An Empirical Approach to Reliability-based Design using Scenario Optimization

Roberto Rocchetta

Department of Mathematics and Computer Science, Security W& I, Technical University of Eindhoven, Eindhoven, The Netherlands. E-mail: roberto.rocchetta@tue.nl

Luis G. Crespo

Dynamic Systems and Control Branch, NASA, Langley Research Center, Hampton, Virginia, USA. E-mail: luis.g.crespo@nasa.gov

Scenario-based approaches to Reliability-Based Design-Optimization were recently proposed by the authors, Rocchetta et al. (2019). Scenario theory makes direct use of the available data thereby eliminating the need for creating a probabilistic model of the uncertainty in the parameters. This feature makes the resulting design exempt from the subjectivity caused by prescribing an uncertainty model from insufficient data. Most importantly, scenario theory renders a formally verifiable bound to the probability of failure. This bound is non-asymptotic and holds for any probabilistic model consistent with the available data. In this article we seek designs that minimize a combination of cost and penalty terms caused by violating reliability constraints. Similar to Conditional-Value-at-Risk programs, the proposed optimization approach is convex, thereby easing its numerical implementation. As opposite to a Conditional-Value-at-Risk method, a model for the uncertainty is not required and the method provides bounds on the reliability, which is valuable information to assess the robustness of the prescribed design. Furthermore, the proposed approach enables the analyst to shape the distribution of the design's performance according to a given value-at-risk. This is done by minimizing the empirical approximation of the integral of the design's performance in the loss/failure region. The effectiveness of the approach is tested on an easily reproducible numerical example with its strengths discussed in comparison to traditional methods.

Keywords: Scenario Optimization, Reliability Bounds, Conditional-Value-at-Risk, Constraints Relaxation, Uncertainty

1. Introduction

Reliability-Based Design-Optimization (RBDO) seeks a design which minimizes cost such that a set of reliability-based constraints are satisfied. For instance, geometries of components must be selected to minimize manufacturing costs while the probability of facing structural failures remains within specification. A standard approach to RBDO problems involves two nested loops, an outer loop searches for an optimal design whereas an inner loop evaluates its cost and reliability. The majority of the existing RBDO methods rely on a suitable model of the uncertainty, which is generally needed to estimate the failure probability in the inner loop, e.g., via Monte Carlo integration. However, a model of the uncertainty can be challenging to prescribe because of complex parameter dependencies and lack of data in low-probability regions. This will have an impact in a design optimized using such a model and, in practice, can potentially lead to unexpected reliability performance and, in the worst case, to hazardous situations and severe failures. Furthermore, the cost and the reliability requirements often define conflicting objectives. To overcome these difficulties, a chance-constrained and distribution-free reformulation of the RBDO program is advisable.

Chance-Constrained Programs (CCPs) minimize the design cost while constraining its reliability to a satisfactory level. CCPs are generally NP-hard, usually non-convex, and chance-constraints on joint probabilistic requirements are significantly more challenging than individual constraints, Özgün Elçi et al. (2018). Non convexity is caused by the form of the requirement functions, whereas sampling-based approaches to the estimation of the failure probability introduce discontinuities that make gradient-based algorithms inapplicable. Intractability of CCPs lead to alternative solution techniques, e.g., convexification techniques have been introduced to replace the non-convex set of feasible designs with a convex inner approximation. Specifically, replacing a failure probability constraint with the Conditional Value-at-Risk (CVaR) improve the numerical tractability of the optimization, see for instance the work of Rockafellar and Royset (2010) on buffered failure probabilities. CVaR

Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference
Edited by Piero Baraldi, Francesco Di Maio and Enrico Zio
Published by Research Publishing, Singapore.
ISBN/DOI: TBA

constraints are guaranteed to be convex in the uncertainty space and offer control on a whole portion of the tail of the distribution and not only a quantile.

One of the main drawbacks of a CVaR constraint versus a failure probability constraint is that the former is statistically less stable, i.e., an outlier can significantly change the value of the estimated CVaR. Both failure probability and CVaR estimation rely on a good model describing the tails of the distribution (which can be hard to provide). This limitation pushed research toward the development of distributionally robust CCPs. These optimization methods seek a design which is reliable for a whole set of uncertainty models. Evidence theory, Possibility theory, Fuzzy sets and Credal sets theory are some of the most widely applied concepts, see, e.g., Beer et al. (2013). Only a limited amount of works investigated given-data (model-free) CCPs to identify an optimal design by making direct use of the data and without prescribing an uncertainty model.

Scenario-based decision-making theory offers a powerful framework to solve CCP according to data. Scenario optimization has been applied to tackle prediction Campi et al. (2015), regression, machine learning, and optimal control problems Rocchetta et al. (2019). To the best of the authors' knowledge, only a few works investigated applicability of Scenario theory to RBDO problems. Rocchetta et al. (2020) developed a Scenario RBDO framework to solve RBDO problems and a powerful prospective-reliability certificate was obtained for the optimized design, i.e., an upper bound on the probability of facing catastrophic failures (failures of magnitude greater than the historically observed worst-case). However this certificate only focused on extreme cases, and a prospective-reliability bound on the failure probability was not provided.

In this work we extend the work of Rocchetta et al. (2020) to provide an upper and lower bounds on the probability of failure. A novel scenario program for RBDO is proposed based on the theoretical results of Garatti and Campi (2019). This optimization scheme minimizes a weighted sum of the cost and penalty terms for constraint violations. The proposed scenario program shares similar features when compared to a traditional Conditional-Value-Risk reliability optimization, i.e., it ensures convexity for a wide class of reliability performance functions and offer control on a whole portion of the reliability performance tail and not only a quantile. In contrast, a prospective-reliability certificate can be obtained for the optimized design. This certificate bounds the probability of exceeding a predefined value-at-risk level, i.e., it is a lower and an upper bound on the failure probability or on the tails of the distribution of the reliability performance function. This certificate is obtained given-data and without the need to prescribe a model (or a set of models) of the uncertainty. Thus, it is exempt from the subjectivity caused by prescribing a model for the distribution tails from insufficient data.

The rest of the paper is organized as follows: Section 2 presents the mathematical background on RBDO and CVaR approximation. Section 3 introduces Scenario optimization theory and theoretical robustness guarantees. In Section 4 the newly proposed scenario RBDO programs are presented. Section 5 exemplifies the method on an easily reproducible case study. Section 6 closes the paper with a discussion on the results.

2. Mathematical background

A reliability chance-constrained program seeks an optimal design $d^\star$ as follows:

$$d^\star = \arg\min_{d\in\Theta}\{J(d) : P_f(d) < 1-\alpha\}$$
$$P_f(d) = \int_{\mathcal{F}(d)} f_\delta(\delta) d\delta \tag{1}$$

where d are the optimization variables constrained in a closed convex set $\Theta \subseteq \mathbb{R}^{n_d}$, $J(d) : \mathbb{R}^{n_d} \to \mathbb{R}$ is a convex cost function, $\alpha \in [0,1]$ is a probabilistic level constraining the design's failure probability $P_f(d)$. Note that $\alpha = 1$ corresponds to an admissible failure probability equal to zero. $P_f(d)$ is a multidimensional integral of the uncertainty model, $f_\delta(\delta)$, a joint PDF of uncertain factors $\delta \in \Delta \subseteq \mathbb{R}^{n_\delta}$, computed over the composite failure domain $\mathcal{F}(d)$. The domain $\mathcal{F}(d)$ is generally defined by n_g reliability requirements $\mathcal{F}(d) = \bigcup_{j=1}^{n_g} \mathcal{F}^j(d)$, where

$$\mathcal{F}^j(d) = \{\delta \in \Delta : g_j(d,\delta) \geq 0\}$$

are the individual failure regions induced by the *reliability performance functions* $g_j : \mathbb{R}^{n_d} \times \mathbb{R}^{n_\delta} \to \mathbb{R}$. A design d satisfies all requirements for δ if $g_j(d,\delta) < 0,\ \forall j \in \{1,..,n_g\}$. An equivalent formulation of program (1) is[a],

$$d^\star = \arg\min_{d\in\Theta}\{J(d) : F_w^{-1}(\alpha) < 0\} \tag{2}$$

where $F_w^{-1}(\alpha)$ is the *Value-at-Risk* at level α, i.e., the α-percentile of the distribution of the worst-case performance function

$$w(d,\delta) = \max_{j\in\{1,..,n_g\}} g_j(d,\delta)$$

[a]The constraint $P_f(d) < 1-\alpha$ implies $\mathbb{P}[w(d,\delta) \geq 0] < 1-\alpha$ which is equivalent to $F_w^{-1}(\alpha) < 0$

induced by the design d and for the uncertainty model f_δ. When $w(d, \delta) < 0$ the design d satisfies all the reliability requirements for the uncertainty realization δ.

A solution of (2) is often computationally demanding to obtain. This is due to the multidimensional integral (1) which must be solved numerous times. This issue has led to the development of approximations, e.g., Most Probable Point or sample-based methods. Given $\mathcal{D}_N = \{\delta^{(i)}\}_{i=1}^N$ samples drawn from f_δ, the integral in (1) can be approximated by,

$$\hat{P}_f(d) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{w(d,\delta^{(i)})\geq 0} \tag{3}$$

where $\mathbf{1}_{w(d,\delta^{(i)})\geq 0}$ is the indicator function for the failure condition $w(d, \delta^{(i)}) \geq 0$. Similarly, the empirical CDF of w is computed by,

$$F_{w(d,\mathcal{D}_N)}(W) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{w(d,\delta^{(i)})\leq W} \tag{4}$$

which gives the empirical quantile at level α, $F^{-1}_{w(d,\mathcal{D}_N)}(\alpha)$. Once an estimator $F_w^{-1}(\alpha)$ or the empirical P_f based on a Monte Carlo sample is obtained, an optimization routine can be readily applied. Note that an uncertainty model f_δ is a key component of the process. Furthermore, $F^{-1}_{w(d,\mathcal{D}_N)}(\alpha)$ is a discontinuous function and this further complicates solution of program (2).

The feasibility set of the chance-constrained program (2) for a given α level is $\Theta_\alpha \subseteq \Theta$, i.e., the set of designs satisfying the given constraints, $\Theta_\alpha = \{d \in \Theta : F^{-1}_{w(d,\mathcal{D}_N)}(\alpha) \leq 0\}$. With the exception of logarithmically concave distributions f_δ, Θ_α is non-convex. Hence, the optimization problem is generally non-convex even when the reliability performance functions are convex in d for any fixed δ. Also, the constraint in (2) gives no guarantees on the severity of failures, i.e., the value of $w(d, \delta)$ can take arbitrarily large values. If the value of w when $w > 0$ is a measure of the severity of the reliability violation, the analyst might want to control not only the failure probability but also the shape of the upper tail of w. This design criterion will be considered below.

This concept is depicted in Figure 1 which shows an example of reliability CCP. The CDFs of w for three feasible designs are presented. The designs satisfies the probabilistic constraint for the level α and d_2 leads to the lower failure probability. However, this does not imply d_2 might lead to failures with lower severity where w can take arbitrarily large values.

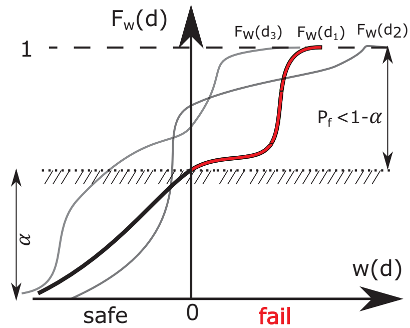

Fig. 1. CDFs of the worst-case performance w associated to 3 feasible designs, i.e., d for which $F_w^{-1}(\alpha) \leq 0$. d_2 is the most reliable but with no guarantee on the severity of failures.

2.1. *CVaR approximation*

Conditional Value-at-Risk has been used to approximate the probabilistic constraint in (2) and for any continuous distributions f_δ is defined as follows:

$$CV_\alpha^{f_\delta}(w) = \frac{1}{1-\alpha}\int_\alpha^1 F_w^{-1}(\beta)d\beta \tag{5}$$

$CV_\alpha^{f_\delta}(w)$ is a continuous function of α and is an expectation over a 'portion' of the upper tail of the distribution of w, i.e., $\mathbb{E}[w|w \geq F_w^{-1}(\alpha)]$. When the constraint in (2) is satisfied this does not imply the integral (5) is less or equal to zero. On the other hand, for the non-decreasing inverse CDF we have $\int_\alpha^1 F_w^{-1}(\beta)d\beta > F_w^{-1}(\alpha)$ and, thus, a $CV_\alpha^{f_\delta} \leq 0$ implies $F_w^{-1}(\alpha) \leq 0$. A CVaR constrained approximation of program (2) is as follows:

$$d^\star = \arg\min_{d\in\Theta}\{J(d) : CV_\alpha^{f_\delta}(w(d,\delta)) \leq 0\} \tag{6}$$

where $CV_\alpha^{f_\delta} \leq 0$ offers a convex inner approximation of the feasibility set Θ_α. This makes optimization problem (6) convex for any performance functions $g_i(d, \delta)$ convex in d[b], thus simplifying its solution. Furthermore, this formulation guarantees a conservative result in terms of failure probability, see e.g., Rockafellar and Royset (2010). Figure 2 illustrate this concept by presenting feasible/unfeasible regions for a VaR constrained program (red regions) a CVaR program (6) when applied to a linear limit state function $w = d_1 + \delta_1 - d_2\delta_2$ with $\alpha = 0.1$ and $\alpha = 0.9$. It can be noticed that the feasible set induced by (6) is always contained in the feasibility set of

[b]This implies that w is also convex as the maximum operator preserves convexity

program (2). Moreover, even for a linear w the feasibility set of program (2) is non-convex, see $\alpha = 0.1$.

A sampling-based estimator of the CVaR is

$$CV_{\alpha}^{\mathcal{D}_N}(w) = \frac{\sum_{i=1}^{N} w(d, \delta^{(i)})\mathbf{1}_{c^{(i)}}}{\sum_{i=1}^{N} \mathbf{1}_{c^{(i)}}} \quad (7)$$

where $\mathbf{1}_{c^{(i)}}$ is the indicator function for the condition $c^{(i)} = \{w(d, \delta^{(i)}) \geq F^{-1}_{w(d,\mathcal{D}_N)}(\alpha)\}$, i.e., a condition for which $w(d, \delta^{(i)})$ exceeds the α-quantile of the empirical distribution of w evaluated for $\delta^{(i)} \in \mathcal{D}_N$. Program (6) has, however, some drawbacks: 1) CVaR estimation is sensitive to the uncertainty model f_δ especially in the tails regions; 2) A CVaR constraint can be very stringent and the convex inner approximation of Θ_α potentially empty for a level α.

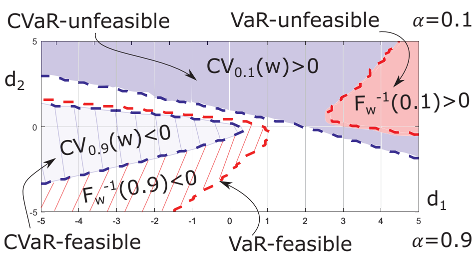

Fig. 2. Feasible/unfeasible set of designs (d_1, d_2) induced by VaR and CVaR constrained programs (2) and (6).

3. Scenario theory

Consider the probability space Δ with associated σ-algebra $\mathcal{F}$ and a stationary probability measure $\mathbb{P}$, see Calafiore and Campi (2006). In practice, the probability $\mathbb{P}$ is unknown and only a data set $\mathcal{D}_N = \{\delta^{(i)}\}_{i=1}^{N} \in \Delta^N$ containing N realizations of the uncertain parameters is available. In scenario optimization programs each realization $\delta^{(i)} \in \mathcal{D}_N$ is a *scenario* and a scenario RBDO program $\mathcal{SP}(\mathcal{D}_N)$ can defined as follows:

$$\min_{d \in \Theta} \{J(d) : \ d \in \Theta_{\delta^{(i)}}, \ \forall \delta^{(i)} \in \mathcal{D}_N\} \quad (8)$$

where $\Theta_{\delta^{(i)}} = \{d \in \Theta : w(d, \delta^{(i)}) \leq 0\}$ is the feasibility set induced by the i^{th} scenario and the set $\mathcal{D}_N$ defines N scenario constraints which can be used to replace probabilistic constraints in classical CCPs. The optimized design solution of $\mathcal{SP}(\mathcal{D}_N)$ is $d^\star$.

Scenario theory can be used to obtain a prospective-reliability certificate for $d^\star$, i.e., how well it generalizes to yet unseen situations $\delta \in \Delta$, see, e.g., Campi and Garatti (2008) . To explain the scenario prospective-reliability it is useful to introduce the concepts of violation probability and support set.

Definition: *(Violation probability, or risk) The probability*

$$V(d^\star) = \mathbb{P}[\delta \in \Delta : d^* \notin \Theta_\delta]$$

is called violation probability. Given a reliability parameter $\epsilon \in (0, 1)$, a design $d^\star$ is called ϵ-robust (or ϵ-feasible) if $V(d) \leq \epsilon$, Campi and Garatti (2018). An ϵ-robust solution will comply with the requirements induced by new scenarios with probability no less than 1-ϵ.

Definition: *A set of support constraints (or support set) $\mathcal{S} \subseteq \mathcal{D}_N$ is a k-tuple $\mathcal{S} = \{\delta^{(i1)}, ..., \delta^{(ik)}\}$ for which the solution of $\mathcal{SP}(\mathcal{S})$ is the same as $\mathcal{SP}(\mathcal{D}_N)$.* The set $\mathcal{S}$ is of minimal cardinality if no further scenarios can be removed without changing the optimal design $d^\star$. The cardinality of the set of support constraints is $s_N^\star = |\mathcal{S}|$, where $|\cdot|$ is the cardinality operator.

The violation probability, $V(d^\star)$ and support set size $s_N^\star$ are inherently stochastic, as they depend on the random set of scenarios $\mathcal{D}_N$. However, it is proven that for convex scenario programs[c], the distribution $V(d^\star)$ is dominated by the Beta distribution, Campi and Garatti (2008). This result provides a powerful robustness-monitoring capability, a bound on $V(d^\star)$ which is guaranteed to hold distribution-free and non-asymptotically. However, a $\Theta_{\delta^{(i)}}$, defined as in (8), implies that a feasible optimal design $d^\star$ will have an empirical failure probability $\hat{P}_f(d^\star) = 0$, see eq.(3). This requirement might not only make the program (8) infeasible, i.e., $\bigcup_{i=1}^{N} \Theta_{\delta^{(i)}} = \emptyset$, but might also lead to high cost values. In the next section we adopt the strategy proposed by Garatti and Campi (2019) to mend for these deficiencies.

4. The proposed Scenario program

Consider the scenario program:

$$\langle d^\star, \zeta^\star \rangle = \arg\min_{\substack{d \in \Theta \\ \zeta \geq \lambda}} \{J(d) + \rho \sum_{i=1}^{N} (\zeta^{(i)} - \lambda) : \\ w(d, \delta^{(i)}) \leq \zeta^{(i)}, \ \delta^{(i)} \in \mathcal{D}_N\} \quad (9)$$

where $\zeta \in \mathbb{R}^N$ represents a vector of slack variables associated to the N reliability constraints,

[c]Under the existence, uniqueness and non-degeneracy assumptions and for any stationary $\mathbb{P}$ and N iid δ.

$\rho > 0$ is a constant value used to penalize constraint violations, an optimization parameter weighting the importance of minimizing reliability violation $\sum_{i=1}^{N}(\zeta^{(i)} - \lambda)$ versus the design cost $J(d)$, and $\lambda \in \mathbb{R}$ is a value-at-risk level which define a lower bound on the slack variables. Program (9) shares some similarity with the CVaR constrained optimization program (6). Program (9) with $\lambda = 0$ seeks a design which minimizes the cost $J(d)$ and the sum of reliability violations. This implies a minimization of the integral of the distribution of w in the failure region, i.e., a combination of cost and constraint in (6).

A $\lambda = 0$ implies that all the non-zero terms in the vector $\zeta^\star$ correspond to scenarios falling into the failure region. The magnitude of $\zeta^{(i)} > 0$ is an indicator of the severity of the reliability violation, i.e., scenarios for which $\zeta^{(i)} = w(d, \delta^{(i)})$. In contrast, a $\lambda \neq 0$ defines a program that seeks an optimal design which minimizes a combination of cost and violation of the constraints $w(d, \delta) \leq \lambda$. Hence, a $\lambda < 0$ means that program (9) is imposing a more stringent constraint that $w(d, \delta) \leq 0$ on each scenario. Conversely, $\lambda > 0$ indicates a program that relaxes the constraint violation. In the context of CVaR this constraint can be interpreted as$F_w^{-1}(\alpha) \leq \lambda$, where values of λ different from zero denote a relaxation or tightening of the constraints.

Note that the penalty terms in (9) enable the analyst to exercise a certain degree of control over the failure event besides the failure probability and can be conveniently used to: 1) Improve feasibility of reliability chance-constrained programs; 2) Modulate the distribution of w in the loss/failure region, i.e., minimizing $CV_\alpha^{\mathcal{D}_N}(w)$ where the value at risk is $F_w^{-1}(\alpha) = \lambda$; 3) Trade reliability to improve the cost function, by tuning ρ. When $\rho \to \infty$ the program goes back to the original formulation in (8) where no constraint violation is allowed.

The work of Garatti and Campi (2019) provides a way to quantify the prospective-reliability in the context of optimization with relaxation of scenario constraints. Specifically, for any Δ and $\mathbb{P}$ it holds that:

$$\mathbb{P}^N[\underline{\epsilon}(s_N^\star) \leq V(d^\star) \leq \overline{\epsilon}(s_N^\star)] \geq 1 - \beta \quad (10)$$

where $\underline{\epsilon}$ and $\overline{\epsilon}$ are lower and upper bounds on the violation probability $V(d^\star) = \mathbb{P}[\delta \in \Delta : w(d^\star, \delta) > \lambda]$, $\beta \in [0, 1]$ is a confidence level set by the user, e.g., $\beta = 10^{-8}$ means almost certainty, and $s_N^\star = |\mathcal{S}|$ is the number of support constraints of the scenario program (9). The support set $\mathcal{S}$ accounts for violated constraints, $\zeta^{(i)} > \lambda$, and active constraints, $w(d, \delta^{(i)}) = \lambda$, as follows:

$$\mathcal{S} = \{\delta^{(i)} : \zeta^{(i)} > \lambda \vee w(d, \delta^{(i)}) = \zeta^{(i)}\} \quad (11)$$

Fixing a confidence level β and for $\lambda = 0$ Eq. (10) gives a lower and upper bound on the design failure probability, i.e.,

$$\mathbb{P}[\delta \in \Delta : w(d^\star, \delta) \geq 0] \in [\underline{\epsilon}(s_N^\star), \overline{\epsilon}(s_N^\star)]$$

where $\underline{\epsilon}(k) = \max\{0, 1 - \overline{t}(k)\}$, $\overline{\epsilon}(k) = 1 - \underline{t}(k)$ and $[\underline{t}, \overline{t}]$ are solutions a polynomial equation in t, see Theorem 4 Garatti and Campi (2019):

$$\binom{N}{k} t^{N-k} - \frac{\beta}{2N} \sum_{i=k}^{N-1} \binom{i}{k} t^{i-k} - \frac{\beta}{6N} \sum_{i=N+1}^{4N} \binom{i}{k} t^{i-k} = 0 \quad (12)$$

The bounds $[\underline{\epsilon}, \overline{\epsilon}]$ hold non-asymptotically, free from a model of the uncertainty and for any RBDO problem which the performance functions $g_j(d, \delta)$, $j = 1, .., n_g$ and the cost $J(d)$ are convex in d.

5. Numerical Example

The proposed method is tested on a simple RBDO problem having multiple, competitive, algebraic performance functions. The problem has $n_g = 2$ reliability requirements, modified from Grooteman (2011), $n_d = 2$ design variables, and $n_\delta = 2$ uncertain factors. A low dimensional uncertainty space is selected to ease the visualization of the results, however, for scenario programs the dimension n_δ is inconsequential Rocchetta et al. (2019). The two reliability performance functions are,

$$g_1(d, \delta) = -d_1 + \delta_1 - 0.532\delta_2 - 2d_2(\delta_1 - \delta_2)^2$$
$$g_2(d, \delta) = -d_1(1 - \delta_2) - 0.1064\delta_1^2 - d_2\delta_1^3$$

and the worst-case performance is just $w(d, \delta) = \max(g_1, g_2)$ and is a convex function in d but not in δ. The optimization problem seeks a reliable design (d_1, d_2) in $[-5, +5]^2$ so that a cost function $J(d) = -\sum_{i=1}^{2} d_i$ is minimized. The MATLAB's *fmincon* optimizer and the *'sqp'* algorithm are the numerical tools used to solve the problem. A baseline design $d_{bl} = [2, 3]$ is arbitrarily selected for comparison and is the initial guess provided to the solver.

We consider two sets of scenarios $\mathcal{D}_{10^3}$ and $\mathcal{D}_{10^5}$ obtained from an unknown stationary Data-Generating Mechanism (DGM). Figure 3 displays the sets with black round markers and red diamond markers, respectively. The data set with

$N = 10^3$ samples is used the optimization routines whilst the set with $N = 10^5$ scenarios is considered unavailable for the RBDO tasks and used to validate the prospective-reliability bounds. It can be noticed that the DGM (designed by the authors but not used for RBDO) generates outliers with a small probability and those are not captured adequately by the smaller set $\mathcal{D}_{10^3}$. This stationary DGM was designed to test the scenario prospective-reliability bounds $[\underline{\epsilon}, \overline{\epsilon}]$, see Section 4, which must hold with high confidence and independently from the stationary probability $\mathbb{P}$ generating the data.

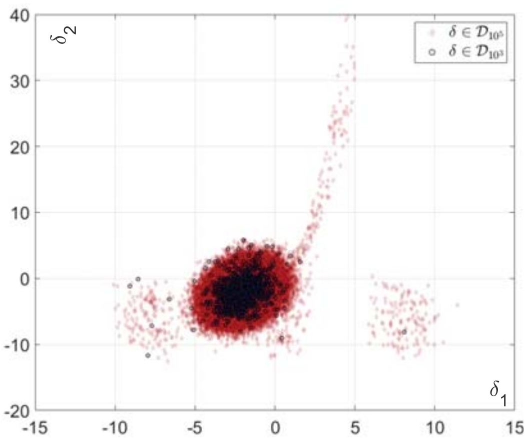

Fig. 3. The two sets of scenarios $\mathcal{D}_{10^3}$, round black markers, and the validation set $\mathcal{D}_{10^5}$, diamond red markers.

The proposed scenario program (9) is used to solve the proposed RBDO problem. The weight of the reliability violation is set to $\rho = 100$ and $N = 10^3$ scenario constraints on w have to be met for the new design to be feasible.

5.1. *Results for program (6)*

Program (6) is used to solve the RBDO problem and an $\alpha = 0.85$ is selected to constrain the probability of facing $w > 0$. A Gaussian mixture model with 5 mixtures is fitted to $\mathcal{D}_{10^3}$ and used to estimate the CVaR constraint on w. Table 1 compares the reliability performances of the baseline design, d_{bl}, and the optimized d^* resulting from program (6). The design cost $J(d)$, the CVaR $CV_{0.95}^{\mathcal{D}_{10^3}}(w)$ the failure probability estimators for joint and individual requirements and the empirical maximum of worst-case performance $\overline{w}(d) = \max_i w(d, \delta^{(i)})$ are proposed as figures of merit. Program (6) improves the reliability of the system substantially from $\hat{P}_f = 0.833$ for d_{bl} to $\hat{P}_f = 0.139$ which results, as expected, $0.139 \leq 1 - \alpha$. The CVAR is also reduced, from $CV_{0.95}^{\mathcal{D}_{10^3}}(w) = 328$ to only $CV_{0.95}^{\mathcal{D}_{10^3}}(w) = 1.27$, thus leading to an overall mitigation of the severity of failures. However, the total cost of the design $J(d)$ increased from -5 to about -0.57 and this is due to the existing trade-off between system cost and its reliability.

5.2. *Results for $\lambda = 0$ and increasing N*

Program (9) with $\lambda = 0$ is used to solve the RBDO problem. A violation of a scenario constraints occurs when $\zeta^{(i)} > 0$, that is, the i^{th} scenario fails to comply with at least one of the reliability requirements g_1 or g_2. Figure 4 presents the optimized vector of slack variables $\zeta^\star$ (red dashed line) and $w(d^\star, \delta^{(i)})$ (blue solid line) for the scenarios $\delta^{(i)} \in \mathcal{D}_{10^3}$ and in correspondence of the optimum design $d^\star$. The empirical CDFs of $w(d^\star, \delta^{(i)})$ and $\zeta^\star$ are presented in the bottom panel and compared to the result of the CVaR program (dashdotted line). It can be observed that for each $\zeta^{(i)} > 0$ the corresponding reliability violation is $w(d, \delta^{(i)}) = \zeta^{(i)}$ and thus, as expected, the proposed method minimizes a combination of $J(d)$ and the integral of w in the failure region expressed as a sum of $\zeta^{(i)}$. The last column in Table 1 presents the reliability performances of the proposed program, $d^\star$, which slightly penalizes the price for a gain in reliability when compared to program (6) and greatly improves the reliability compared to d_{bl}.

The prospective-reliability of $d^\star$ depends on the number of active and violated constraints, see equation 10, which results $s_N^\star = 146$. For a high confidence $\beta = 10^{-8}$ this leads to $\mathbb{P}[\delta \in \Delta : w(d^\star, \delta) \geq 0] \in [\underline{\epsilon}(s_N^\star), \overline{\epsilon}(s_N^\star)]$, where $[\underline{\epsilon}(s_N^\star), \overline{\epsilon}(s_N^\star)] = [0.0834, 0.2282]$. This is a powerful result which assures that the 'true' $P_f(d^\star)$ will result at worst 0.2282 and not better than 0.0834, hence informing the analyst on the robustness of $d^\star$ against the uncertainty affecting the DGM (due limited availability of data).

To investigate prospective-reliability bounds on P_f we solve problem (9) for an increasing availability of scenarios from $N = 100$ to $N = 2500$. Figure 5 shows the given data estimator $\hat{P}_f(d^\star)$, solid line, and its scenario bounds, red area, while Table 2 presents the numerical results of the analysis. As expected, the prospective-reliability bounds get tighter for an increasing availability of samples and always include the failure probability estimator. Tighter bounds are due to the increasing knowledge on the underlying $\mathbb{P}$ generating the data for higher N.

Table 1. Comparison between the reliability of d_{bl}, the optimal designs resulting from programs (6) and (9).

Design from program	d_{bl} -	d^* (6)	$d^\star$ (9)
(d_1, d_2)	(2,3)	(0.565,0.012)	(0.40,0.021)
Reliability & Performance (with $\mathcal{D}_{10^3}$)			
$J(d)$	-5	-0.577	-0.425
$CV_{0.95}^{\mathcal{D}_{10^3}}(w)$	328.2	1.27	1.084
$\hat{P}_f(d)$	0.833	0.139	0.146
$\hat{P}_{f,1}(d)$	0	0.081	0.08
$\hat{P}_{f,2}(d)$	0.833	0.058	0.066
$\overline{w}(d)$	2231	5.75	6.67

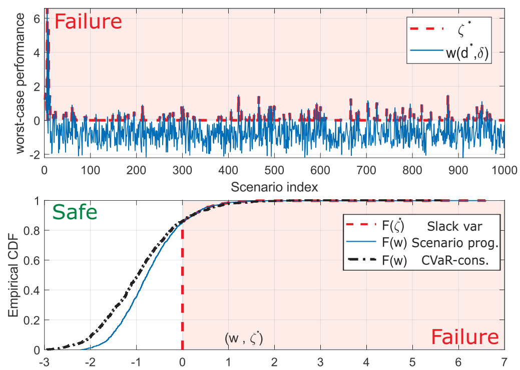

Fig. 4. The worst-case performance $w(d^\star, \delta)$, solid line, and $\zeta^\star$, dashed line, from $\mathcal{D}_{10^3}$ (the top panel). Their empirical CDFs and $w(d^*, \delta)$ from program (9), the bottom panel.

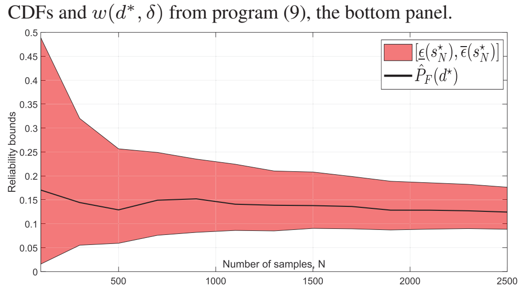

Fig. 5. The estimated failure probability $\hat{P}(d^\star)$ and the scenario bounds on the failure probability for an increasing number of scenarios N.

5.3. *Results for* $\lambda \neq 0$

Scenario program (9) is tested on the RBDO problem for $\lambda \in [-1, +1]$ and Table 3 summarizes the results. It can be observed that small λ values lead to wider scenario bounds on $\mathbb{P}[w(d^\star, \delta) \geq \lambda]$. For instance $\lambda = -1$ leads to a (random) number of support constraints $s^\star_{10^3} = 203$, which correspond to a $\mathbb{P}[w(d^\star, \delta) \geq -1] \in [0.129, 0.294]$. In contrast, $\lambda = 1$ leads to $s^\star_{10^3} = 24$ and thus a tighter bound $\mathbb{P}[w(d^\star, \delta) \geq 1] \in [0.004, 0.069]$. The latter means that no more than 6.9% of the unobserved scenarios will result in a worst-case performance $w \geq 1$ for the corresponding $d^\star$. Intuitively, the tighter bounds on $\lambda = 1$ are due to the weaker statement on the tails of w. Differently, the probabilistic statement $\mathbb{P}[w(d^\star, \delta) \geq -1]$ is 'stronger' as it is made on a wider portion of the tail of w. However, it is also less guaranteed and results in wider prospective-reliability bounds. A validation analysis of the scenario bounds is proposed and 'true' violation probability $\mathbb{P}[w(d^\star, \delta) \geq \lambda]$ estimated over the larger data set $\mathcal{D}_{10^5}$, see Figure 3. Results of this analysis are presented in Table 3 and it can be seen that the scenario bounds always includes the 'true' violation probability.

The scenario bounds are further investigated by solving the optimization problem for $N = 100$ and $N = 10^3$. Figure 6 shows the resulting bounds $[\bar{\epsilon}(s^\star_N), \bar{\epsilon}(s^\star_N)]$ (red areas) and the true violation probability estimated with $\mathcal{D}_{10^5}$. It can be observed that in correspondence of $\lambda = 0$ we have bounds on the system failure probability. In general, these bound might be slightly different from the one obtained in Section 5.2 since $s^\star_N$ is a random number which depends on the available $\mathcal{D}_N$.

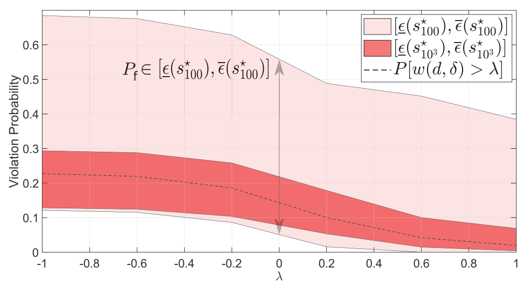

Fig. 6. Prospective-reliability bounds for $N = 100$ and $N = 10^3$ scenarios $\lambda \in [-1, +1]$ and 'true' violation probability $\mathbb{P}[\delta \in \Delta : w(d^\star, \delta) \geq \lambda]$ (dashed line).

6. Discussion and Conclusion

This paper introduces a new scenario program with relaxed constraints to solve RBDO problems given-data. The proposed method can be used to identify a reliable design which minimizes a combination of system cost and violation of reliability requirements without prescribing a model for the uncertainty. The relaxation of constraints can be conveniently used to control the severity of

Table 2. Optimal $d^\star$ for $\lambda = 0$ and increasing N. Comparison between reliability and scenario ϵ-feasibility ($\beta = 10^{-8}$).

N	100	600	900	1500	2000
d_1	0.1973	0.4047	0.4519	0.4617	0.5245
d_2	0.0319	0.0210	0.0162	0.0227	0.0236
$J(d)$	-0.2293	-0.4257	-0.468	-0.4844	-0.5481
P_f	0.17	0.153	0.146	0.142	0.13
$s^\star_N$	18	92	133	214	261
$[\bar{\epsilon}(s^\star_N), \bar{\epsilon}(s^\star_N)]$	[0.016,0.489]	[0.075,0.2634]	[0.082,0.235]	[0.09,0.208]	[0.086,0.185]

Table 3. Scenario bounds on $d^\star$ obtained for $\lambda \in [-1, +1]$ and comparison with the 'true' $V(d^\star)$ estimated using $\mathcal{D}_{10^5}$.

			Scenario optimization (with $\mathcal{D}_{10^3}$)			
λ	-1	-0.6	-0.2	+0.2	+0.6	+1
$d^\star$	[1.35,0.023]	[0.925,0.023]	[0.925,0.023]	[0.36,0.022]	[0.29,0.02]	[0.34,0.015]
$J(d^\star)$	-1.37	-0.95	-0.56	-0.378	-0.31	-0.35
$s^\star_{10^3}$	203	198	172	105	45	24
$[\underline{\epsilon}(s^\star_{10^3}), \bar{\epsilon}(s^\star_{10^3})]$	[0.129,0.294]	[0.124,0.288]	[0.104,0.259]	[0.053,0.179]	[0.015,0.1]	[0.004,0.069]
			Reliability validation using the set $\mathcal{D}_{10^5}$			
$\mathbb{P}[w(d^\star, \delta) > \lambda]$	0.228	0.219	0.186	0.1	0.042	0.02

violation in the failure region and recent results in scenario theory allows computing bounds on the failure probability of the optimized design for any function $w(d, \delta)$ convex in d. These bounds are independent of the underlying data-generating mechanism (hold for any model f_δ consistent with the data) and are non-asymptotic. This is a powerful robustness monitoring capability which clearly reflects the current state of knowledge and helps to avoid overconfidence on the reliability of the optimized design.

Acknowledgements

The authors would like to acknowledge the gracious support of this work through the local authorities under grant agreement "ITEA-2018-17030-Daytime"

References

Beer, M., S. Ferson, and V. Kreinovich (2013). Imprecise probabilities in engineering analyses. *Mechanical Systems and Signal Processing 37*(1-2), 4–29.

Calafiore, G. and M. C. Campi (2006, May). The scenario approach to robust control design. *IEEE Transactions on Automatic Control 51*(5), 742–753.

Campi, M. and S. Garatti (2008). The exact feasibility of randomized solutions of uncertain convex programs. *SIAM Journal on Optimization 19*(3), 1211–1230.

Campi, M. and S. Garatti (2018, Jan). Wait-and-judge scenario optimization. *Mathematical Programming 167*(1), 155–189.

Campi, M. C., S. Garatti, and F. A. Ramponi (2015, Dec). Non-convex scenario optimization with application to system identification. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pp. 4023–4028.

Garatti, S. and M. Campi (2019, Nov). Risk and complexity in scenario optimization. *Mathematical Programming*.

Grooteman, F. (2011). An adaptive directional importance sampling method for structural reliability. *Probabilistic Engineering Mechanics 26*(2), 134 – 141.

Rocchetta, R., L. Crespo, and S. Kenny (2019). Solution of the benchmark control problem by scenario optimization. *Proceedings of the ASME Dynamic Systems and Control Conference, DSCC, October,*.

Rocchetta, R., L. Crespo, and S. Kenny (2020). A scenario optimization approach to reliability-based design. *Reliability Engineering & System Safety 196*, 106755.

Rockafellar, R. and J. Royset (2010). On buffered failure probability in design and optimization of structures. *Reliability Engineering & System Safety 95*(5), 499 – 510.

Özgün Elçi, N. Noyan, and K. Bülbül (2018). Chance-constrained stochastic programming under variable reliability levels with an application to humanitarian relief network design. *Computers & Operations Research 96*, 91 – 107.