}, \mathbf{q}) := \frac{1}{2}\sum_{\omega\in\Omega} (p_\omega - q_\omega)_+$ and

$\mathrm{V^L}(\mathbf{p}, \mathbf{q}) := \frac{1}{2}\sum_{\omega\in\Omega}(q_\omega - p_\omega)_+$, which we call *right-sided* and *left-sided* variation distance, respectively. In our notation, $(\cdot)_+ = \max\{0, \cdot\}$. Note that while $\mathrm{V}(\mathbf{p}, \mathbf{q})$ is bounded above by 1, $\mathrm{V^R}(\mathbf{p}, \mathbf{q})$ and $\mathrm{V^L}(\mathbf{p}, \mathbf{q})$ are bounded above by $\frac{1}{2}$. Moreover, $\mathrm{V}(\mathbf{p}, \mathbf{q})$ is a metric on $\mathbb{R}^n$, while the other two are not metrics.

3.1 Risk-averse interpretation

Because the ambiguity sets $\mathrm{V^R}(\mathbf{p}, \mathbf{q})$, $\mathrm{V^L}(\mathbf{p}, \mathbf{q})$, and $\mathrm{V}(\mathbf{p}, \mathbf{q})$ are convex compact, the corresponding worst-case expected problems represent the dual formulation of a real-valued coherent risk measure [3,37]. As a result, their corresponding DRSPs are equivalent to a risk-averse stochastic program involving a coherent risk measure. Proposition 3 presents equivalent risk-averse formulations when variation-type distances are used to model the distributional ambiguity. Below, CVaR is taken with respect to the nominal distribution $\mathbf{q}$.

Proposition 3 *Consider* (DRSP-V) *with* $f_\gamma(\cdot)$. *Let* $f_\gamma^{V^R}(\cdot)$ *and* $f_\gamma^{V^L}(\cdot)$ *denote the corresponding worst-case expected values using the right- and left-sided variation distances, respectively. For a fixed* $x \in \mathbb{X}$ *and* $0 \le \gamma \le 1$, *we have*

$$f_{\frac{\gamma}{2}}^{V^R}(x) = f_{\frac{\gamma}{2}}^{V^L}(x) = f_\gamma(x) = \gamma \sup_{\omega\in\Omega} h_\omega(x) + (1-\gamma)\,\mathrm{CVaR}_\gamma\,[\mathbf{h}(x)]\,.$$

Per usual convention, we set $\mathrm{CVaR}_0\,[\mathbf{h}(x)] = \sum_{\omega\in\Omega} q_\omega h_\omega(x)$ and $\mathrm{CVaR}_1\,[\mathbf{h}(x)] = \sup_{\xi\in\Omega} h_\omega(x)$. So, when $\gamma = 0$, (DRSP-V) reduces to the risk-neutral model (1). As γ increases, more weight is put on the worst-case cost, and (DRSP-V) becomes more conservative. At the highest level of robustness, i.e., $\gamma = 1$, (DRSP-V) reduces to minimizing the worst-case cost.

The right end of Proposition 3 is derived in [25, Theorem 1] , where an earlier version appeared in dissertation [23]. Recent work in [40] shows the above risk-averse interpretation for the right-sided variation distance using a similar method as in [25]. For completeness, we prove this proposition in the Online Supplement with a different approach, following the proof of Proposition 4. The equality between $f_{\frac{\gamma}{2}}^{\mathrm{V^R}}(x)$ and $f_{\frac{\gamma}{2}}^{\mathrm{V^L}}(x)$ is implied by the equivalence between their corresponding ambiguity sets (see Proposition A-1 in the Appendix). Proposition 3 also shows that using the right- or left-sided variation distance results in the same risk-averse model as the total variation distance, but with an adjustment to the level of robustness. Thus, for the rest of the paper we focus on the total variation distance, although all results are applicable to right- and left-sided variation distances with the adjusted level of robustness.

3.2 Primal categories

Before we proceed, let us define some notation for a fixed $x \in \mathbb{X}$. For any $B \subset \mathbb{R}$, we use $[\mathbf{h}(x) \in B]$ as shorthand notation for the set $\{\omega \in \Omega : h_\omega(x) \in B\}$. Also, for a fixed $\eta \in \mathbb{R}$, we use $\Psi(x, \eta)$ to denote $\sum_{\omega\in[\mathbf{h}(x)\le\eta]} q_\omega$.

Given $\beta \in [0, 1]$, let $\mathrm{VaR}_\beta\,[\mathbf{h}(x)]$ be the left-side β-quantile of distribution of $\mathbf{h}(x)$ or equivalently, Value-at-Risk (VaR) of $\mathbf{h}(x)$ at level β: $\mathrm{VaR}_\beta\,[\mathbf{h}(x)] := \inf\{\eta : \Psi(x, \eta) \geq \beta\}$ [34]. By our convention, $\mathrm{VaR}_\beta\,[\mathbf{h}(x)] = -\infty$ for $\beta = 0$ and $\mathrm{VaR}_\beta\,[\mathbf{h}(x)] = \sup_{\omega\in\Omega} h_\omega(x)$ for $\beta = 1$. For the rest of the paper, we assume $\gamma > 0$, unless stated explicitly otherwise. We define the following sets that partition the scenario set Ω:

$\Omega_1(x) :=$ $[\mathbf{h}(x) < \mathrm{VaR}_\gamma\,[\mathbf{h}(x)]]$, i.e., the set of scenarios strictly below $\mathrm{VaR}_\gamma\,[\mathbf{h}(x)]$,
$\Omega_2(x) :=$ $[\mathbf{h}(x) = \mathrm{VaR}_\gamma\,[\mathbf{h}(x)]]$, i.e., the set of scenarios at $\mathrm{VaR}_\gamma\,[\mathbf{h}(x)]$,
$\Omega_3(x) :=$ $[\mathrm{VaR}_\gamma\,[\mathbf{h}(x)] < \mathbf{h}(x) < \sup_{\omega\in\Omega} h_\omega(x)]$, i.e., the set of scenarios strictly between $\mathrm{VaR}_\gamma\,[\mathbf{h}(x)]$ and the worst-case cost at x, and
$\Omega_4(x) :=$ $[\mathbf{h}(x) = \sup_{\omega\in\Omega} h_\omega(x)]$, i.e., the set of scenarios at the worst-case cost at x.

Throughout the rest of the paper, we will make references to the dual variables of the worst-case expected problem. We introduce them next. Let $\lambda(x)$ and $\mu(x)$ denote the optimal dual variables for the first and second constraints in (4), respectively, for a given $x \in \mathbb{X}$. The proof of Proposition 3 (provided in the Online Supplement) gives the values of $\lambda(x)$ and $\mu(x)$ as

$$\lambda(x) = \sup_{\omega\in\Omega} h_\omega(x) - \mathrm{VaR}_\gamma\,[\mathbf{h}(x)]\,, \quad \mu(x) = \frac{1}{2}\left(\sup_{\omega\in\Omega} h_\omega(x) + \mathrm{VaR}_\gamma\,[\mathbf{h}(x)]\right). \tag{5}$$

Throughout the paper, we will be repeatedly referring to the cases $\lambda(x) > 0$ and $\lambda(x) = 0$, which, using (5), correspond to the conditions $\mathrm{VaR}_\gamma\,[\mathbf{h}(x)] < \sup_{\omega\in\Omega} h_\omega(x)$ and $\mathrm{VaR}_\gamma\,[\mathbf{h}(x)] = \sup_{\omega\in\Omega} h_\omega(x)$, respectively. When $\lambda(x) = 0$, we have $\Omega_2(x) = \Omega_4(x)$ and $\Omega_3(x) = \emptyset$. Thus, $\Omega_4(x)$ and its complement, i.e., $\Omega_4^{\mathsf{c}}(x) (= \Omega_1(x))$, partition the scenario set Ω when $\lambda(x) = 0$. We collectively refer to these sets as *primal categories*. In our notation, we suppress the dependence of these sets on γ for ease of exposition.

4 Identifying effectiveness of scenarios for DRSP-V

In this section we investigate the effectiveness of scenarios for (DRSP-V). First, we derive the optimal (i.e., worst-case) probabilities of the worst-case expected problem in (DRSP-V) in Sect. 4.1. We then discuss the assessment problem of (DRSP-V) in Sect. 4.2. These two sections provide the necessary foundations to identify the effectiveness of scenarios. We state our main results regarding the effectiveness of scenarios in Sect. 4.3, with proofs in Sect. 4.4.

4.1 Characterization of the worst-case probabilities

For a fixed $x \in \mathbb{X}$, suppose $\mathbf{p} := \mathbf{p}(x)$ solves the worst-case expected problem in (DRSP-V) at x, $\lambda(x)$ and $\mu(x)$ are the corresponding optimal dual variables, given by (5). Proposition 4 shows that the worst-case probabilities of scenarios in sets $\Omega_1(x)$

and $\Omega_3(x)$ are precisely determined when $\lambda(x) > 0$. Moreover, the worst-case probabilities of scenarios in $\Omega_2(x)$ and $\Omega_4(x)$ are bounded by the nominal probabilities.

Proposition 4 *Consider a fixed $x \in \mathbb{X}$ and $\lambda(x)$ given in (5).*

(i) When $\lambda(x) > 0$, the following characterizes the worst-case probabilities corresponding to x:

$$\begin{cases} p_\omega = 0, & \omega \in \Omega_1(x), \\ p_\omega \leq q_\omega, & \omega \in \Omega_2(x), \\ p_\omega = q_\omega, & \omega \in \Omega_3(x), \\ p_\omega \geq q_\omega, & \omega \in \Omega_4(x), \end{cases} \tag{6}$$

coupled with additional constraints

$$\sum_{\omega \in \Omega_2(x)} p_\omega = \sum_{\omega \in \Omega_1(x) \cup \Omega_2(x)} q_\omega - \gamma, \tag{7}$$

$$\sum_{\omega \in \Omega_4(x)} p_\omega = \gamma + \sum_{\omega \in \Omega_4(x)} q_\omega, \tag{8}$$

in addition to $\sum_{\omega \in \Omega} p_\omega = 1$ and $p_\omega \geq 0$, for all $\omega \in \Omega$.

(ii) When $\lambda(x) = 0$, the conditions can be written as $p_\omega \geq 0$ for all $\omega \in \Omega_4(x)$, and $p_\omega = 0$ otherwise (i.e., $\omega \in \Omega_4^{\mathsf{c}}(x)$), in addition to $\sum_{\omega \in \Omega} p_\omega = 1$ and $\frac{1}{2} \sum_{\omega \in \Omega} |p_\omega - q_\omega| \leq \gamma$.

The above proposition can be proved by writing the KKT conditions for a linearized version of the worst-case expected problem in (DRSP-V). We skip the proof for brevity and present it in the Appendix. See also [25] for a related result that characterizes one particular worst-case distribution.

With $\lambda(x) > 0$, when the sets $\Omega_2(x)$ and $\Omega_4(x)$ are singletons, the worst-case probabilities of scenarios in $\Omega_2(x)$ and $\Omega_4(x)$ are uniquely determined. A similar argument holds when $\lambda(x) = 0$. Corollary 2 summarizes this result. Then, Corollary 3 sheds more light on the worst-case probabilities of scenarios at the worst-case cost. The proofs of the corollaries are immediate from Proposition 4 and hence skipped.

Corollary 2 *Consider a fixed $x \in \mathbb{X}$. The worst-case probability of scenario ω' is uniquely determined under the following conditions:*

(i) When $\lambda(x) > 0$, whenever $\Omega_2(x) = \{\omega'\}$, we have $p_{\omega'} = \sum_{\omega \in \Omega_1(x) \cup \Omega_2(x)} q_\omega - \gamma$.
(ii) When $\lambda(x) > 0$, whenever $\Omega_4(x) = \{\omega'\}$, we have $p_{\omega'} = \gamma + q_{\omega'}$.
(iii) When $\lambda(x) = 0$, whenever $\Omega_4(x) = \{\omega'\}$, we have $p_{\omega'} = 1$.

Corollary 3 *For a fixed $x \in \mathbb{X}$, there is at least one scenario $\omega \in \Omega_4(x)$ with $p_\omega > q_\omega$ when $\lambda(x) > 0$. Similarly, there is at least one scenario $\omega \in \Omega_4(x)$ with $p_\omega > 0$ when $\lambda(x) = 0$.*

4.2 Assessment problem: formulation and notation

For (DRSP-V), the assessment problem of scenarios in $\mathcal{F}$, defined in (2), can be equivalently written as

$$\min_{x\in\mathbb{X}} \left\{ f_\gamma^{\mathrm{A}}(x;\mathcal{F}) = \max_{\mathbf{p}\in\mathcal{P}_\gamma^{\mathrm{A}}(\mathcal{F})} \sum_{\omega\in\mathcal{F}^{\mathsf{C}}} p_\omega h_\omega(x) \right\}, \tag{9}$$

where

$$\mathcal{P}^{\mathrm{A}}(\mathcal{F}) = \left\{ \mathbf{p} \,:\, \mathrm{V}(\mathbf{p},\mathbf{q};\mathcal{F}) \le \gamma - \frac{1}{2}\sum_{\omega\in\mathcal{F}} q_\omega,\ \sum_{\omega\in\mathcal{F}^{\mathsf{C}}} p_\omega = 1,\ p_\omega \ge 0,\ \forall \omega\in\mathcal{F}^{\mathsf{C}} \right\}. \tag{10}$$

Above, $\mathrm{V}(\mathbf{p},\mathbf{q};\mathcal{F}) := \frac{1}{2}\sum_{\omega\in\mathcal{F}^{\mathsf{C}}} |p_\omega - q_\omega|$. Let $\mathbb{Q}(\mathcal{F}) := \sum_{\omega\in\mathcal{F}} q_\omega$. If $0 < \gamma < \mathbb{Q}(\mathcal{F})$, then $\mathcal{P}_\gamma^{\mathrm{A}}(\mathcal{F})$ is empty because the first constraint in (10) is violated.[1] Hence, for such γ, the assessment problem of scenarios in $\mathcal{F}$ is not well defined. In this case, by convention set forth in Sect. 2, $f_\gamma^{\mathrm{A}}(x;\mathcal{F}) = -\infty$, and such $\mathcal{F}$ is effective at these relatively small γ values.

We need to introduce some more notation for the assessment problem in order to state our main results in Sect. 4.3 and provide the proofs in Sect. 4.4. Let us define

$$\gamma_{\mathcal{F}} := \frac{\gamma - \mathbb{Q}(\mathcal{F})}{1 - \mathbb{Q}(\mathcal{F})} \tag{11}$$

for $\mathbb{Q}(\mathcal{F}) \le \gamma \le 1$, i.e., $0 \le \gamma_{\mathcal{F}} \le 1$. For a fixed $\eta \in \mathbb{R}$, let $\Psi_{|\mathcal{F}^{\mathsf{C}}}(x,\eta)$ denote the conditional version of $\Psi(x,\eta)$. That is, $\Psi_{|\mathcal{F}^{\mathsf{C}}}(x,\eta) := \sum_{\omega\in\mathcal{F}^{\mathsf{C}}\cap[\mathbf{h}(x)\le\eta]} q_{\omega|\mathcal{F}^{\mathsf{C}}}$, where $q_{\omega|\mathcal{F}^{\mathsf{C}}} := \frac{q_\omega}{1-\mathbb{Q}(\mathcal{F})}$ is the probability mass function of scenario ω conditional on $\mathcal{F}^{\mathsf{C}}$. Also, let $\inf\{\eta \,:\, \Psi_{|\mathcal{F}^{\mathsf{C}}}(x,\eta) \ge \gamma_{\mathcal{F}}\}$ be the VaR of $\mathbf{h}(x)$ at level $0 \le \gamma_{\mathcal{F}} \le 1$ *conditioned on* $\mathcal{F}^{\mathsf{C}}$, denoted by $\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{C}}\right]$.

4.3 Main results

We summarize our main results in three categories: (i) First, we study the effectiveness of a single scenario, i.e., $\mathcal{F} = \{\omega'\}$ (Theorems 1–3); we call these results "easy-to-check" conditions. (ii) Then, we study the effectiveness of a non-singleton subset of scenarios (Theorem 4). (iii) Finally, we show that existence of alternative optimal worst-case probabilities is a necessary condition for having a zero-probability effective scenario and/or a positive-probability ineffective scenario (Theorem 5; see also Example 1).

---

[1] Observe that by the triangle inequality, $\mathrm{V}(\mathbf{p},\mathbf{q};\mathcal{F}) \ge \frac{1}{2}|\sum_{\omega\in\mathcal{F}^{\mathsf{C}}}(p_\omega - q_\omega)|$. Because the second constraint in (10) dictates $\sum_{\omega\in\mathcal{F}^{\mathsf{C}}} p_\omega = 1$, we have $\mathrm{V}(\mathbf{p},\mathbf{q};\mathcal{F}) \ge \frac{1}{2}\mathbb{Q}(\mathcal{F})$. When $0 < \gamma < \mathbb{Q}(\mathcal{F})$, we have $\gamma - \frac{1}{2}\mathbb{Q}(\mathcal{F}) < \frac{1}{2}\mathbb{Q}(\mathcal{F})$, resulting in an infeasibility in (10).

Let us first consider a singleton subset $\mathcal{F} = \{\omega'\}$. When $\gamma = 0$, every scenario with a positive nominal probability is effective because the ambiguity set (10) is an empty set, and hence the assessment problem is not well defined for such scenarios. On the other hand, every scenario ω' with a zero nominal probability is ineffective because $S = S^{\mathrm{A}}(\{\omega'\})$. We suppose that we are not in the trivial case $\gamma = 0$. As discussed in Sect. 2.2, we want to provide easy-to-check conditions that help decision makers identify the effectiveness of a scenario by using the information they obtain from an optimal solution $(x^*, \mathbf{p}^*)$ to (DRSP-V) with $\gamma > 0$. Theorems 1–3 below summarize the characteristics of the ineffective and effective scenarios that are identifiable using easy-to-check conditions.

Theorem 1 (Easy-to-check conditions for ineffective scenarios) *Suppose* $(x^*, \mathbf{p}^*)$ *solves* (DRSP-V)*, and let* $\lambda^* := \lambda(x^*)$ *and* $\mu^* := \mu(x^*)$ *be the optimal dual variables at* x^**, given by* (5)*. For a scenario* ω' *with* $q_{\omega'} \leq \gamma$*, consider the following conditions:*

(i) $\omega' \in \Omega_1(x^*)$,
(ii) $\omega' \in \Omega_2(x^*)$ *and* $q_{\omega'} = 0$,
(iii) $\omega' \in \Omega_2(x^*)$ *and* $\sum_{\omega\in\Omega_2(x^*)} p^*_\omega = 0$,
(iv) $\omega' \in \Omega_3(x^*)$ *and* $q_{\omega'} = 0$.

When $\lambda^* > 0$*, scenario* ω' *is ineffective for* (DRSP-V) *if any of conditions (i)–(iv) holds. When* $\lambda^* = 0$*, scenario* ω' *is ineffective for* (DRSP-V) *if condition (i) holds.*

Theorem 2 (Easy-to-check conditions for effective scenarios) *Consider* (DRSP-V) *and notation defined in Theorem* 1*. For a scenario* ω'*, consider the following conditions:*

(i) $q_{\omega'} > \gamma$,
(ii) $\Omega_2(x^*) = \{\omega'\}$ *and* $p^*_{\omega'} > 0$,
(iii) $\omega' \in \Omega_3(x^*)$ *and* $q_{\omega'} > 0$,
(iv) $\omega' \in \Omega_4(x^*)$ *and* $q_{\omega'} > 0$,
(v) $\Omega_4(x^*) = \{\omega'\}$.

When $\lambda^* > 0$*, scenario* ω' *is effective for* (DRSP-V) *if any of the conditions (i)–(v) holds. When* $\lambda^* = 0$*, scenario* ω' *is effective for* (DRSP-V) *if either condition (i) or (v) holds.*

Theorems 1 and 2 identify only a subset of ineffective and effective scenarios. For instance, let us consider a scenario ω' with $q_{\omega'} \leq \gamma$ in the non-singleton VaR and/or the worst-case-cost sets. When $\lambda^* = 0$, the effectiveness of scenario $\omega' \in \Omega_2(x^*)(= \Omega_4(x^*))$ is "undetermined". When $\lambda^* > 0$, the effectiveness of scenario $\omega' \in \Omega_2(x^*)$ with $q_{\omega'} > 0$ and $\sum_{\omega\in\Omega_2(x^*)} p^*_\omega > 0$, and the effectiveness of scenario $\omega' \in \Omega_4(x^*)$ with $q_{\omega'} = 0$ are undetermined. All the above sufficient conditions are based on the cost and probability (worst-case p^*_ω or nominal q_ω) of a scenario. Theorem 3 below uses additional information about $\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right]$ to identify the effectiveness of an undetermined scenario. This allows us to classify additional scenarios as effective that were not identified by Theorems 1 and 2.

Theorem 3 *Consider* (DRSP-V) *and notation defined in Theorem* 1*. In addition, for a scenario* $\omega' \in \Omega_2(x^*)$ *with* $q_{\omega'} > 0$*, suppose* $\gamma_{\mathcal{F}}$ *is defined as in* (11) *for* $\mathcal{F} = \{\omega'\}$*. Suppose that the effectiveness of scenario* ω' *is not identified by Theorems* 1 *and* 2*. If* *(i)* $\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{C}}\right] < \mathrm{VaR}_{\gamma}\left[\mathbf{h}(x^*)\right]$, *and*

(ii) *either there exists a scenario* $\omega \in \left[\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{C}}\right] < \mathbf{h}(x^*) < \mathrm{VaR}_{\gamma}\left[\mathbf{h}(x^*)\right]\right]$ *with* $q_\omega > 0$ *or* $\Psi_{|\mathcal{F}^{\mathsf{C}}}\left(x^*, \mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{C}}\right]\right) > \gamma_{\mathcal{F}}$,

then scenario ω' *is effective for* (DRSP-V).

Together, Theorems 1–3 identify the effectiveness of all scenarios, except for some scenarios in the non-singleton VaR and worst-case-cost sets, with positive and zero nominal probabilities, respectively. We study the performance of these theorems in identifying the effectiveness of scenarios numerically in Sect. 5.2. Our numerical results indicate that the number of undetermined scenarios can be relatively small.

All the results we have stated so far considered the effectiveness of a single scenario. We already know the effectiveness of union of effective scenarios and intersection of ineffective scenarios with other subsets from Proposition 2. The following theorem addresses the effectiveness of a union of ineffective scenarios for (DRSP-V).

Theorem 4 *For* (DRSP-V)*, the union of any of the ineffective scenarios identified in Theorem* 1 *is ineffective.*

One might ask what happens to the effectiveness of the union of ineffective scenarios that are *not* identified in Theorem 1. For instance, one can ask whether the union of two ineffective scenarios $\omega', \omega'' \in \Omega_4(x^*)$ with $q_{\omega'} = q_{\omega''} = 0$ is ineffective or not. The following example shows the union of such scenarios might be effective.

Example 2 Let us consider an inventory problem with the following specifics: $n = 4$, $\mathbf{q} = (0, 0.5, 0.5, 0)$, and $\mathbf{d} = (1, 2, 3, 4)$. Also, ordering cost c is 1, backorder cost b is 4, and holding cost m is 8. For this problem, the cost function is written as $h_\omega(x) = cx + b(d_\omega - x)_+ + m(x - d_\omega)_+$. At $\gamma = 0.15$, we have $x^* = 2$, $f_\gamma(x^*) = 5.2$, and scenarios $d = 1$ and $d = 4$ belong to $\Omega_4(x^*)$. Note that Theorem 1 cannot identify the effectiveness of scenarios $d = 1$ and $d = 4$. However, one can observe by solving the assessment problems of scenario $d = 1$ and $d = 4$ that both scenarios are individually ineffective. However, when the assessment problem of the union of scenarios $d = 1$ and $d = 4$ is solved, the optimal solution and the optimal value are 2 and 4.6, respectively. Thus, the union of two ineffective scenarios $d = 1$ and $d = 4$ is effective. □

We now present two necessary conditions to observe either a zero-probability effective scenario or a positive-probability ineffective scenario for (DRSP-V) as motivated by Example 1.

Theorem 5 *Suppose* $(x^*, \mathbf{p}^*)$ *solves* (DRSP-V)*. If there is either a zero-probability effective scenario* ω' *(*$p^*_{\omega'} = 0$*) or a positive-probability ineffective scenario* ω' *(*$p^*_{\omega'} > 0$*) for* (DRSP-V)*, then, there is a scenario* ω'' *with the same cost as that of* ω'*, i.e.,* $h_{\omega'}(x^*) = h_{\omega''}(x^*)$*, and the (primal) worst-case expected problem at* x^* *has multiple optimal solutions.*

Theorem 5 ties a zero-probability effective scenario and a positive-probability ineffective scenario to multiple optimal solutions of the worst-case expected problem (recall Example 1). To have multiple optimal solutions, the probability of some scenarios must have some leeway (see, for instance, Proposition 4). Thus, it is possible to observe a zero-probability effective scenario and/or a positive-probability ineffective scenario among the set of scenarios at VaR and/or the worst-case cost. In Sect. 4.4, we look into these sets in more detail in order to detect these anomalies. In addition, the proof of Theorem 5 provides guidelines to construct an alternative optimal solution to the worst-case expected problem, for which the worst-case probabilities are more indicative of the effectiveness of scenarios on the optimal value.

4.4 Proofs

First, we discuss the inner maximization problem in the assessment problem of a subset of scenarios in more detail. Then, we prove our main results. For the rest of this section, we consider the notation defined in Theorem 1 for (DRSP-V), unless stated explicitly otherwise.

4.4.1 Preliminary results for the assessment problem

Proposition 5 below shows that the objective function of the assessment problem (9) has a similar interpretation in terms of risk measures as problem (DRSP-V) (cf. Proposition 3), with the nominal distribution $\mathbf{q}$ replaced by a conditional distribution. Notice however that the objective function of the assessment problem is not a direct version of the objective function in Proposition 3 for the conditional distribution $\mathbf{q}$. Here, even though the weights on the sup and CVaR terms are the same as their original counterparts, the level of CVaR for the assessment problem is adjusted from γ to $\gamma_{\mathcal{F}}$ defined in (11).

Proposition 5 *Consider a subset $\mathcal{F} \subset \Omega$ and a fixed $x \in \mathbb{X}$, and suppose $\gamma_{\mathcal{F}}$ is defined as in* (11). *Then, for $\mathbb{Q}(\mathcal{F}) \leq \gamma \leq 1$, the inner maximization problem in* (9) *can be written as*

$$f_{\gamma}^{A}(x;\mathcal{F}) = \gamma \sup_{\omega \in \mathcal{F}^{\mathsf{c}}} h_{\omega}(x) + (1-\gamma)\,\mathrm{CVaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right], \tag{12}$$

where $\mathrm{CVaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]$ is the conditional CVaR of $\mathbf{h}(x)$ at level $\gamma_{\mathcal{F}}$ with respect to the conditional distribution $\mathbf{q}_{\cdot|\mathcal{F}^{\mathsf{c}}}$ defined in Sect. 4.2.

Again, we adopted the convention that $\mathrm{CVaR}_{\beta}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right] = \mathbb{E}_{\mathbf{q}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]$ for $\beta = 0$ and $\mathrm{CVaR}_{\beta}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right] = \sup_{\omega \in \mathcal{F}^{\mathsf{c}}} h_{\omega}(x)$ for $\beta = 1$.

Proof For a fixed $x \in \mathbb{X}$, the Lagrangian dual formulation of the inner maximization problem in (9) can be derived as[2]

[2] For details of this derivation, we refer to [5,27,28] for the Lagrangian dual formulation of (DRSP) via ϕ-divergences.

$$f_\gamma^{\mathrm{A}}(x;\mathcal{F}) = \min_{\lambda,\mu}\left\{\mu - \frac{\lambda}{2}(1-2\gamma) + \sum_{\omega\in\mathcal{F}^{\mathsf{c}}} q_\omega \left(h_\omega(x) - \mu + \frac{\lambda}{2}\right)_+ : \right. \\ \left. \mu + \frac{\lambda}{2} \geq \sup_{\omega\in\mathcal{F}^{\mathsf{c}}} h_\omega(x),\ \lambda \geq 0\right\} \tag{13}$$

for $\mathbb{Q}(\mathcal{F}) \leq \gamma \leq 1$, where λ and μ denote Lagrange multipliers for the first and second constraints in (10), respectively. This problem can be linearized as

$$\min_{\lambda,\mu,\mathbf{r}}\left\{\mu - \frac{\lambda}{2}(1-2\gamma) + \sum_{\omega\in\mathcal{F}^{\mathsf{c}}} q_\omega r_\omega : \right. \\ \left. \mu + \frac{\lambda}{2} \geq \sup_{\omega\in\mathcal{F}^{\mathsf{c}}} h_\omega(x),\ \mu - \frac{\lambda}{2} + r_\omega \geq h_\omega(x),\ r_\omega \geq 0,\ \forall\omega\in\mathcal{F}^{\mathsf{c}},\ \lambda\geq 0\right\}.$$

Let t and z_ω denote the corresponding dual variables for the first and second constraint above, respectively. The dual of the above problem can be derived as

$$\max_{t,\mathbf{z}}\left\{t \sup_{\omega\in\mathcal{F}^{\mathsf{c}}} h_\omega(x) + \sum_{\omega\in\mathcal{F}^{\mathsf{c}}} z_\omega h_\omega(x) : \right. \\ \left. t + \sum_{\omega\in\mathcal{F}^{\mathsf{c}}} z_\omega = 1,\ t - \sum_{\omega\in\mathcal{F}^{\mathsf{c}}} z_\omega \leq 2\gamma - 1,\ 0 \leq z_\omega \leq q_\omega,\ \forall\omega\in\mathcal{F}^{\mathsf{c}},\ t\geq 0\right\}. \tag{14}$$

Let us define $t_{\mathcal{F}} := \frac{t-\mathbb{Q}(\mathcal{F})}{\mathbb{Q}(\mathcal{F}^{\mathsf{c}})}$ and $\zeta := \inf\{0 \leq t \leq 1 : \mathrm{VaR}_{t_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right] = \sup_{\omega\in\mathcal{F}^{\mathsf{c}}} h_\omega(x)\}$. We claim the set of optimal solutions $(t,\mathbf{z})$ to the above problem is given by $\mathcal{T}\times\mathcal{Z}_\gamma^{\mathrm{A}}(x;\mathcal{F})$, where $\mathcal{T} := [\min\{\zeta,\gamma\},\ \gamma\}]$ and $\mathcal{Z}_\gamma^{\mathrm{A}}(x;\mathcal{F})$ is given by

$$\mathcal{Z}_\gamma^{\mathrm{A}}(x;\mathcal{F}) := \left\{\begin{array}{ll} z_\omega = 0, & \text{if } \omega\in\mathcal{F}^{\mathsf{c}}\cap\left[\mathbf{h}(x) < \mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]\right], \\ z_\omega = q_\omega, & \text{if } \omega\in\mathcal{F}^{\mathsf{c}}\cap\left[\mathbf{h}(x) > \mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]\right], \\ \mathbf{z} : z_\omega\in[0,q_\omega], & \text{if } \omega\in\mathcal{F}^{\mathsf{c}}\cap\left[\mathbf{h}(x) = \mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]\right], \\ \sum_{\omega\in\mathcal{W}} z_\omega & = \mathbb{Q}(\mathcal{F}^{\mathsf{c}})\left[\Psi_{|\mathcal{F}^{\mathsf{c}}}\left(x, \mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]\right) - \gamma_{\mathcal{F}}\right], \end{array}\right. \tag{15}$$

in which the last summation is taken over the set $\mathcal{W} := \mathcal{F}^{\mathsf{c}}\cap\left[\mathbf{h}(x) = \mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]\right]$. To see this claim, note that we can write the dual problem as

$$\max_{0\leq t\leq\gamma} t \sup_{\omega\in\mathcal{F}^{\mathsf{c}}} h_\omega(x) + \max_{\mathbf{z}}\left\{\sum_{\omega\in\mathcal{F}^{\mathsf{c}}} z_\omega h_\omega(x) : \sum_{\omega\in\mathcal{F}^{\mathsf{c}}} z_\omega = 1-t,\ 0\leq z_\omega\leq q_\omega,\ \forall\omega\in\mathcal{F}^{\mathsf{c}}\right\}.$$

The maximization problem over $\mathbf{z}$ in the above problem can be written as

$$\mathbb{Q}(\mathcal{F}^{\mathsf{c}})(1-t_{\mathcal{F}}) \max_{\mathbf{z}'}\left\{\sum_{\omega\in\mathcal{F}^{\mathsf{c}}} z'_\omega h_\omega(x) : \sum_{\omega\in\mathcal{F}^{\mathsf{c}}} z'_\omega = 1,\ 0\leq z'_\omega \leq \frac{q_{\omega|\mathcal{F}^{\mathsf{c}}}}{1-t_{\mathcal{F}}},\ \forall\omega\in\mathcal{F}^{\mathsf{c}}\right\},$$

where $\mathbf{z}' := \frac{\mathbf{z}}{\mathbb{Q}(\mathcal{F}^{\mathsf{c}})(1-t_{\mathcal{F}})}$. By Lemma A-1, the maximization problem over $\mathbf{z}'$ above is equivalent to $\mathrm{CVaR}_{t_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]$ and the set of its optimal solutions is given by

$$\mathcal{Z}'_t(x;\mathcal{F}) := \left\{ \mathbf{z}' : \begin{array}{ll} z'_\omega = 0, & \text{if } \omega \in \mathcal{F}^{\mathsf{c}} \cap \left[\mathbf{h}(x) < \mathrm{VaR}_{t_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]\right], \\ z'_\omega = \frac{q_{\omega|\mathcal{F}^{\mathsf{c}}}}{1-t_{\mathcal{F}}}, & \text{if } \omega \in \mathcal{F}^{\mathsf{c}} \cap \left[\mathbf{h}(x) > \mathrm{VaR}_{t_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]\right], \\ z'_\omega \in [0, \frac{q_{\omega|\mathcal{F}^{\mathsf{c}}}}{1-t_{\mathcal{F}}}], & \text{if } \omega \in \mathcal{F}^{\mathsf{c}} \cap \left[\mathbf{h}(x) = \mathrm{VaR}_{t_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]\right], \\ \multicolumn{2}{l}{\sum_{\omega \in \mathcal{F}^{\mathsf{c}} \cap \left[\mathbf{h}(x) = \mathrm{VaR}_{t_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]\right]} z'_\omega = \frac{1}{1-t_{\mathcal{F}}}\left[\Psi_{|\mathcal{F}^{\mathsf{c}}}\left(x, \mathrm{VaR}_{t_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]\right) - t_{\mathcal{F}}\right].} \end{array} \right.$$

Now, let $B(t) := t \sup_{\omega \in \mathcal{F}^{\mathsf{c}}} h_\omega(x) + \mathbb{Q}(\mathcal{F}^{\mathsf{c}})(1 - t_{\mathcal{F}})\mathrm{CVaR}_{t_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]$. We can equivalently write $B(t) = t \sup_{\omega \in \mathcal{F}^{\mathsf{c}}} h_\omega(x) + \mathbb{Q}(\mathcal{F}^{\mathsf{c}}) \int_{t_{\mathcal{F}}}^1 \mathrm{VaR}_\beta\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right] d\beta$ using the definition of $\mathrm{CVaR}_{t_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]$. Observe that $B(t)$ is non-decreasing and concave in t because $B'(t) = \sup_{\omega \in \mathcal{F}^{\mathsf{c}}} h_\omega(x) - \mathrm{VaR}_{t_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right] \geq 0$ and $B'(t)$ is non-increasing in t. Because $\zeta = \inf\{0 \leq t \leq 1 : B'(t) = 0\}$ by definition, the maximum of $B(t)$ over $0 \leq t \leq \gamma$ is achieved on $\mathcal{T}$. Now, if $\zeta < \gamma$, then for all $t \in [\zeta, \ \gamma]$, we have $\mathrm{VaR}_{t_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right] = \mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right] = \sup_{\omega \in \mathcal{F}^{\mathsf{c}}} h_\omega(x)$ and also $\Psi_{|\mathcal{F}^{\mathsf{c}}}\left(x, \mathrm{VaR}_{t_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]\right) = \Psi_{|\mathcal{F}^{\mathsf{c}}}\left(x, \mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]\right) = 1$. Hence, $\mathcal{Z}'_t(x;\mathcal{F}) = \mathcal{Z}'_\gamma(x;\mathcal{F})$ for all $t \in [\zeta, \ \gamma]$. As a result, the set of optimal solutions $(t, \mathbf{z}')$ to the dual problem is given by $\mathcal{T} \times \mathcal{Z}'_\gamma(x;\mathcal{F})$. Because for such an optimal $\mathbf{z}'$ we have $\mathbf{z}' = \frac{\mathbf{z}}{\mathbb{Q}(\mathcal{F}^{\mathsf{c}})(1-\gamma_{\mathcal{F}})}$, the set of optimal solutions $(t, \mathbf{z})$ to (14) is given by $\mathcal{T} \times \mathcal{Z}^{\mathrm{A}}_\gamma(x;\mathcal{F})$. Now, by substituting any optimal $(t, \mathbf{z})$ in the objective function of (14) we obtain (12), where $\mathrm{CVaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right] = \mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right] + \frac{1}{1-\gamma_{\mathcal{F}}} \sum_{\omega \in \mathcal{F}^{\mathsf{c}}} q_{\omega|\mathcal{F}^{\mathsf{c}}}\left(h_\omega(x) - \mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]\right)_+$. $\square$

Remark 3 Consider a subset $\mathcal{F} \subset \Omega$ and a fixed $x \in \mathbb{X}$, and suppose $\gamma_{\mathcal{F}}$ is defined as in (11). Then, for $\mathbb{Q}(\mathcal{F}) \leq \gamma \leq 1$, some of the proofs that come later need to calculate $\mathrm{CVaR}_\gamma\,[\mathbf{h}(x)]$, $\mathrm{CVaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]$, and their subdifferentials. By the proof of Proposition 5, we have $(1-\gamma)\mathrm{CVaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right] = \sum_{\omega \in \mathcal{F}^{\mathsf{c}}} z_\omega h_\omega(x)$ for any $\mathbf{z} \in \mathcal{Z}^{\mathrm{A}}_\gamma(x;\mathcal{F})$, as defined in (15). That is,

$$\begin{aligned}(1-\gamma)\,\mathrm{CVaR}_{\gamma_{\mathcal{F}}}&\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right] \\ &= \mathbb{Q}(\mathcal{F}^{\mathsf{c}})\Bigg(\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]\left[\Psi_{|\mathcal{F}^{\mathsf{c}}}\left(x, \mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]\right) - \gamma_{\mathcal{F}}\right] \\ &\quad + \sum_{\omega \in \mathcal{F}^{\mathsf{c}} \cap \left[\mathbf{h}(x) > \mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]\right]} q_{\omega|\mathcal{F}^{\mathsf{c}}} h_\omega(x)\Bigg).\end{aligned} \tag{16}$$

By an application of [42, Theorem 6.14], we also have $(1-\gamma)\,\partial\mathrm{CVaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right] = \bigcup_{\mathbf{z} \in \mathcal{Z}^{\mathrm{A}}_\gamma(x;\mathcal{F})} \sum_{\omega \in \mathcal{F}^{\mathsf{c}}} z_\omega \partial h_\omega(x)$. That is,

$$(1-\gamma)\,\partial\mathrm{CVaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right] = \bigcup_{\mathbf{z}\in\mathcal{Z}_{\gamma}^{\mathrm{A}}(x;\mathcal{F})} \sum_{\omega\in\mathcal{F}^{\mathsf{c}}\cap\left[\mathbf{h}(x)=\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]\right]} z_{\omega}\partial h_{\omega}(x) + \sum_{\omega\in\mathcal{F}^{\mathsf{c}}\cap\left[\mathbf{h}(x)>\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x)|\mathcal{F}^{\mathsf{c}}\right]\right]} q_{\omega}\partial h_{\omega}(x). \tag{17}$$

In particular, when $\mathcal{F} = \emptyset$, we can obtain $(1-\gamma)\,\mathrm{CVaR}_{\gamma}\,[\mathbf{h}(x)]$ from (16) and $(1-\gamma)\,\partial\mathrm{CVaR}_{\gamma}\,[\mathbf{h}(x)]$ from (17). For future references, we denote $\mathcal{Z}_{\gamma}^{\mathrm{A}}(x;\emptyset)$ as $\mathcal{Z}_{\gamma}(x)$.

4.4.2 Proof of easy-to-check conditions

Before we proceed with proving our main results stated in Sect. 4.3, we comment on the proof technique of the following intermediate results. Many of these proofs share a similar foundation as follows: First, for a subset $\mathcal{F}$ of scenarios, if the primal category of scenario $\omega' \in \mathcal{F}$ is not given, we determine the primal category to which ω' belongs using the optimality conditions in Proposition 4, given $p^*_{\omega'}$ and/or $q_{\omega'}$. Next, if needed, we calculate $\mathrm{CVaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right]$, $\sup_{\omega\in\mathcal{F}^{\mathsf{c}}} h_{\omega}(x^*)$, $f_{\gamma}^{\mathrm{A}}(x^*;\mathcal{F})$, and their subdifferentials and compare them to $\mathrm{CVaR}_{\gamma}\left[\mathbf{h}(x^*)\right]$, $\sup_{\omega\in\Omega} h_{\omega}(x^*)$, $f_{\gamma}(x^*)$, and their subdifferentials, respectively, to check the optimality of x^* for the assessment problem. This is summarized in Lemma 1. Finally, we use Definition 1 to identify the effectiveness of subset $\mathcal{F}$.

We first state a useful lemma. Note that the lemma makes references to the sets $\mathcal{Z}_{\gamma}^{\mathrm{A}}(x;\mathcal{F})$ and $\mathcal{Z}_{\gamma}(x) = \mathcal{Z}_{\gamma}^{\mathrm{A}}(x;\emptyset)$, defined in (15). Propositions 6 and 7 then show sufficient conditions for a scenario to be effective or ineffective.

Lemma 1 *Consider a subset $\mathcal{F}$ such that $\Omega_4(x^*)\cap\mathcal{F} = \emptyset$, and suppose $\gamma_{\mathcal{F}}$ is defined as in (11). If the set $\mathcal{Z}_{\gamma}^{\mathrm{A}}(x^*;\mathcal{F})$ coincides with the restriction of the set $\mathcal{Z}_{\gamma}(x^*)$ to $\mathcal{F}^{\mathsf{c}}$, then*

(i) $\mathrm{CVaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right] = \mathrm{CVaR}_{\gamma}\left[\mathbf{h}(x^*)\right]$ *and* $\partial\mathrm{CVaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right] = \partial\mathrm{CVaR}_{\gamma}[\mathbf{h}(x^*)]$,
(ii) $\sup_{\omega\in\mathcal{F}^{\mathsf{c}}} h_{\omega}(x^*) = \sup_{\omega\in\Omega} h_{\omega}(x^*)$ *and* $\partial\sup_{\omega\in\mathcal{F}^{\mathsf{c}}} h_{\omega}(x^*) = \partial\sup_{\omega\in\Omega} h_{\omega}(x^*)$, *and*
(iii) *$\mathcal{F}$ is ineffective for* (DRSP-V).

Proof Part (i) is immediate from Remark 3. In part (ii), the fact that $\Omega_4(x^*)\cap\mathcal{F} = \emptyset$ implies $\sup_{\omega\in\mathcal{F}^{\mathsf{c}}} h_{\omega}(x^*) = \sup_{\omega\in\Omega} h_{\omega}(x^*)$ and also $\partial\sup_{\omega\in\mathcal{F}^{\mathsf{c}}} h_{\omega}(x^*) = \partial\sup_{\omega\in\Omega} h_{\omega}(x^*)$ by [19, Corollary 4.3.2]. Finally, using parts (i) and (ii), we have $\partial f_{\gamma}^{\mathrm{A}}(x^*;\mathcal{F}) = \gamma\partial\sup_{\omega\in\mathcal{F}^{\mathsf{c}}} h_{\omega}(x^*) + (1-\gamma)\,\partial\mathrm{CVaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right] = \gamma\partial\sup_{\omega\in\Omega} h_{\omega}(x^*) + (1-\gamma)\,\partial\mathrm{CVaR}_{\gamma}\left[\mathbf{h}(x^*)\right] = \partial f_{\gamma}(x^*)$ and $f_{\gamma}^{\mathrm{A}}(x^*;\mathcal{F}) = f_{\gamma}(x^*)$. By optimality of x^* to (DRSP-V), there exists a subgradient $s \in \partial f_{\gamma}^{\mathrm{A}}(x^*;\mathcal{F}) = \partial f_{\gamma}(x^*)$ with $s^{\top}(x-x^*) \geq 0$ for all $x \in \mathbb{X}$. Then, by convexity of $f_{\gamma}^{\mathrm{A}}(x;\mathcal{F})$ in x, we have $f_{\gamma}^{\mathrm{A}}(x;\mathcal{F}) \geq f_{\gamma}^{\mathrm{A}}(x^*;\mathcal{F}) + s^{\top}(x-x^*) \geq f_{\gamma}^{\mathrm{A}}(x^*;\mathcal{F})$ for any $x \in \mathbb{X}$. This implies x^* is a global minimum to the assessment problem with the optimal value $f_{\gamma}^{\mathrm{A}}(x^*;\mathcal{F}) = f_{\gamma}(x^*)$. Therefore, by Definition 1, subset $\mathcal{F}$ is ineffective. □

Proposition 6 *Consider a scenario ω' with $q_{\omega'} = 0$. If $\omega' \notin \Omega_4(x^*)$, then scenario ω' is ineffective for* (DRSP-V) *with $p^*_{\omega'} = 0$.*

Proof First, note that we must have $p^*_{\omega'} = 0$ by Proposition 4. Let $\mathcal{F} = \{\omega'\}$. Because $q_{\omega'} = 0$, we have $\gamma_{\mathcal{F}} = \gamma$ and $q_{\omega|\mathcal{F}^{\mathsf{c}}} = q_\omega$ on $\mathcal{F}^{\mathsf{c}}$. As a result, $\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right] = \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]$ and the set $\mathcal{Z}^{\mathrm{A}}_\gamma(x^*; \mathcal{F})$ coincides with the restriction of the set $\mathcal{Z}_\gamma(x^*)$ to $\mathcal{F}^{\mathsf{c}}$. Now, Lemma 1 completes the proof. □

Proposition 7 *Consider a scenario ω' with $q_{\omega'} > 0$. If $\omega' \notin \Omega_2(x^*)$, then scenario ω' is ineffective for* (DRSP-V) *if $p^*_{\omega'} = 0$; and scenario ω' is effective for* (DRSP-V) *if $p^*_{\omega'} > 0$.*

The proof of the above proposition uses Lemma A-2.

Proof First, suppose $\omega' \notin \Omega_2(x^*)$ and $p^*_{\omega'} = 0$. Then, we must have $\omega' \in \Omega_1(x^*)$ by Proposition 4. Let $\mathcal{F} = \{\omega'\}$. We first claim $\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right] = \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]$. To see the claim, note by the definition of $\mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]$ that $\Psi\left(x^*, \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]\right) = \sum_{\omega\in\Omega_1(x^*)\cup\Omega_2(x^*)} q_\omega \geq \gamma$ and thus, $\sum_{\omega\in\mathcal{F}^{\mathsf{c}}\cap(\Omega_1(x^*)\cup\Omega_2(x^*))} q_\omega \geq \gamma - \mathbb{Q}(\mathcal{F})$, which implies $\Psi_{|\mathcal{F}^{\mathsf{c}}}\left(x^*, \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]\right) \geq \gamma_{\mathcal{F}}$. So, $\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right] \leq \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]$. By definition, we have $\Psi_{|\mathcal{F}^{\mathsf{c}}}\left(x^*, \mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right]\right) \geq \gamma_{\mathcal{F}}$, which implies $\mathbb{Q}(\mathcal{F}) + \sum_{\omega\in\mathcal{F}^{\mathsf{c}}\cap\left[\mathbf{h}(x^*)\leq\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right]\right]} q_\omega \geq \gamma$. Suppose now that

$$\omega' \in \left[\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right] < \mathbf{h}(x^*) < \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]\right].$$

Then, we have $\sum_{\omega\in[\mathbf{h}(x^*)\leq h_{\omega'}(x^*)]} q_\omega \geq \mathbb{Q}(\mathcal{F}) + \sum_{\omega\in\mathcal{F}^{\mathsf{c}}\cap\left[\mathbf{h}(x^*)\leq\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right]\right]} q_\omega \geq \gamma$, which is a contradiction to the definition of $\mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]$ because $h_{\omega'}(x^*) < \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]$. It follows that we must have $\omega' \in \left[\mathbf{h}(x^*) \leq \mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right]\right]$ and thus, $\Psi\left(x^*, \mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right]\right) = \mathbb{Q}(\mathcal{F}) + \sum_{\omega\in\mathcal{F}^{\mathsf{c}}\cap\left[\mathbf{h}(x^*)\leq\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right]\right]} q_\omega \geq \gamma$, i.e., $\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right] \geq \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]$. Therefore, $\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right] = \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]$. Finally, because $\omega' \in \Omega_1(x^*)$, we have that

$$\mathbb{Q}(\mathcal{F}^{\mathsf{c}})\left[\Psi_{|\mathcal{F}^{\mathsf{c}}}\left(x^*, \mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right]\right) - \gamma_{\mathcal{F}}\right] = \Psi\left(x^*, \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]\right) - \gamma.$$

Therefore, the set $\mathcal{Z}^{\mathrm{A}}_\gamma(x^*; \mathcal{F})$ coincides with the restriction of the set $\mathcal{Z}_\gamma(x^*)$ to $\mathcal{F}^{\mathsf{c}}$. The result follows then from Lemma 1.

Now, suppose $\omega' \notin \Omega_2(x^*)$ and $p^*_{\omega'} > 0$. By Proposition 4, we must have $\lambda^* > 0$ and $\omega' \in \Omega_3(x^*) \cup \Omega_4(x^*)$. Then, by Lemma A-2, we conclude that scenario ω' is effective. □

Remark 4 Propositions 6 and 7 hold in both cases $\lambda^* > 0$ and $\lambda^* = 0$. Recall that when $\lambda^* = 0$, we have $\mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right] = \sup_{\omega\in\Omega} h_\omega(x^*)$ by (5). Thus, in that case, if $\omega' \notin \Omega_4(x^*)$, then by Propositions 6 and 7, we have that scenario ω' is ineffective and $p^*_{\omega'} = 0$, regardless of $q_{\omega'}$.

The next proposition states a sufficient condition for a scenario at the worst-case cost to be effective.

Proposition 8 *Suppose that $\Omega_4(x^*) = \{\omega'\}$ for some scenario ω'. Then, scenario ω' is effective for* (DRSP-V).

Proof Let $\mathcal{F} = \{\omega'\}$. Then, we must have $p^*_{\omega'} > 0$ by Corollary 3, and $p^*_{\omega'} > 0$ is uniquely determined by Corollary 2. Thus, for any feasible solution $\hat{\mathbf{p}}$ to the worst-case expected problem at x^* with $\hat{p}_{\omega'} = 0$, we must have $\sum_{\omega \in \mathcal{F}^{\mathsf{c}}} \hat{p}_\omega h_\omega(x^*) < \sum_{\omega \in \Omega} p^*_\omega h_\omega(x^*) = f_\gamma(x^*)$. Otherwise, $\hat{\mathbf{p}}$ would be optimal to the worst-case expected problem at x^*, which is a contradiction with Corollary 3. Now, if $\bar{\mathbf{p}}$, with $\bar{p}_{\omega'} = 0$, solves the inner maximization problem of the assessment problem of scenario ω' at x^*, we have $f^{\mathrm{A}}_\gamma(x^*; \mathcal{F}) = \sum_{\omega \in \mathcal{F}^{\mathsf{c}}} \bar{p}_\omega h_\omega(x^*) < \sum_{\omega \in \Omega} p^*_\omega h_\omega(x^*) = f_\gamma(x^*)$. Hence, scenario ω' is effective by Definition 1. □

In the following proposition, we provide a sufficient condition to observe only ineffective scenarios in $\Omega_2(x^*)$. When this sufficient condition is satisfied, sum of nominal probabilities of the scenarios in $\Omega_1(x^*) \cup \Omega_2(x^*)$, i.e., the set of scenarios below and at $\mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]$, is exactly γ.

Proposition 9 *Suppose that $\lambda^* > 0$. If $\sum_{\omega \in \Omega_2(x^*)} p^*_\omega = 0$, then the subset $\mathcal{F} = \Omega_2(x^*)$ is ineffective for* (DRSP-V).

Proof Let $\mathcal{F} = \Omega_2(x^*)$. We first show the set $\mathcal{Z}^{\mathrm{A}}_\gamma(x^*; \mathcal{F})$ coincides with the restriction of the set $\mathcal{Z}_\gamma(x^*)$ to $\mathcal{F}^{\mathsf{c}}$. Then, by Lemma 1, we conclude $\mathcal{F}$ is ineffective.

Because $\sum_{\omega \in \Omega_2(x^*)} p^*_\omega = 0$, we have $\sum_{\omega \in \Omega_1(x^*) \cup \Omega_2(x^*)} q_\omega = \gamma$ by (7). This implies for any $\mathbf{z} \in \mathcal{Z}_\gamma(x^*)$, we must have $z_\omega = 0$ on $\mathcal{F}$ because $\sum_{\omega \in \Omega_2(x^*)} z_\omega = \sum_{\omega \in \Omega_1(x^*) \cup \Omega_2(x^*)} q_\omega - \gamma$. Also, note that $\sum_{\omega \in \mathcal{F}^{\mathsf{c}} \cap (\Omega_1(x^*) \cup \Omega_2(x^*))} q_\omega = \gamma - \mathbb{Q}(\mathcal{F})$ implies $\Psi_{|\mathcal{F}^{\mathsf{c}}}\left(x^*, \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]\right) = \gamma_{\mathcal{F}}$. Thus, $\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right] < \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]$ by definition and the fact that $\mathcal{F} = \Omega_2(x^*)$. Moreover, we must have $q_\omega = 0$ for all $\omega \in \mathcal{F}^{\mathsf{c}} \cap \left[\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right] < \mathbf{h}(x^*) < \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]\right]$ because

$$\gamma_{\mathcal{F}} \leq \Psi_{|\mathcal{F}^{\mathsf{c}}}\left(x^*, \mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right]\right) \leq \Psi_{|\mathcal{F}^{\mathsf{c}}}\left(x^*, \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]\right) = \gamma_{\mathcal{F}}.$$

Now, it is easy to verify that the set $\mathcal{Z}^{\mathrm{A}}_\gamma(x^*; \mathcal{F})$ coincides with the restriction of the set $\mathcal{Z}_\gamma(x^*)$ to $\mathcal{F}^{\mathsf{c}}$. □

Note that the condition in Proposition 9 is not necessary. Moreover, we restricted the proposition to the case $\lambda^* > 0$ because otherwise the sufficient condition is not satisfied according to Corollary 3.

In the following proposition, we provide a sufficient and necessary condition so that a scenario ω' in the singleton set $\Omega_2(x^*)$ is effective. We restrict the proposition to the case $\lambda^* > 0$ because otherwise scenario ω' is effective by Proposition 8.

Proposition 10 *Suppose that $\lambda^* > 0$ and $\Omega_2(x^*) = \{\omega'\}$ for some scenario ω'. Then, scenario ω' is effective for* (DRSP-V) *if and only if $p^*_{\omega'} > 0$.*

Proof Let $\mathcal{F} = \{\omega'\}$.

"$\Longrightarrow$": Follows directly from Proposition 9.

"$\Longleftarrow$": By Corollary 2(i), $p^*_{\omega'}$ is uniquely determined. So, a similar argument as the proof of Proposition 8 proves the result. □

Proof of Theorem 1 Part (i) is immediate from Propositions 6 and 7; part (ii) is immediate from Proposition 6; part (iii) is immediate from Proposition 9 and Corollary 1; and part (iv) is immediate from Proposition 6. Case $\lambda^* = 0$ is proved by Remark 4. □

Proof of Theorem 2 Part (i) is trivial by definition; part (ii) is immediate from Proposition 10; part (iii) is immediate from Proposition 7; part (iv) is immediate from Proposition 7; and part (v) is immediate from Proposition 8. When $\lambda^* = 0$, the non-trivial part (v) is proved by Proposition 8. □

Proof of Theorem 3 Because the effectiveness of scenario ω' is not identified by Theorems 1 and 2, then there exists a scenario ω'' such that $h_{\omega'}(x^*) = h_{\omega''}(x^*)$. So, regardless of λ^*, we have $\sup_{\omega\in\mathcal{F}^{\mathsf{c}}} h_\omega(x^*) = \sup_{\omega\in\Omega} h_\omega(x^*)$, where $\mathcal{F} = \{\omega'\}$. Also, notice that $\sum_{\omega\in\mathcal{F}^{\mathsf{c}}\cap(\Omega_1(x^*)\cup\Omega_2(x^*))} q_\omega \geq \gamma - \mathbb{Q}(\mathcal{F})$ implies $\Psi_{|\mathcal{F}^{\mathsf{c}}}\big(x^*, \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]\big) \geq \gamma_{\mathcal{F}}$. So, $\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right] \leq \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]$.

Suppose now that scenario ω' is ineffective. By (3) and Definition 1, we must have $f^{\mathrm{A}}_\gamma(x^*; \mathcal{F}) = f_\gamma(x^*)$. Consequently, $\mathrm{CVaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right] = \mathrm{CVaR}_\gamma\left[\mathbf{h}(x^*)\right]$ because $\sup_{\omega\in\mathcal{F}^{\mathsf{c}}} h_\omega(x^*) = \sup_{\omega\in\Omega} h_\omega(x^*)$. Let us define $\mathcal{V} := \Big[\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right] < \mathbf{h}(x^*) < \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]\Big]$. By Remark 3, the expression $(1-\gamma)\Big[\mathrm{CVaR}\gamma_{\mathcal{F}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right] - \mathrm{CVaR}_\gamma\left[\mathbf{h}(x^*)\right]\Big]$ can be written as

$$
\begin{aligned}
&\mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right] \sum_{\omega\in\mathcal{F}^{\mathsf{c}}\cap\Omega_2(x^*)} q_\omega + \sum_{\omega\in\mathcal{V}} q_\omega h_\omega(x^*) \\
&\quad + \mathbb{Q}(\mathcal{F}^{\mathsf{c}})\Big[\Psi_{|\mathcal{F}^{\mathsf{c}}}\Big(x^*, \mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right]\Big) - \gamma_{\mathcal{F}}\Big]\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right] \\
&\quad - \Big[\Psi\Big(x^*, \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]\Big) - \gamma\Big]\mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right],
\end{aligned}
$$

which can be rearranged as

$$
\begin{aligned}
&\sum_{\omega\in\mathcal{V}} q_\omega\Big(h_\omega(x^*) - \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]\Big) \\
&\quad + \mathbb{Q}(\mathcal{F}^{\mathsf{c}})\Big[\Psi_{|\mathcal{F}^{\mathsf{c}}}\Big(x^*, \mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right]\Big) - \gamma_{\mathcal{F}}\Big] \\
&\quad \times \Big(\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right] - \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]\Big).
\end{aligned}
$$

Now, the results of the theorem will be implied by setting the above equation equal to zero to ensure $\mathrm{CVaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{c}}\right] = \mathrm{CVaR}_\gamma\left[\mathbf{h}(x^*)\right]$. □

4.4.3 Proof of effectiveness of union of scenarios

Proof of Theorem 4 Let $\mathcal{F}$ denote the union of all ineffective scenarios identified in Theorem 1. We first show $\mathcal{F}$ is ineffective. Then, Corollary 1 completes the proof.

Suppose $\lambda^* > 0$. Let us consider two further cases.

Case 1. $\sum_{\omega\in\Omega_2(x^*)} p^*_\omega > 0$. Consider the union of scenarios in parts (i), (ii), and (iv) in Theorem 1. Then, we can argue $\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{C}}\right] = \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]$ and the set $\mathcal{Z}^{\mathrm{A}}_\gamma(x^*;\mathcal{F})$ coincides with the restriction of the set $\mathcal{Z}_\gamma(x^*)$ to $\mathcal{F}^{\mathsf{C}}$, with a similar argument as that in the proof of Proposition 7. Therefore, we conclude $\mathcal{F}$ is ineffective by Lemma 1.

Case 2. $\sum_{\omega\in\Omega_2(x^*)} p^*_\omega = 0$. Consider the union of scenarios in parts (i), (iii), and (iv) in Theorem 1. Note that $\sum_{\omega\in\Omega_1(x^*)\cup\Omega_2(x^*)} q_\omega = \gamma$ by (7); so, $\gamma_{\mathcal{F}} = 0$. Thus, we have $\mathrm{VaR}_{\gamma_{\mathcal{F}}}\left[\mathbf{h}(x^*)|\mathcal{F}^{\mathsf{C}}\right] = -\infty$ following our convention. As a result, for any $\mathbf{z} \in \mathcal{Z}^{\mathrm{A}}_\gamma(x^*;\mathcal{F})$, we have $z_\omega = q_\omega$ for all $\omega \in \mathcal{F}^{\mathsf{C}}$, including $\omega \in \left[\mathbf{h}(x^*) > \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]\right]$. Now, it is easy to verify that the set $\mathcal{Z}^{\mathrm{A}}_\gamma(x^*;\mathcal{F})$ coincides with the restriction of the set $\mathcal{Z}_\gamma(x^*)$ to $\mathcal{F}^{\mathsf{C}}$; therefore, $\mathcal{F}$ is ineffective by Lemma 1.

Now, suppose $\lambda^* = 0$ and let $\mathcal{F} = \Omega_1(x^*)$. The rest of the proof follows a similar argument as that for Case 1. □

4.4.4 Proof of alternative optima

The proof of Theorem 5 is based on Lemmas A-3 and A-4.

Proof of Theorem 5 Suppose that ω' is effective and $p^*_{\omega'} = 0$. We first claim that there is a scenario ω'' such that $h_{\omega'}(x^*) = h_{\omega''}(x^*)$ and $p^*_{\omega''} > 0$. Then, we show that there is $\epsilon > 0$ such that solution $\mathbf{p}^{**}$, with $p^{**}_{\omega'} = \epsilon$, $p^{**}_{\omega''} = p^*_{\omega''} - \epsilon$, and $p^{**}_\omega = p^*_\omega$, $\omega \notin \{\omega', \omega''\}$, is another optimal solution to the (primal) worst-case expected problem at x^*.

By Theorems 1 and 2 and Proposition 4, when $\lambda^* > 0$, either $\omega' \in \Omega_2(x^*)$ with $q_{\omega'} > 0$, or $\omega' \in \Omega_4(x^*)$ with $q_{\omega'} = 0$. First, suppose $\omega' \in \Omega_2(x^*)$ and $q_{\omega'} > 0$. By Theorem 1(iii), there is a scenario $\omega'' \in \Omega_2(x^*)$ with $0 < p^*_{\omega''} \le q_{\omega''}$, otherwise all scenarios $\omega \in \Omega_2(x^*)$, including ω', would be ineffective. Now, we choose $\epsilon \le \min\{q_{\omega'}, p^*_{\omega''}\}$ so that $\mathbf{p}^{**}$ is feasible to the worst-case expected problem at x^*. On the other hand, the change in the optimal value of the worst-case expected problem at x^* is zero for $\mathbf{p}^{**}$ because both scenarios ω' and ω'' belong to the same primal category $\Omega_2(x^*)$. Now, suppose $\omega' \in \Omega_4(x^*)$ and $q_{\omega'} = 0$. By Corollary 3, there is a scenario $\omega'' \in \Omega_4(x^*)$ with $p^*_{\omega''} > q_{\omega''} \ge 0$. Now, choosing $\epsilon \le p^*_{\omega''} - q_{\omega''}$ yields the desired result with a similar argument as above. When $\lambda^* = 0$, according to Remark 4, we must have $\omega' \in \Omega_4(x^*)$. By Corollary 3, there is scenario $\omega'' \in \Omega_4(x^*)$, with $p^*_{\omega''} > 0$. Now, we choose ϵ following similar steps as in the proof of Lemma A-3.

Now, suppose ω' is ineffective and $p^*_{\omega'} > 0$. We first claim that there is a scenario ω'' such that $h_{\omega'}(x^*) = h_{\omega''}(x^*)$. Then, we show that there is $\epsilon > 0$ such that either $\mathbf{p}^{**}$ or $\bar{\mathbf{p}}^{**}$ is optimal to the (primal) worst-case expected problem at x^*, where $\mathbf{p}^{**}$ is defined as $p^{**}_{\omega'} = p^*_{\omega'} + \epsilon$, $p^{**}_{\omega''} = p^*_{\omega''} - \epsilon$, and $p^{**}_{\omega} = p^*_{\omega}$, $\omega \notin \{\omega', \omega''\}$, and $\bar{\mathbf{p}}^{**}$ is defined as $\bar{p}^{**}_{\omega'} = p^*_{\omega'} - \epsilon$, $\bar{p}^{**}_{\omega''} = p^*_{\omega''} + \epsilon$, and $\bar{p}^{**}_{\omega} = p^*_{\omega}$, $\omega \notin \{\omega', \omega''\}$.

When $\lambda^* > 0$, we consider two cases based on Proposition 4.

Case 1. $q_{\omega'} = 0$. According to Proposition 6, we must have $\omega' \in \Omega_4(x^*)$. By Proposition 8, there is a scenario $\omega'' \in \Omega_4(x^*)$ with $p^*_{\omega''} \geq q_{\omega''}$. Now, we choose $\epsilon \leq p^*_{\omega''} - q_{\omega''}$ for $\mathbf{p}^{**}$ or $\epsilon \leq p^*_{\omega'}$ for $\bar{\mathbf{p}}^{**}$.

Case 2. $q_{\omega'} > 0$. According to Proposition 7, we must have $\omega' \in \Omega_2(x^*)$. By Proposition 10, there is a scenario $\omega'' \in \Omega_2(x^*)$ with $p^*_{\omega''} \leq q_{\omega''}$. Now, we choose $\epsilon \leq \min\{q_{\omega'} - p^*_{\omega'}, p^*_{\omega''}\}$ for $\mathbf{p}^{**}$ or $\epsilon \leq \min\{q_{\omega''} - p^*_{\omega''}, p^*_{\omega'}\}$ for $\bar{\mathbf{p}}^{**}$.

When $\lambda^* = 0$, according to Remark 4, we must have $\omega' \in \Omega_4(x^*)$. By Proposition 8, there is a scenario $\omega'' \in \Omega_4(x^*)$ with $\omega' \neq \omega''$. Now, we choose ϵ following similar steps as in the proof of Lemma A-4. □

5 Numerical illustration

In this section we present numerical results. First, we explain our experimental setup in Sect. 5.1. We study the efficacy of our proposed easy-to-check conditions in identifying the effectiveness of scenarios in Sect. 5.2. Then, we investigate how the effective and ineffective scenarios can be used to gain insight about the underlying uncertainties of the problems. Finally, we examine the sets of effective and ineffective scenarios at different levels of robustness in Sect. 5.4.

5.1 Experimental setup

To conduct numerical analysis, we used a number of problems from the literature, including APL1P, PGP2, and Water Allocation problem. APL1P [22] and PGP2 [18] are power capacity expansion models. In both problems, the first stage decides on what capacities must be installed on the generators. The second stage decides on the purchase of additional capacities to fulfill the unmet demands. APL1P and PGP2 have 1280 and 576 scenarios in total, respectively. The Water Allocation problem is a water resources allocation problem [47]. This problem has originally four stages, and its goal is to allocate Colorado River water among different users, while meeting uncertain water demand and not exceeding uncertain water supply over the next 41 years. We generated a smaller two-stage variant of this problem, with a 16-year planning horizon and 200 scenarios.

The level of robustness parameter, γ, is set between 0 and 1, and in increment of 0.05. Hence, each test problem is solved a total of 21 times. All problems were solved on a 64-bit Windows environment using C++\CPLEX 12.6 on a single core of a PC with an Intel 2.30 GHz processor and 4.00 GB of RAM. To solve the problems, we used the decomposition method discussed Sect. A.2 of the Appendix. Alternatively, a variant of the L-shaped method, capable of handling the CVaR function, can be used. We refer to [47] for a survey of different decomposition schemes to handle the CVaR function.

Table 1 Number of scenarios in each primal category, and number of ineffective, effective, and undetermined scenarios for APL1P, $n = 1280$

γ	# of scenarios $\Omega_1(x^*)$	$\Omega_2(x^*)$	$\Omega_3(x^*)$	$\Omega_4(x^*)$	# of scenarios Ineffective	Effective	Undetermined
0.00	0	3	1276	1	0	1280	0
0.05	74	2	1203	1	74	1205	1
0.10	136	1	1142	1	136	1144	0
0.15	189	1	1089	1	189	1091	0
0.20	226	1	1052	1	226	1054	0
0.25	267	1	1011	1	267	1013	0
0.30	312	4	963	1	312	966	2
0.35	353	4	922	1	353	924	3
0.40	384	3	892	1	384	893	3
0.45	431	6	842	1	431	843	6
0.50	471	6	802	1	471	803	6
0.55	510	7	762	1	510	763	7
0.60	561	7	711	1	561	712	7
0.65	600	6	673	1	600	674	6
0.70	671	3	605	1	671	609	0
0.75	728	11	540	1	728	541	11
0.80	804	10	465	1	804	466	10
0.85	899	9	371	1	899	379	2
0.90	988	12	279	1	988	280	12
0.95	1076	12	191	1	1076	192	12
1.00	1279	1	–	–	1279	1	0

5.2 Performance of the easy-to-check conditions

We tested the performance of easy-to-check conditions in identifying the effectiveness of scenarios on APL1P, PGP2, and the Water Allocation problem. Table 1 presents the number of scenarios in each primal category, as well as the number of ineffective and effective scenarios identified by the easy-to-check conditions for APL1P. The results for PGP2 and the Water Allocation problem are presented in the Online Supplement. The trends are similar for these two problems except that there are no undetermined scenarios for the Water Allocation problem.

Several trends are clear from results. As stated before, the easy-to-check conditions might not be able to identify the effectiveness of all scenarios. However, for APL1P, the undetermined scenarios form a small portion of total number of scenarios. Moreover, as the level of robustness increases, the total number of scenarios in $\Omega_1(x^*)$ and $\Omega_2(x^*)$ decreases, and the total number of scenarios in $\Omega_3(x^*)$ and $\Omega_4(x^*)$ increases. This is because as the level of the robustness increases, the problem is concentrated more on the tail of the cost distribution (i.e., $\Omega_3(x^*)$ and $\Omega_4(x^*)$), and the number of scenarios in the tail decreases. On the other hand, roughly speaking, effective scenarios are those in the tail of the distribution, and ineffective scenarios are those that are not in the tail of the distribution. Consequently, as shown in Table 1, as the level of robustness increases, the number of ineffective scenarios decreases and the number of effective scenarios increases. We further investigate this behavior in Sect. 5.4.

5.3 Managerial insights

In Sect. 5.2, we reported the number of effective scenarios at each level of robustness. However, we have not commented yet on the specifics of effective scenarios. In this section we provide insight on the effective scenarios for APL1P, PGP2, and the Water Allocation problem in order to help decision makers recognize most important scenarios on the optimal value and gain more information about the underlying uncertainties of their problems. In particular, we are interested in the insight obtained from the effective scenarios identified by the easy-to-check conditions. To this end, we considered each scenario in the problem. Then, we reported the percentage of cases that scenario is identified as effective by the easy-to-check conditions out of the 21 DRSPs with varying γ. Ideally, we like to report these percentages in the space of all random variables that represent the uncertainty in the problem. However, the uncertainty is expressed by multi-dimensional random variables, and it is not possible to visualize the results. On the other hand, the uncertainty in these problems can be categorized into supply and/or demand stochasticity. To visualize the results, for each scenario, we reported the percentage of cases a scenario is identified effective versus the average supply and/or demand corresponding to that scenario.

The results for APL1P are depicted in Fig. 1. Generally speaking, at a fixed level of demand, the percentage of cases a scenario is effective increases as the generator availability decreases (Fig. 1a). In addition, generally speaking, at a fixed level of the generator availability, the percentage of cases a scenario is effective increases as the demand increases. Fig. 1b shows that a scenario with a higher demand and a lower supply is more likely to be an effective scenario. On the other hand, a scenario with a lower demand and a higher supply is less likely to be effective.

The results for PGP2 are depicted in Fig. 2. Generally speaking, at a fixed level for each demand, the percentage of cases a scenario is effective increases as the average of the other two demands increases. In addition, generally speaking, at a fixed level of average demand 2 and 3, the percentage of cases a scenario is effective increases as demand 1 increases (Fig. 2a). However, as shown in Fig. 2b, c, the percentage of cases a scenario is effective remains constant as demand 2 and 3 increases, respectively. These imply demand 1 is most critical to the optimal value.

A similar trend occurs for the Water Allocation problem, as shown in Fig. 3. Generally speaking, at a fixed level of the water supply, the percentage of cases a scenario is effective increases as the demand increases. In addition, a scenario with a higher demand is more likely to be an effective scenario (Fig. 3). This is because a higher demand scenario is more costly to the system, and therefore the problem must consider such a scenario. This is even true when there is normal supply of Colorado River

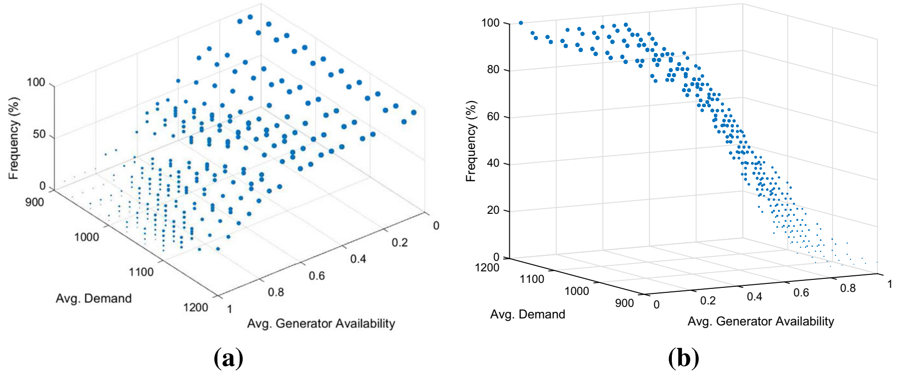

Fig. 1 The percentage of cases a scenario is effective for APL1P

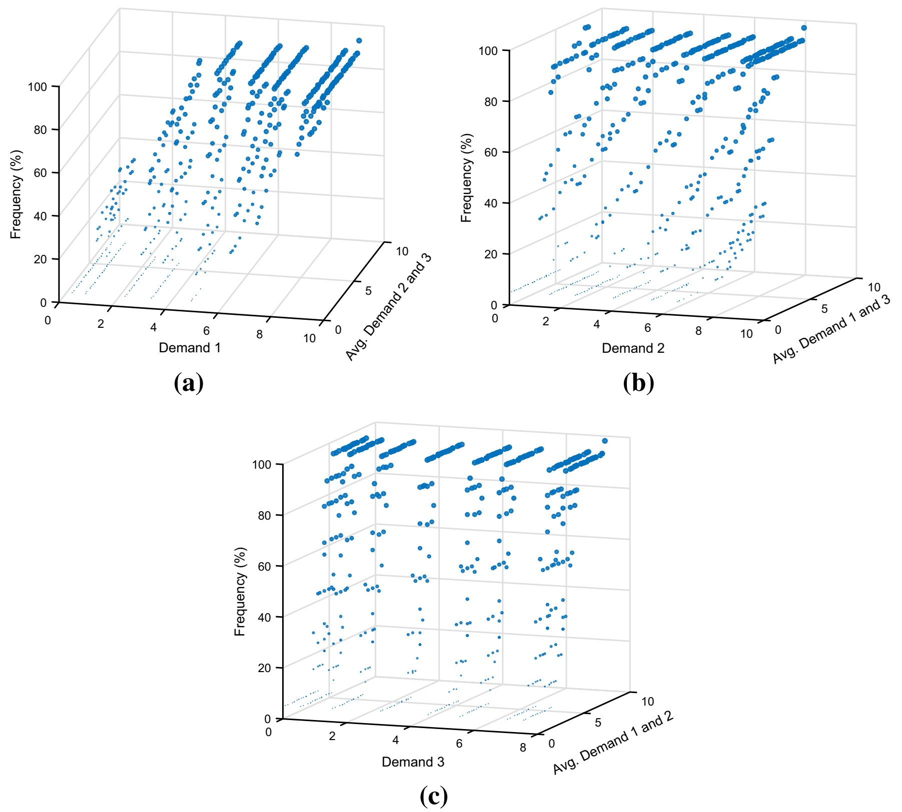

Fig. 2 The percentage of cases a scenario is effective for PGP2

water (higher avg. supply in Fig. 3) because the system is quite constrained even under normal water supply conditions.

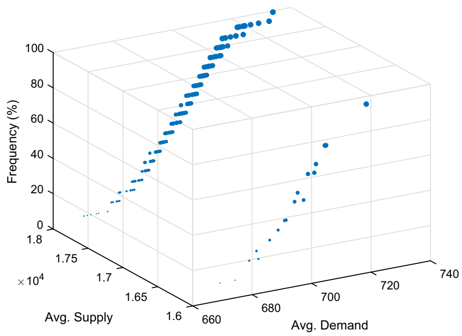

Fig. 3 The percentage of cases a scenario is effective for Water Allocation problem across different levels of robustness. The two distinct portions in the graph correspond to a drought in the Colorado River (low avg. supply) and no drought in the Colorado River (high avg. supply)

5.4 Monotonicity and nestedness of effective/ineffective scenarios

In this section, we analyze the following three conjectures on effectiveness of scenarios:

(C1) *Monotonicity in the number of effective/ineffective scenarios* The number of ineffective scenarios is non-decreasing and the number of effective scenarios is non-increasing as the level of robustness increases (see, for instance, Table 1).

(C2) *Nestedness of the sets of effective/ineffective scenarios* If a scenario is ineffective at some level of robustness, it must be ineffective at all larger levels of robustness, and if a scenario is effective at some level of robustness, it must be effective at all smaller levels of robustness.

(C3) *Monotonicity in the effectiveness of scenarios* If a scenario switches its effectiveness from ineffective to effective (or from effective to ineffective) at some level of robustness, it will keep its effectiveness the same and will not change anymore as the level of robustness increases.

Note that (C1) and (C3) are implied by (C2). However, neither of them imply (C2); see, for instance, Table 2a.

For APL1P, PGP2, and the Water Allocation problem, not only is the number of effective/ineffective scenarios monotone, but also the sets of effective/ineffective scenarios are nested and the effectiveness of scenarios is monotone. It might be thought that such trends happen in all problems. In the following, we present some counterexamples to demonstrate the breakage of the monotonicity and nestedness.

Example 3 (*Breakage of (C1), and hence breakage of (C2), while (C3) does hold.*) Let us consider an inventory problem with the following specifics: $n = 6$, $\mathbf{q} = (0, 0.2, 0.25, 0.2, 0.35, 0)$, and $\mathbf{d} = (1, 2, 3, 4, 5, 6)$. Also, ordering cost c is 4, backorder cost b is 5, and holding cost m is 5. For this problem, the cost function is written as $h_\omega(x) = cx + b(d_\omega - x)_+ + m(x - d_\omega)_+$. Table 2a shows that the number of effective scenarios (equivalently, ineffective scenarios) is not monotone since $\gamma = 0.95$ breaks down the monotonicity. Moreover, the nestedness of the sets of effective/ineffective scenarios breaks down, because, for instance, the first scenario is not effective at the levels of robustness smaller than 0.95. However, the monotonicity in the effectiveness of scenarios holds. □

Table 2 Counterexamples

(a) Breakage of (C1) and (C2), while (C3) holds for the inventory problem.

γ	# of scen.s I	# of scen.s E	scen. # 1	2	3	4	5	6
0.00	2	4	I	E	E	E	E	I
0.05	1	5	I	E	E	E	E	E
0.10	1	5	I	E	E	E	E	E
0.15	1	5	I	E	E	E	E	E
0.20	1	5	I	E	E	E	E	E
0.25	1	5	I	E	E	E	E	E
0.30	1	5	I	E	E	E	E	E
0.35	2	4	I	E	I	E	E	E
0.40	2	4	I	E	I	E	E	E
0.45	2	4	I	E	I	E	E	E
0.50	2	4	I	E	I	E	E	E
0.55	3	3	I	E	I	I	E	E
0.60	3	3	I	E	I	I	E	E
0.65	4	2	I	I	I	I	E	E
0.70	4	2	I	I	I	I	E	E
0.75	4	2	I	I	I	I	E	E
0.80	4	2	I	I	I	I	E	E
0.85	4	2	I	I	I	I	E	E
0.90	5	1	I	I	I	I	I	E
0.95	4	2	E	I	I	I	I	E
1.00	4	2	E	I	I	I	I	E

I: Ineffective; E: Effective.

(b) Breakage of (C3) and (C2), while (C1) holds for the variant of APL1P.

γ	# of scen.s I	# of scen.s E	scen. # 5
0.00	0	16	E
0.05	1	15	E
0.10	2	14	I
0.15	2	14	E
0.20	3	13	I
0.25	3	13	I
0.30	4	12	I
0.35	5	11	I
0.40	5	11	I
0.45	6	10	I
0.50	8	8	I
0.55	8	8	I
0.60	9	7	I
0.65	9	7	I
0.70	9	7	I
0.75	10	6	I
0.80	10	6	I
0.85	11	5	I
0.90	12	4	I
0.95	13	3	I
1.00	15	1	I

I: Ineffective; E: Effective.

Example 4 (Breakage of (C3), and hence breakage of (C2), while (C1) does hold.) For this counterexample, we generated a variant of APL1P, but with a smaller number of scenarios. In this variant, the availability of the second generator as well as the first and third demands are set to their expected values. Thus, we have two random variables, one representing the uncertainty in the availability of the first generator, and the other representing the uncertainty in the second demand. Thus, in total, we have 16 scenarios. Table 2b shows that the effectiveness of scenario # 5, with 900 units of demand and 90% generator availability, is not monotone. However, the number of effective/ineffective scenarios is monotone. □

The two examples above illustrate different mechanisms monotonicity can break. Example 3 contains a high-cost scenario with zero nominal probability ($d_1 = 1$). For high-enough γ, this scenario is "popped" (that is, given a positive probability p_1^*), and it eventually becomes an effective scenario. Observe that total variation distance can pop scenarios, and only the highest-cost scenarios can be popped [4]. Thus, this first type of monotonicity breakage may be observed if there are high-cost scenarios with zero nominal probabilities.

In contrast, Example 4 contains no scenarios with zero nominal probabilities. Here, there are several scenarios of similar cost (e.g., 1000 units of demand with 100% generator availability and 900 units of demand with 90% generator availability). When x^* changes at different γ values, these similar scenarios exchange primal categories (e.g., from $\Omega_2(x^*)$ to $\Omega_1(x^*)$ and back) because the order of their costs change at different x^*. This exchange of primal categories causes an effective scenario to become ineffective and return back to being effective. This is exactly what happens to scenario # 5 in Table 2b. Observe that for a lower resolution of 0.1-steps in the γ values, we do not observe breakage of monotonicity. In our numerical experiments with other problems, we observed similar breakage of monotonicity at higher resolutions of 0.005-steps in the γ values, depending on the total number of scenarios and the nominal probabilities of similar-cost scenarios. While obtaining general conditions under which one can guarantee monotonicity may be difficult, in [33] we show monotonicity for newsvendor problems under appropriate assumptions.

6 Discussion and future research

In this paper, we investigated distributionally robust convex stochastic optimization problems with finite support to hedge against the distributional ambiguity. We analyzed the critical scenarios we are concerned about in this setting, and we illustrated that one cannot necessarily identify critical scenarios by examining the worst-case probabilities. This led us to define the notions of effective and ineffective scenarios for a general class of distributionally robust convex optimization problems. In particular, we studied problems where the distributional ambiguity is modeled by the total variation distance, motivated by the conceptual, computational, and interpretative benefits of the resulting model. By exploiting the specific structures of the ambiguity set formed via the total variation distance, we proposed easy-to-check conditions to identify effectiveness of scenarios for DRSP-V.

A summary of conclusions from this study are as follows:

- The notion of effectiveness of scenarios is applicable to a general class of problems. For DRSP, the effectiveness of a scenario can be identified by re-solving the corresponding assessment problem and using the general definition.
- For DRSP-V, we proposed easy-to-check conditions. Theoretically, easy-to-check conditions identify the effectiveness of all scenarios, except for some scenarios in certain primal categories. Numerical results showed that these conditions work quite well and the number of undetermined scenarios is small relative to the total number of scenarios for the tested problems.
- We investigated three conjectures on the effectiveness of scenarios as the level of robustness increases: monotonicity in the number of effective/ineffective scenarios, nestedness of the corresponding sets, and monotonicity in the effectiveness of scenarios. We illustrated through counterexamples that these conjectures do not hold in general. However, we observed that for APL1P, PGP2, and the Water Allocation problem they do hold.
- Effective scenarios can provide managerial insight into the underlying uncertainties of the problems and encourage decision makers to collect more accurate information surrounding certain scenarios.

Future work includes investigating the effectiveness of scenarios for other DRSPs formed by e.g., other ϕ-divergences, or DRSPs that result from the dualization of other commonly used risk measures. One example is minimization of spectral risk measures, which can be written as convex combinations of CVaRs [1]. For this type of problems, as well as general DRSPs, a large portion of the scenarios may be effective.

So, one might further explore the amount of change in the optimal value to differentiate between effective scenarios. Another topic for future research is the study of conditions under which monotonicity properties of effective scenarios hold. Moreover, the notion of effective scenarios opens avenues to study the value of data from this perspective. Other future work includes studying distributionally robust multistage stochastic programs, in particular the notion of effectiveness of scenarios in that setting.

It would also be interesting to further explore the notion of effective scenarios in practical applications where the uncertainty appears in forms other than demand or supply as in the examples discussed in this paper. For example, in mine planning geological uncertainty plays an important role, as the actual concentrations of the minerals in the ground are not known until the material is extracted. Despite the vast literature on geostatistics models to represent that uncertainty (see, e.g., [36]), the corresponding probability distributions are ultimately approximations resulting from sampled data obtained from drill holes and as such may not represent reality accurately. In such cases, a distributional robust model can be useful, and the analysis of effective scenarios could potentially guide the extraction process in terms of the level of robustness set by the decision maker.

Acknowledgements This work has been partially supported by the National Science Foundation through Grants CMMI-1345626 and CMMI-1563504 of the second author and CONICYT PIA Anillo ACT1407 (Chile) of the third author.

Appendix

In Sect. A.1, we provide the results and proofs that were relegated to the Appendix. Then, in Sect. A.2, we propose a decomposition method to solve (DRSP-V).

A-1 Proofs

In this section, we first provide the proof of Proposition 4. Then, in Sect. A.1.2, we present Proposition A-1. In Sect. A.1.3, we state Lemma A-1, used in the proof of Proposition 5. In Sect. A.1.4, we provide Lemma A-2, used in the proof of Proposition 7. Finally, in Sect. A.1.5, we preset Lemmas A-3 and A-4, which are used in the proof of Theorem 5.

A-1.1 Proof of Proposition 4

Proof of Proposition 4 For a fixed $x \in \mathbb{X}$, the worst-case expected problem can be written as follows

$$\max_{\mathbf{p},\mathbf{e}} \sum_{\omega\in\Omega} p_\omega h_\omega(x) \tag{A-1a}$$

$$\text{s.t.} \quad p_\omega - e_\omega \le q_\omega, \ \forall\omega, \ : [u'_\omega] \tag{A-1b}$$

$$- p_\omega - e_\omega \le -q_\omega, \ \forall\omega, \ : [u''_\omega] \tag{A-1c}$$

$$\frac{1}{2}\sum_{\omega\in\Omega} e_\omega \le \gamma, \; : [\lambda(x)] \tag{A-1d}$$

$$\sum_{\omega\in\Omega} p_\omega = 1, \; : [\mu(x)] \tag{A-1e}$$

$$p_\omega, e_\omega \ge 0, \; \forall\omega. \tag{A-1f}$$

where u'_ω, u''_ω, $\lambda(x)$, and $\mu(x)$ are the associated dual variables for the constraints. Define now

$$\lambda(x) = \sup_{\omega\in\Omega} h_\omega(x) - \mathrm{VaR}_\gamma\left[\mathbf{h}(x)\right], \; \mu(x) = \frac{1}{2}\left(\sup_{\omega\in\Omega} h_\omega(x) + \mathrm{VaR}_\gamma\left[\mathbf{h}(x)\right]\right).$$

We consider two cases (i) $\mathrm{VaR}_\gamma[\mathbf{h}(x)] < \sup_{\omega\in\Omega} h_\omega(x)$ and (ii) $\mathrm{VaR}_\gamma[\mathbf{h}(x)] = \sup_{\omega\in\Omega} h_\omega(x)$. In case (i), we have $\lambda(x) > 0$. Then,

$$\begin{cases} p_\omega = 0, \; u'_\omega = 0, \; u''_\omega = \frac{\lambda(x)}{2}, \; e_\omega = q_\omega, & \omega \in \Omega_1(x), \\ p_\omega \le q_\omega, \; u'_\omega = 0, \; u''_\omega = \frac{\lambda(x)}{2}, \; e_\omega = q_\omega - p_\omega, & \omega \in \Omega_2(x), \\ p_\omega = q_\omega, \; u'_\omega = \frac{h_\omega(x)-\mu(x)+\frac{\lambda(x)}{2}}{2}, \; u''_\omega = \frac{\lambda(x)}{2} - \frac{h_\omega(x)-\mu(x)+\frac{\lambda(x)}{2}}{2}, \; e_\omega = 0, & \omega \in \Omega_3(x), \\ p_\omega \ge q_\omega, \; u'_\omega = \frac{\lambda(x)}{2}, \; u''_\omega = 0, \; e_\omega = p_\omega - q_\omega, & \omega \in \Omega_4(x), \end{cases}$$

together with $\sum_{\omega\in\Omega} p_\omega = 1$, is primal-dual feasible. For the complementary slackness condition to hold, (A-1d) must hold with equality at an optimal $(\mathbf{p}, \mathbf{e}, \mu(x), \lambda(x), \mathbf{u}', \mathbf{u}'')$ when $\lambda(x) > 0$. Indeed, we have

$$\begin{aligned} \sum_{\omega\in\Omega} e_\omega &= \sum_{\omega\in\Omega_1(x)} q_\omega + \sum_{\omega\in\Omega_2(x)} (q_\omega - p_\omega) + \sum_{\omega\in\Omega_4(x)} (p_\omega - q_\omega) \\ &= \sum_{\omega\in\Omega_1(x)\cup\Omega_2(x)} q_\omega - \sum_{\omega\in\Omega_4(x)} q_\omega - \sum_{\omega\in\Omega_2(x)} p_\omega + \sum_{\omega\in\Omega_4(x)} p_\omega \\ &= 1 - 2\sum_{\omega\in\Omega_4(x)} q_\omega - \sum_{\omega\in\Omega_3(x)} q_\omega - \sum_{\omega\in\Omega_2(x)} p_\omega + \sum_{\omega\in\Omega_4(x)} p_\omega \\ &\qquad \left[\text{because} \sum_{\omega\in\Omega} q_\omega = 1\right] \\ &= 1 - 2\sum_{\omega\in\Omega_4(x)} q_\omega - \sum_{\omega\in\Omega_3(x)} p_\omega - \sum_{\omega\in\Omega_2(x)} p_\omega + \sum_{\omega\in\Omega_4(x)} p_\omega \\ &\qquad [\text{because } p_\omega = q_\omega \text{ on } \Omega_3(x)] \\ &= 2\sum_{\omega\in\Omega_4(x)} (p_\omega - q_\omega) \left[\text{because} \sum_{\omega\in\Omega} p_\omega = 1\right] \\ &= 2\gamma, \end{aligned} \tag{A-2}$$

where the last equality follows from (8). Now, by substituting (8) in (A-2) we obtain (7).

On the other hand, in case (ii), we have $\lambda(x) = 0$, $\Omega_3(x) = \emptyset$ and $\Omega_2(x) = \Omega_4(x)$. Then, all u'_ωs and u''_ωs are zero, and we have $\mu(x) = \sup_{\omega\in\Omega} h_\omega(x)$. In this case, if $h_\omega(x) = \sup_{\omega\in\Omega} h_\omega(x)$, we have $e_\omega = |p_\omega - q_\omega|$ and $p_\omega \geq 0$ is chosen so that it satisfies $\sum_{\omega\in\Omega} p_\omega = 1$ and $\frac{1}{2}\sum_{\omega\in\Omega} e_\omega \leq \gamma$. Otherwise, $p_\omega = 0$. □

A-1.2 Proposition A-1

Proposition A-1 shows the equivalence of the ambiguity sets induced by the right- and left-sided variation distances.

Proposition A-1 *The ambiguity sets induced by the right- and left-sided variation distances are the same.*

Proof of Proposition A-1 Let $\mathcal{P}_\gamma^{\mathrm{V^R}}$ and $\mathcal{P}_\gamma^{\mathrm{V^L}}$ denote the ambiguity set induced by the right- and left-sided variation distances, respectively. Recall $\mathrm{V^R}(\mathbf{p}, \mathbf{q}) = \frac{1}{2}\sum_{\omega\in\Omega}(p_\omega - q_\omega)_+$ and $\mathrm{V^L}(\mathbf{p}, \mathbf{q}) = \frac{1}{2}\sum_{\omega\in\Omega}(q_\omega - p_\omega)_+$. Now, if we show $\mathrm{V^R}(\mathbf{p}, \mathbf{q})$ and $\mathrm{V^L}(\mathbf{p}, \mathbf{q})$ are equal, it would show the equivalence of $\mathcal{P}_\gamma^{\mathrm{V^R}}$ and $\mathcal{P}_\gamma^{\mathrm{V^L}}$. Let $M := \{\omega \in \Omega \;:\; p_\omega \geq q_\omega\}$. So, we can write $\mathrm{V^R}(\mathbf{p}, \mathbf{q}) = \frac{1}{2}\sum_{\omega\in M}(p_\omega - q_\omega) = \frac{1}{2}\sum_{\omega\in M^c}(q_\omega - p_\omega) = \mathrm{V^L}(\mathbf{p}, \mathbf{q})$ by using the facts that $\sum_{\omega\in\Omega} p_\omega = 1$ and $\sum_{\omega\in\Omega} q_\omega = 1$. □

A-1.3 Lemma A-1

In this section, we state Lemma A-1, which is used in the proof of Proposition 5. The proof of this lemma is presented in the Online Supplement.

Lemma A-1 *Consider a fixed $x \in \mathbb{X}$ and level $\beta \in [0, 1)$. Then,* $\mathrm{CVaR}_\beta[\mathbf{h}(x)] = \max_{\mathbf{z}\in\mathfrak{A}_\beta} \sum_{\omega\in\Omega} z_\omega h_\omega(x)$, *where*

$$\mathfrak{A}_\beta := \left\{ \mathbf{z} \;:\; \sum_{\omega\in\Omega} z_\omega = 1,\; 0 \leq z_\omega \leq \frac{q_\omega}{1-\beta},\; \forall\omega\in\Omega \right\}. \tag{A-3}$$

Moreover, the set of all optimal solutions $\mathbf{z}$ *is given by*

$$\mathfrak{D}_\beta(x) := \left\{ \mathbf{z} : \begin{array}{ll} z_\omega = 0, & \textit{if } \omega \in \big[\mathbf{h}(x) < \mathrm{VaR}_\beta[\mathbf{h}(x)]\big], \\ z_\omega = \frac{q_\omega}{1-\beta}, & \textit{if } \omega \in \big[\mathbf{h}(x) > \mathrm{VaR}_\beta[\mathbf{h}(x)]\big], \\ z_\omega \in [0, \frac{q_\omega}{1-\beta}], & \textit{if } \omega \in \big[\mathbf{h}(x) = \mathrm{VaR}_\beta[\mathbf{h}(x)]\big], \\ \sum_{\omega\in[\mathbf{h}(x)=\mathrm{VaR}_\beta[\mathbf{h}(x)]]} z_\omega = \frac{1}{1-\beta}\Big[\Psi\Big(x, \mathrm{VaR}_\beta[\mathbf{h}(x)]\Big) - \beta\Big]. & \end{array} \right. \tag{A-4}$$

A-1.4 Lemma A-2

In this section, we present Lemma A-2, which is used in the proof of Proposition 7.

Lemma A-2 *Consider a scenario ω' with $q_{\omega'} > 0$. If $h_{\omega'}(x^*) > \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]$, then scenario ω' is effective for* (DRSP-V).

Proof of Lemma A-2 Suppose $\bar{x}$ solves the assessment problem of $\mathcal{F} = \{\omega'\}$ with $q_{\omega'} > 0$. By (3), we have $f_\gamma^{\mathrm{A}}(\bar{x}; \mathcal{F}) \le f_\gamma^{\mathrm{A}}(x^*; \mathcal{F}) \le f_\gamma(x^*)$. Consequently, $f_\gamma(x^*) - f_\gamma^{\mathrm{A}}(x^*; \mathcal{F})$ gives a lower bound on $f_\gamma(x^*) - f_\gamma^{\mathrm{A}}(\bar{x}; \mathcal{F})$. If this lower bound is positive, then scenario ω' is effective. To check this lower bound, let us consider the objective function of the dual problem of (A-1a) (presented in the proof of Proposition 3 in the Online Supplement) and (13), both evaluated at x^*. Note that (μ^*, λ^*) belongs to the feasible region of (13) because $\mu^* + \frac{\lambda^*}{2} = \sup_{\omega\in\Omega} h_\omega(x^*) \ge \sup_{\omega\in\mathcal{F}^{\mathrm{c}}} h_\omega(x^*)$ and $\lambda^* \ge 0$. Thus, we have

$$\begin{aligned} f_\gamma(x^*) - f_\gamma^{\mathrm{A}}(x^*; \mathcal{F}) \ge{}& \mu^* - \frac{\lambda^*}{2}(1-2\gamma) + \sum_{\omega\in\Omega} q_\omega \left(h_\omega(x^*) - \mu^* + \frac{\lambda^*}{2}\right)_+ \\ & - \left[\mu^* - \frac{\lambda^*}{2}(1-2\gamma) + \sum_{\omega\in\mathcal{F}^{\mathrm{c}}} q_\omega \left(h_\omega(x^*) - \mu^* + \frac{\lambda^*}{2}\right)_+\right] \\ ={}& q_{\omega'}\left(h_{\omega'}(x^*) - \mu^* + \frac{\lambda^*}{2}\right)_+ \\ ={}& q_{\omega'}\Big(h_{\omega'}(x^*) - \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]\Big)_+, \end{aligned}$$

where the last equality follows from the fact $\mu^* - \frac{\lambda^*}{2} = \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]$ by (5). Now, dividing by $q_{\omega'}$ shows that the lower bound is positive if $h_{\omega'}(x^*) > \mathrm{VaR}_\gamma\left[\mathbf{h}(x^*)\right]$. □

A-1.5 Lemmas used in the Proof of Theorem 5

For the following two lemmas, we consider a fixed $x \in X$. Also, we suppose $\mathbf{p}$ solves the worst-case expected problem in (DRSP-V) at x, and (μ, λ) are the optimal dual variables at x, given by (5).

Lemma A-3 *For a fixed $x \in \mathbb{X}$, suppose $\lambda = 0$ and there is a scenario $\omega' \in \Omega_4(x)$ with $p_{\omega'} = 0$. Then, $\Omega_4(x)$ is not a singleton set, and for a scenario $\omega'' \in \Omega_4(x)$ with $p_{\omega''} > 0$, there is always $\epsilon > 0$ such that $\mathbf{p}'$, with $p'_{\omega'} = p_{\omega'} + \epsilon$, $p'_{\omega''} = p_{\omega''} - \epsilon$, and $p'_\omega = p_\omega$, $\omega \notin \{\omega', \omega''\}$, is feasible to the worst-case expected problem at x.*

Proof of Lemma A-3 Note that $\Omega_4(x)$ is not a singleton set by Corollary 3. To show the existence of another feasible solution $\mathbf{p}'$, we first examine the left-hand side of the distance constraint $\frac{1}{2}\sum_{\omega\in\Omega} |p_\omega - q_\omega| \le \gamma$ in (4). Then, we find $\epsilon > 0$ such that the change in the left-hand side of the distance constraint in (4) is smaller than or equal to zero; hence, $\mathbf{p}'$ is feasible to the worst-case expected problem at x. We consider two cases.

Case 1. $q_{\omega'} = 0$. Let us consider two further cases.

Case 1.1. There is a scenario $\omega'' \in \Omega_4(x)$ with $p_{\omega''} > q_{\omega''}$. By choosing $\epsilon \le p_{\omega''} - q_{\omega''}$ for $\mathbf{p}'$, the change in the left-hand side of the distance constraint is zero.

Case 1.2. For all $\omega'' \in \Omega_4(x)$, we have $p_{\omega''} \leq q_{\omega''}$. Then, set $\Omega_4^{\mathsf{c}}(x)$ is either empty or there is no scenario $\omega \in \Omega_4^{\mathsf{c}}(x)$ with $q_\omega > 0$. Otherwise, $\sum_{\omega\in\Omega} p_\omega \leq \sum_{\omega\in\Omega_4(x)} q_\omega < 1$. Moreover, we must have $p_{\omega''} = q_{\omega''}$ for all $\omega'' \in \Omega_4(x)$. Otherwise, $\sum_{\omega\in\Omega} p_\omega < \sum_{\omega\in\Omega_4(x)} q_\omega = 1$. Consequently, the left-hand side of the distance constraint for $\mathbf{p}$ is zero. Because for $\mathbf{p}'$ the change in the left-hand side of the distance constraint is ϵ, we choose $\epsilon \leq \gamma$.

Case 2. $q_{\omega'} > 0$. Then, there is a scenario $\omega'' \in \Omega_4(x)$ with $p_{\omega''} > q_{\omega''}$. Otherwise, $\sum_{\omega\in\Omega} p_\omega \leq \sum_{\substack{\omega\in\Omega_4(x)\\ \omega\neq\omega'}} q_\omega < 1$. By choosing $\epsilon \leq \min\{q_{\omega'}, p_{\omega''} - q_{\omega''}\}$ for $\mathbf{p}'$, the change in the left-hand side of the distance constraint is $-\epsilon$. □

Lemma A-4 *For a fixed $x \in \mathbb{X}$, suppose $\lambda = 0$ and there is a scenario $\omega' \in \Omega_4(x)$ with $p_{\omega'} > 0$. Moreover, suppose there is a scenario $\omega' \neq \omega'' \in \Omega_4(x)$. Then, there is always $\epsilon > 0$ such that either $\mathbf{p}'$ or $\mathbf{p}''$ is feasible to the worst-case expected problem at x, where $\mathbf{p}'$ is defined as $p'_{\omega'} = p_{\omega'} + \epsilon$, $p'_{\omega''} = p_{\omega''} - \epsilon$, and $p'_\omega = p_\omega$, $\omega \notin \{\omega', \omega''\}$, and $\mathbf{p}''$ is defined as $p''_{\omega'} = p_{\omega'} - \epsilon$, $p''_{\omega''} = p_{\omega''} + \epsilon$, and $p''_\omega = p_\omega$, $\omega \notin \{\omega', \omega''\}$.*

Proof of Lemma A-4 Similar to Lemma A-3, we examine the left-hand side of the distance constraint and find $\epsilon > 0$ such that $\mathbf{p}'$ or $\mathbf{p}''$ is feasible to the worst-case expected problem at x. We consider two cases.

Case 1. $p_{\omega'} \geq q_{\omega'}$. Let us consider two further cases.

Case 1.1. There is a scenario $\omega'' \in \Omega_4(x)$ with $p_{\omega''} \geq q_{\omega''}$. By choosing $\epsilon \leq p_{\omega''} - q_{\omega''}$ for $\mathbf{p}'$ or $\epsilon \leq p_{\omega'} - q_{\omega'}$ for $\mathbf{p}''$, the change in the left-hand side of the distance constraint is zero.

Case 1.2. For all $\omega' \neq \omega'' \in \Omega_4(x)$, we have $p_{\omega''} < q_{\omega''}$. By choosing $\epsilon \leq \min\{p_{\omega'} - q_{\omega'}, q_{\omega''} - p_{\omega''}\}$ for $\mathbf{p}''$, the change in the left-hand side of the distance constraint is $-\epsilon$.

Case 2. $p_{\omega'} < q_{\omega'}$. Then, there is a scenario $\omega'' \in \Omega_4(x)$ with $p_{\omega''} > q_{\omega''}$. Otherwise, $\sum_{\omega\in\Omega} p_\omega < \sum_{\omega\in\Omega_4(x)} q_\omega \leq 1$. By choosing $\epsilon \leq \min\{p_{\omega''} - q_{\omega''}, q_{\omega'} - p_{\omega'}\}$ for $\mathbf{p}'$, the change in the left-hand side of the distance constraint is $-\epsilon$. □

A-2 Primal decomposition

As the number of scenarios increases, solving (DRSP-V) becomes computationally expensive. Decomposition-based methods could significantly reduce the solution time and allow larger problems to be solved efficiently. In the following, we propose a cutting-plane approach, referred to as *Primal Decomposition*, to solve (DRSP-V) and obtain an optimal solution x^* and an optimal worst-case probability distribution $\mathbf{p}^*$.

Let us consider, $\mathcal{P}_\gamma$, the ambiguity set induced by the total variation distance. Let $\{\mathbf{p}^k\}_{k\in K}$ denote the set of extreme points of the polytope $\mathcal{P}_\gamma$. Then, (DRSP-V) can be written equivalently as

$$\begin{aligned} &\min_{x\in\mathbb{X}} \theta \\ &\text{s.t. } \theta \geq \sum_{\omega\in\Omega} p_\omega^k h_\omega(x), \;\; k \in K. \end{aligned} \tag{A-5}$$

One can solve this problem using a cut generation approach. That is, the restricted master problem is solved with a smaller subset of constraints (A-5) in order to get $(\hat{x}, \hat{\theta})$. Then, the worst-case expected problem is solved at $\hat{x}$ to obtain optimal value $f_\gamma(\hat{x})$. If $\hat{\theta} < f_\gamma(\hat{x})$, then the optimality cut (A-5)—corresponding to the extreme point $\mathbf{p}^k$, $k \in K$, that solves the worst-case expected problem at $\hat{x}$—is added to the restricted master problem. The Primal Decomposition approach to obtain x^* and $\mathbf{p}^*$ is presented in Algorithm A-1.

Algorithm A-1 Primal Decomposition

STEP 0. Initialization
Set $k \leftarrow 0$, $\bar{z} \leftarrow +\infty$. Add initial cuts (e.g., $\theta \geq 0$), if available.
STEP 1. Update lower bound
Set $k \leftarrow k + 1$.
Solve the current restricted master problem and obtain $(\hat{x}, \hat{\theta})$.
Set lower bound $\underline{z} = \hat{\theta}$.
STEP 2. Forward pass to obtain $\mathbf{h}(\hat{x})$
for ω **do**
  Solve $h_\omega(\hat{x})$
end for
STEP 3. Backward pass to obtain $\mathbf{p}^k$
Solve the worst-case expected problem at $\hat{x}$ and obtain $\mathbf{p}^k$ and $f_\gamma(\hat{x})$.
STEP 4. Update upper bound
if $f_\gamma(\hat{x}) < \bar{z}$ **then**
  Set $\bar{z} = f_\gamma(\hat{x})$, $x^* = \hat{x}$, and $p^*_\omega = p^k_\omega$ $\forall \omega$.
end if
STEP 5. Check stopping criterion
if $\bar{z} = \underline{z}$ **then**
  STOP.
  Output x^* and $\mathbf{p}^*$ as an optimal solution/worst-case probability pair.
end if
STEP 6. Backward pass to generate cut
Augment the current restricted master problem's set of cuts with $\theta \geq \sum_{\omega \in \Omega} p^k_\omega h_\omega(x)$.
Go to STEP 1.

Since the restricted master problem is a relaxation of (DRSP-V), $\hat{\theta}$ provides a lower bound on the optimal value of (DRSP-V). Moreover, since $\hat{x} \in \mathbb{X}$, $f_\gamma(\hat{x})$ gives an upper bound on the optimal value of (DRSP-V). The algorithm continues until $\hat{\theta} = f_\gamma(\hat{x})$, where $\hat{x}$ and the extreme point that solves the worst-case expected problem at $\hat{x}$ are an optimal x^* and $\mathbf{p}^*$, respectively. Because the polytope $\mathcal{P}_\gamma$ has finitely many extreme points, finitely many optimality cuts (A-5) are added before x^* and $\mathbf{p}^*$ are obtained. In other words, the Primal Decomposition algorithm converges in a finite number of iterations. However, at each iteration a convex optimization problem is solved. For algorithms to solve these problems, we refer interested readers to [7] and references therein. In the particular case of SLP-2, one can apply a decomposition-based method to obtain an outer approximation to $h_\omega(x)$. In fact, this is what we use in our numerical illustration in Sect. 5.

Algorithm A-1 decomposes primal formulation of (DRSP-V) and obtain x^* and $\mathbf{p}^*$ concurrently and directly. Because the Primal Decomposition exploits the polyhedral structure of the ambiguity set formed using the total variation distance, it might have computational benefits over those decomposition methods that solve the dual formulation. A comprehensive computational study will provide more insight on the performance of these algorithms; however, such a study is out of the scope of this paper.

References

1. Acerbi, C.: Spectral measures of risk: a coherent representation of subjective risk aversion. J. Bank. Financ. **26**(7), 1505–1518 (2002)
2. Analui, B., Pflug, G.C.: On distributionally robust multiperiod stochastic optimization. Comput. Manag. Sci. **11**(3), 197–220 (2014)
3. Artzner, P., Delbaen, F., Eber, J.M., Heath, D.: Coherent measures of risk. Math. Financ. **9**(3), 203–228 (1999)
4. Bayraksan, G., Love, D.K.: Data-driven stochastic programming using phi-divergences. In: Tutorials in Operations Research, INFORMS, pp. 1–19 (2015)
5. Ben-Tal, A., den Hertog, D., De Waegenaere, A., Melenberg, B., Rennen, G.: Robust solutions of optimization problems affected by uncertain probabilities. Manag. Sci. **59**(2), 341–357 (2013)
6. Bonnans, J.F., Shapiro, A.: Perturbation Analysis of Optimization Problems. Springer, Berlin (2013)
7. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)
8. Calafiore, G.: Ambiguous risk measures and optimal robust portfolios. SIAM J. Optim. **18**(3), 853–877 (2007)
9. Calafiore, G., Campi, M.C.: Uncertain convex programs: randomized solutions and confidence levels. Math. Program. **102**(1), 25–46 (2005)
10. Campi, M.C., Garatti, S.: The exact feasibility of randomized solutions of uncertain convex programs. SIAM J. Optim. **19**(3), 1211–1230 (2008)
11. Carè, A., Garatti, S., Campi, M.C.: Scenario min–max optimization and the risk of empirical costs. SIAM J. Optim. **25**(4), 2061–2080 (2015)
12. Delage, E., Ye, Y.: Distributionally robust optimization under moment uncertainty with application to data-driven problems. Oper. Res. **58**(3), 595–612 (2010)
13. Erdoğan, E., Iyengar, G.: Ambiguous chance constrained problems and robust optimization. Math. Program. **107**(1–2), 37–61 (2006)
14. Esfahani, P.M., Kuhn, D.: Data-driven distributionally robust optimization using the Wasserstein metric: performance guarantees and tractable reformulations. Math Program. (2017). https://doi.org/10.1007/s10107-017-1172-1
15. Gao, R., Kleywegt, A.J.: Distributionally robust stochastic optimization with Wasserstein distance (2016). arXiv:160402199v2 [mathOC]
16. Hanasusanto, G.A., Kuhn, D., Wallace, S.W., Zymler, S.: Distributionally robust multi-item newsvendor problems with multimodal demand distributions. Math. Program. **152**(1), 1–32 (2014)
17. Hanasusanto, G.A., Roitch, V., Kuhn, D., Wiesemann, W.: A distributionally robust perspective on uncertainty quantification and chance constrained programming. Math. Program. **151**(1), 35–62 (2015)
18. Higle, J., Sen, S.: Stochastic Decomposition: A Statistical Method for Large Scale Stochastic Linear Programming. Kluwer Academic Publishers, Dordrecht (1996)
19. Hiriart-Urruty, J.B., Lemaréchal, C.: Fundamentals of Convex Analysis. Springer, Berlin (2001)
20. Hu, Z., Hong, L.J.: Kullback–Leibler divergence constrained distributionally robust optimization. Optimization Online (2012). http://www.optimization-online.org/DB_HTML/2012/11/3677.html
21. Hu, Z., Hong, L.J., So, A.M.C.: Ambiguous probabilistic programs. Optimization Online (2013). http://www.optimization-online.org/DB_HTML/2013/09/4039.html
22. Infanger, G.: Monte Carlo (importance) sampling within a Benders decomposition algorithm for stochastic linear programs. Ann. Oper. Res. **39**(1), 69–95 (1992)
23. Jiang, R.: Data-driven stochastic optimization: integrating reliability with cost effectiveness. Ph.D. dissertation, University of Florida, Gainesville (2013)
24. Jiang, R., Guan, Y.: Data-driven chance constrained stochastic program. Math. Program. **158**(1), 1–37 (2015a)
25. Jiang, R., Guan, Y.: Risk-averse two-stage stochastic program with distributional ambiguity. Optimization Online (2015). http://www.optimization-online.org/DB_HTML/2015/05/4908.html

26. Klabjan, D., Simchi-Levi, D., Song, M.: Robust stochastic lot-sizing by means of histograms. Prod. Oper. Manag. **22**(3), 691–710 (2013)
27. Love, D.K.: Data-driven methods for optimization under uncertainty with application to water allocation. Ph.D. dissertation, University of Arizona, Tucson (2013)
28. Love, D.K., Bayraksan, G.: Phi-divergence constrained ambiguous stochastic programs for data-driven optimization. Optimization Online (2015). http://www.optimization-online.org/DB_HTML/2016/03/5350.html
29. Luo, F., Mehrotra, S.: Decomposition algorithm for distributionally robust optimization using Wasserstein metric (2017). arXiv:170403920 [mathOC]
30. Mehrotra, S., Zhang, H.: Models and algorithms for distributionally robust least squares problems. Math. Program. **146**(1–2), 123–141 (2014)
31. Pflug, G.C., Pichler, A.: The problem of ambiguity in stochastic optimization. In: Mikosch, T.V., Resnick, S.I., Robinson, S.M. (eds.) Multistage Stochastic Optimization, pp. 229–255. Springer (2014)
32. Pflug, G.C., Wozabal, D.: Ambiguity in portfolio selection. Quant. Financ. **7**(4), 435–442 (2007)
33. Rahimian, H., Bayraksan, G., Homem-de Mello, T.: Distributionally robust newsvendor problems with variation distance. Optimization Online (2017). http://www.optimization-online.org/DB_HTML/2017/03/5936.html
34. Rockafellar, R.T., Uryasev, S.: Conditional value-at-risk for general loss distributions. J. Bank. Financ. **26**(7), 1443–1471 (2002)
35. Römisch, W.: Stability of stochastic programming problems. In: Ruszczyński, A., Shapiro, A. (eds.) Stochastic Programming, Handbooks in Operations Research and Management Science, vol. 10, pp. 483–554. Elsevier, Amsterdam (2003)
36. Rossi, M.E., Deutsch, C.V.: Mineral Resource Estimation. Springer, Berlin (2014)
37. Ruszczynski, A., Shapiro, A.: Optimization of convex risk functions. Math. Oper. Res. **31**(3), 433–452 (2006)
38. Scarf, H.: A min–max solution of an inventory problem. In: Scarf, H., Arrow, K., Karlin, S. (eds.) Studies in the Mathematical Theory of Inventory and Production, vol. 10, pp. 201–209. Stanford University Press, Stanford (1958)
39. Shapiro, A.: Rectangular sets of probability measures. Oper. Res. **64**(2), 528–541 (2016)
40. Shapiro, A.: Distributionally robust stochastic programming. SIAM J. Optim. **27**(4), 2258–2275 (2017)
41. Shapiro, A., Tekaya, W., Soares, M.P., da Costa, J.P.: Worst-case-expectation approach to optimization under uncertainty. Oper. Res. **61**(6), 1435–1449 (2013)
42. Shapiro, A., Dentcheva, D., Ruszczyński, A.: Lectures on Stochastic Programming: Modeling and Theory. MPS-SIAM Series on Optimization, 2nd edn. Society for Industrial and Applied Mathematics, Philadelphia (2014)
43. Wang, Z., Glynn, P.W., Ye, Y.: Likelihood robust optimization for data-driven problems. Comput. Manag. Sci. **13**(2), 241–261 (2016)
44. Wiesemann, W., Kuhn, D., Sim, M.: Distributionally robust convex optimization. Oper. Res. **62**(6), 1358–1376 (2014)
45. Wozabal, D.: A framework for optimization under ambiguity. Ann. Oper. Res. **193**(1), 21–47 (2012)
46. Xin, L., Goldberg, DA., Shapiro, A.: Time (in)consistency of multistage distributionally robust inventory models with moment constraints (2013). arXiv:13043074 [mathOC]
47. Zhang, W., Rahimian, H., Bayraksan, G.: Decomposition algorithms for risk-averse multistage stochastic programs with application to water allocation under uncertainty. INFORMS J. Comput. **28**(3), 385–404 (2016)
48. Zhao, C., Guan, Y.: Data-driven risk-averse two-stage stochastic program with ζ-structure probability metrics. Optimization Online (2015). http://www.optimization-online.org/DB_HTML/2015/07/5014.html