OPTIMAL UNCERTAINTY QUANTIFICATION

H. OWHADI, C. SCOVEL, T.J. SULLIVAN, M. McKERNS AND M. ORTIZ

Technical Report No. 2010-03
September 2010

Optimal Uncertainty Quantification

H. Owhadi, T. J. Sullivan, M. McKerns, M. Ortiz
California Institute of Technology
owhadi@caltech.edu

C. Scovel
Los Alamos National Laboratory
jcs@lanl.gov

September 1, 2010

Abstract

We propose a rigorous framework for Uncertainty Quantification (UQ) in which the UQ objectives and the assumptions/information set are brought to the forefront. This framework, which we call *Optimal Uncertainty Quantification* (OUQ), is based on the observation that, given a set of assumptions and information about the problem, there exist optimal bounds on uncertainties: these are obtained as extreme values of well-defined optimization problems corresponding to extremizing probabilities of failure, or of deviations, subject to the constraints imposed by the scenarios compatible with the assumptions and information. In particular, this framework does not implicitly impose inappropriate assumptions, nor does it repudiate relevant information.

Although OUQ optimization problems are extremely large, we show that under general conditions, they have finite-dimensional reductions. As an application, we develop *Optimal Concentration Inequalities* (OCI) of Hoeffding and McDiarmid type. Surprisingly, contrary to the classical sensitivity analysis paradigm, these results show that uncertainties in input parameters do not necessarily propagate to output uncertainties.

In addition, a general algorithmic framework is developed for OUQ and is tested on the Caltech surrogate model for hypervelocity impact, suggesting the feasibility of the framework for important complex systems.

Contents

1 Introduction

1.1 The UQ Problem

In this paper, we provide a well-posed framework for the statement and solution of UQ problems. To make the discussion specific, we describe the framework as it applies to the certification problem. This is a good model problem since we can also formulate many important UQ problems, such as prediction, verification and validation as certification problems. This is similar to the point of view of [4], in which formulations of many problem objectives in reliability are shown to be representable in a unified framework.

Suppose that we are interested in certifying that the response function G of a given physical system will not exceed a given safety threshold a with probability at least $1 - \epsilon$. That is, we wish to certify that

$$\mathbb{P}[G(X) \geq a] \leq \epsilon. \tag{1.1}$$

In practice, the event $[G(X) \geq a]$ may represent the crash of an aircraft, the failure of a weapons system, or the average surface temperature on the Earth being too high. The measure $\mathbb{P}$ stands for the measure of probability associated with the randomness of some of the input variables of G.

If enough (on the order of $\frac{1}{\epsilon^2} \log \frac{1}{\epsilon}$) independent $\mathbb{P}$-distributed samples of $G(X)$ are available, then the certification of (1.1) can be obtained through a Monte Carlo estimate. If G is regular and X is controllable with a known distribution and low dimension, then this number can be reduced to $\frac{1}{\epsilon} \ln \frac{1}{\epsilon}$. If G is the solution of a stochastic partial differential equation with regular coefficients and X is controllable and has a known distribution, then stochastic collocation methods can reduce the number of required samples even further. However, in most real applications, only incomplete information on $\mathbb{P}$ and G is available and the number of available samples on G is small or zero. X may be of high dimension, and may include uncontrollable variables and unknown unknowns. G may not be the solution of a PDE and may involve interactions between singular and complex processes such as (for instance) dislocation, fragmentation, phase transitions, physical phenomena in untested regimes, and even human decisions.

The certification problem exhibits one of the main difficulties that face UQ practitioners: many theoretical methods are available, but they require assumptions or conditions that, oftentimes, are not satisfied by the application. More precisely, the characteristic elements distinguishing these different methods are the assumptions upon which they are based, and some methods will be more efficient than others depending on the validity of those assumptions. UQ applications are also characterized by a set of assumptions/information on the response function G and measure $\mathbb{P}$, which varies from application to application. Hence, on the one hand, we have a list of theoretical methods that are applicable or efficient under very specific assumptions; on the other hand, most applications are characterized by an information set or assumptions that, in general, do not match those required by these theoretical methods.

Recently, it has been shown that many of these, and other, difficulties acknowledged in the recent NRC report [37] can be resolved if the certification problem is

3

well-formulated [28, 47]. Moreover, since these resolutions are quite elementary, we are encouraged by John Dewey's assertion that "a problem well formulated is half solved" to develop a unified framework for the formulation and solution of UQ problems. One of the objectives of this work is the development of such a framework, which we call *Optimal Uncertainty Quantification (OUQ)*.

In the OUQ framework, *information on and assumptions about G and $\mathbb{P}$ are placed at the forefront of the UQ problem*. To that end, let $\mathcal{A}$ denote the set of all *admissible scenarios* (f, μ) for the unknown reality $(G, \mathbb{P})$. It then follows from $(G, \mathbb{P}) \in \mathcal{A}$ that

$$\inf_{(f,\mu)\in\mathcal{A}} \mu[f(X) \geq a] \leq \mathbb{P}[G(X) \geq a] \leq \sup_{(f,\mu)\in\mathcal{A}} \mu[f(X) \geq a]. \tag{1.2}$$

Moreover, it is elementary to observe that
- The quantities on the right-hand and left-hand of (1.2) are values of optimization problems and elements of $[0, 1]$.
- Both the right-hand and left-hand inequalities are optimal in the sense that they are the sharpest bounds for $\mathbb{P}[G(X) \geq a]$ that are consistent with the information and assumptions $\mathcal{A}$.

More importantly, in Proposition 2.1, we show that these two inequalities provide sufficient information to produce an optimal solution to the certification problem.

A well-posed formulation of UQ within the OUQ framework using inequalities such as (1.2) resolves more than the issues raised in the NRC report [37]. Indeed, we can show that that difficulties related to the distinction between aleatoric and epistemic uncertainties disappear, since both are gracefully incorporated in the specification of $\mathcal{A}$. In a similar way, difficulties related to the multiple "faces" of UQ (*e.g.* validation, certification, reduced order modeling, prediction, extrapolation) also disappear, since all of these aspects can be formulated in a natural and unified way.

1.2 Outline of the Paper

Although the OUQ optimization problems (1.2) are extremely large, we show in Section 3 that an important subclass enjoys significant and practical finite-dimensional reduction properties. More precisely, although the optimization variables (f, μ) live in a product space of functions and probability measures, for OUQ problems governed by linear inequality constraints on generalized moments, we demonstrate in Theorem 3.1 and Corollary 3.4 that we can reduce to probability measures that are products of finite convex combinations of Dirac masses with explicit upper bounds on the number of Dirac masses. Moreover, all the results in this paper can be extended to sets of extreme points (extremal measures) more general than Dirac masses, such as those described by Dynkin [13]; we have phrased the results in terms of Dirac masses for simplicity.

Furthermore, when all constraints are generalized moments of functions of f, we can also reduce the space of functions to functions on an m-fold product of finite discrete spaces, and reduce the m-fold products of finite convex combinations of Dirac masses to the products of probability measures on this m-fold product of finite discrete spaces.

This latter reduction, presented in Theorem 3.6, completely eliminates dependency on the coordinate positions of the Dirac masses.

In Section 4, the reduction techniques of Section 3 are applied to obtain optimal Mc-Diarmid and Hoeffding inequalities, *i.e.*, optimal concentration-of-measure inequalities with the assumptions of McDiarmid's inequality [31] or Hoeffding's inequality [18]. In particular, in Theorems 4.1, 4.2 and 4.4, we obtain analytic solutions to the McDiarmid problem for dimension $m = 1, 2, 3$, and in Proposition 4.6, we obtain an asymptotic formula for general m, thereby obtaining an optimal McDiarmid inequality in these cases. In Theorems 4.9 and 4.11, we provide analytic solutions under Hoeffding's assumptions. Moreover, all of these solutions provide important information on the effects of the diameter parameters (see Equation (3.4)). For example, for $m = 2$, if D_2 is sufficiently smaller than D_1, then the optimal bound only depends on D_1, and therefore, any decrease in D_2 does not improve the inequality. Hence, contrary to the standard sensitivity analysis paradigm, uncertainties in input parameters do not necessarily propagate to output uncertainties.

The OUQ framework allows the development of an OUQ loop that can be used for experimental design. In Section 2, we show that the problem of predicting optimal bounds on the results of experiments under the assumption that the system is safe (or unsafe) is well-posed and benefits from similar reduction properties as the certification problem. Best experiments are then naturally identified as those whose predicted ranges have minimal overlap between safe and unsafe systems.

In Section 5, we give examples of numerical OUQ calculations for a surrogate hyper-velocity impact model; these calculations are made possible by the reduction theorems of Section 3. We observe that, even if the numerical optimization is allowed to search over a larger search space than the reduction theorems show is sufficient, the reduced-order extrema are attractors. This suggests that adequate numerical implementation of OUQ problems can detect and use "hidden" reduction properties even in the absence of theorems proving them to be true. Based on these observations, we propose an OUQ optimization algorithm for arbitrary constraints based on a coagulation/fragmentation of probability distributions, leveraging possibly hidden reduction properties.

2   Optimal Uncertainty Quantification

In this section, we describe more formally the *Optimal Uncertainty Quantification* framework. In particular, we describe what it means to give optimal bounds on the probability of failure in (1.1) given information/assumptions about the system of interest, and hence how to rigorously certify or de-certify that system.

For the sake of clarity, we will start the description of OUQ with deterministic information and assumptions (when $\mathcal{A}$ is a deterministic set of functions and probability measures). We refer to Section 7 for an initial description of OUQ with sample data.

2.1 First Description

We propose to reverse the standard UQ strategy of developing methods with their attendant implicit and explicit assumptions, which may or may not match the relevant assumptions for the application. In particular, we begin by placing information and assumptions at the core of UQ. As noted by Hoeffding [17], assumptions about the system of interest play a central and sensitive role in any statistical decision problem, even though the assumptions are often only approximations of reality.

A simple example of an information/assumptions set is given by constraining the mean and range of the response function. For example, let $\mathcal{M}(\mathcal{X})$ be the set of probability measures on the set $\mathcal{X}$, and let $\mathcal{A}_1$ be the set of pairs of probability measures μ and real-valued measurable functions f such that the mean value of f with respect to μ is b and the diameter of the range of f is at most D;

$$\mathcal{A}_1 := \left\{ (f, \mu) \,\middle|\, \begin{array}{l} f \colon \mathcal{X} \to \mathbb{R}, \\ \mu \in \mathcal{M}(\mathcal{X}), \\ \mathbb{E}_\mu[f] = b, \\ (\sup f - \inf f) \leq D \end{array} \right\}. \tag{2.1}$$

Let us assume that $(G, \mathbb{P}) \in \mathcal{A}_1$. Then, any other pair $(f, \mu) \in \mathcal{A}_1$ constitutes an admissible scenario representing a valid possibility for the "reality" $(G, \mathbb{P})$. If asked to bound $\mathbb{P}[G(X) \geq a]$, should we apply different methods and obtain different bounds on $\mathbb{P}[G(X) \geq a]$? Since some methods will distort this information set and others are only using part of it, we instead let the set $\mathcal{A}_1$ be viewed as a feasible set for an optimization problem.

The general OUQ framework. In the general case, we regard the response function G as an unknown measurable function, with some possibly known characteristics, from one measurable space $\mathcal{X}$ of inputs to a second measurable space $\mathcal{Y}$ of values. The input variables are generated randomly according to an unknown random variable X with values in $\mathcal{X}$ according to a law $\mathbb{P} \in \mathcal{M}(\mathcal{X})$, also with some possibly known characteristics. We let a measurable subset $\mathcal{Y}_0 \subseteq \mathcal{Y}$ define the *failure region*; in the example given above, $\mathcal{Y} = \mathbb{R}$ and $\mathcal{Y}_0 = [a, +\infty)$. When there is no danger of confusion, we shall simply write $[G \text{ fails}]$ for the event $[G(X) \in \mathcal{Y}_0]$.

Let $\epsilon \in [0, 1]$ denote the *greatest acceptable probability of failure*. We say that the system is *safe* if $\mathbb{P}[G \text{ fails}] \leq \epsilon$ and the system is *unsafe* if $\mathbb{P}[G \text{ fails}] > \epsilon$. By *information*, or a *set of assumptions*, we mean a subset

$$\mathcal{A} \subseteq \left\{ (f, \mu) \,\middle|\, \begin{array}{l} f \colon \mathcal{X} \to \mathcal{Y} \text{ is measurable} \\ \mu \in \mathcal{M}(\mathcal{X}) \end{array} \right\} \tag{2.2}$$

that contains, at the least, $(G, \mathbb{P})$. The set $\mathcal{A}$ encodes all the information that we have about the real system $(G, \mathbb{P})$, information that may come from known physical laws, past experimental data, and expert opinion. In the example $\mathcal{A}_1$ above, the only information that we have is that the mean performance of the system is b and that the diameter of

6

its range is at most D; any pair (f, μ) that satisfies these two criteria is an *admissible scenario* for the unknown reality $(G, \mathbb{P})$. Since some admissible scenarios may be safe (*i.e.* have $\mu[f \text{ fails}] \leq \epsilon$) whereas other admissible scenarios may be unsafe (*i.e.* have $\mu[f \text{ fails}] > \epsilon$), we decompose $\mathcal{A}$ into the disjoint union $\mathcal{A} = \mathcal{A}_{\text{safe},\epsilon} \uplus \mathcal{A}_{\text{unsafe},\epsilon}$, where

$$\mathcal{A}_{\text{safe},\epsilon} := \{(f, \mu) \in \mathcal{A} \mid \mu[f \text{ fails}] \leq \epsilon\}, \tag{2.3a}$$

$$\mathcal{A}_{\text{unsafe},\epsilon} := \{(f, \mu) \in \mathcal{A} \mid \mu[f \text{ fails}] > \epsilon\}. \tag{2.3b}$$

Now observe that, given such an information/assumptions set $\mathcal{A}$, there exist optimal upper and lower bounds on $\mathbb{P}[G(X) \geq a]$ corresponding to the scenarios compatible with assumptions, *i.e.* the lower $\mathcal{L}(\mathcal{A})$ and upper $\mathcal{U}(\mathcal{A})$ values of the naturally defined optimization problems:

$$\mathcal{L}(\mathcal{A}) := \inf_{(f,\mu) \in \mathcal{A}} \mu[f \text{ fails}] \tag{2.4a}$$

$$\mathcal{U}(\mathcal{A}) := \sup_{(f,\mu) \in \mathcal{A}} \mu[f \text{ fails}]. \tag{2.4b}$$

Since $\mathcal{L}(\mathcal{A})$ and $\mathcal{U}(\mathcal{A})$ are well-defined in $[0, 1]$, and approximations are sufficient for most purposes and are necessary in general, the difference between sup and max should not be much of an issue. Of course, some of the work that follows is concerned with the attainment of maximizers, and whether those maximizers have any simple structure that can be exploited for the sake of computational efficiency. Now, since $(G, \mathbb{P}) \in \mathcal{A}$, it follows that

$$\mathcal{L}(\mathcal{A}) \leq \mathbb{P}[G \text{ fails}] \leq \mathcal{U}(\mathcal{A}).$$

Moreover, the upper bound $\mathcal{U}(\mathcal{A})$ is *optimal* in the sense that

$$\mu[f \text{ fails}] \leq \mathcal{U}(\mathcal{A}) \text{ for all } (f, \mu) \in \mathcal{A}$$

and, if $U' < \mathcal{U}(\mathcal{A})$, then there is an admissible scenario $(f, \mu) \in \mathcal{A}$ such that

$$U' < \mu[f \text{ fails}] \leq \mathcal{U}(\mathcal{A}).$$

That is, although $\mathbb{P}[G \text{ fails}]$ may be much smaller than $\mathcal{U}(\mathcal{A})$, there is a pair (f, μ) which satisfies the same assumptions as $(G, \mathbb{P})$ such that $\mu[f \text{ fails}]$ is approximately equal to $\mathcal{U}(\mathcal{A})$. Similar remarks apply for the lower bound $\mathcal{L}(\mathcal{A})$.

Moreover, the values $\mathcal{L}(\mathcal{A})$ and $\mathcal{U}(\mathcal{A})$, defined in (2.4) can be used to construct a solution to the certification problem. Let the certification problem be defined by an error function that gives an error whenever 1) the certification process produces "safe" and there exists an admissible point which is unsafe, 2) the certification process produces "unsafe" and there exists an admissible point which is safe, or 3) the certification process produces "cannot decide" and all admissible points are safe or all admissible points are unsafe. Otherwise, the certification process produces no error. The following proposition demonstrates that, except in the special case $\mathcal{L}(\mathcal{A}) = \epsilon$, that these values determine an optimal solution to this certification problem.

7

	$\mathcal{L}(\mathcal{A}) := \inf\limits_{(f,\mu)\in\mathcal{A}} \mu\big[f(X) \geq a\big]$	$\mathcal{U}(\mathcal{A}) := \sup\limits_{(f,\mu)\in\mathcal{A}} \mu\big[f(X) \geq a\big]$
$\leq \epsilon$	**Cannot decide** Insufficient Information	**Certify** **Safe** even in the Worst Case
$> \epsilon$	**De-certify** **Unsafe** even in the Best Case	**Cannot decide** Insufficient Information

Table 2.1: The OUQ certification process provides a rigorous certification criteria whose outcomes are of three types: "Certify", "De-certify" and "Cannot decide".

Proposition 2.1. *If* $(G, \mathbb{P}) \in \mathcal{A}$ *and*

- $\mathcal{U}(\mathcal{A}) \leq \epsilon$ *then* $\mathbb{P}[G \text{ fails}] \leq \epsilon$.

- $\epsilon < \mathcal{L}(\mathcal{A})$ *then* $\mathbb{P}[G \text{ fails}] > \epsilon$.

- $\mathcal{L}(\mathcal{A}) < \epsilon < \mathcal{U}(\mathcal{A})$ *the there exists* $(f_1, \mu_1) \in \mathcal{A}$ *and* $(f_2, \mu_2) \in \mathcal{A}$ *such that* $\mu_1[f_1 \text{ fails}] < \epsilon < \mu_2[f_2 \text{ fails}]$.

In other words, provided that our information set $\mathcal{A}$ is valid (in the sense that $(G, \mathbb{P}) \in \mathcal{A}$) then if $\mathcal{U}(\mathcal{A}) \leq \epsilon$, then, the system is provably safe; if $\epsilon < \mathcal{L}(\mathcal{A})$, then the system is provably unsafe; and if $\mathcal{L}(\mathcal{A}) < \epsilon < \mathcal{U}(\mathcal{A})$, then the safety of the system cannot be decided due to lack of information. The corresponding certification process and its optimality are represented in Table 2.1. Hence, solving the optimization problems (2.4) determines an optimal solution to the certification problem, under the condition that $\mathcal{L}(\mathcal{A}) \neq \epsilon$. When $\mathcal{L}(\mathcal{A}) = \epsilon$ we can still produce an optimal solution if we obtain further information. That is, when $\mathcal{L}(\mathcal{A}) = \epsilon = \mathcal{U}(\mathcal{A})$, then the optimal process produces "safe." On the other hand, when $\mathcal{L}(\mathcal{A}) = \epsilon < \mathcal{U}(\mathcal{A})$, the optimal solution depends on whether or not there exists a minimizer $(f, \mu) \in \mathcal{A}$ such that $\mu[f \text{ fails}] = \mathcal{L}(\mathcal{A})$; if so, the optimal process should declare "cannot decide," otherwise, the optimal process should declare "unsafe." Observe that, in Table 2.1, we have classified $\mathcal{L}(\mathcal{A}) = \epsilon < \mathcal{U}(\mathcal{A})$ as "cannot decide." This "nearly optimal" solution appears natural and conservative without the knowledge of the existence or non-existence of optimizers.

Example 2.1. The bounds $\mathcal{L}(\mathcal{A})$ and $\mathcal{U}(\mathcal{A})$ can be computed exactly — and are non-trivial — in the case of the simple example $\mathcal{A}_1$ given in (2.1). Indeed, the optimal upper bound is given by

$$\mathcal{U}(\mathcal{A}_1) = p_{\max} := \left(1 - \frac{(b-a)_+}{D}\right)_+, \tag{2.5}$$

where the maximum is achieved at the weighted sum of two weighted Dirac delta masses

$$p_{\max}\delta_a + (1 - p_{\max})\delta_{a-D}.$$

This simple example demonstrates an extremely important point: even if the function G is extremely expensive to evaluate, certification can be accomplished without recourse to the expensive evaluations of G. Furthermore, the simple structure of the maximizers motivates the reduction theorems later in Section 3.

Example 2.2. Concentration-of-measure inequalities can be used to obtain upper bounds on $\mathcal{U}(\mathcal{A})$ and lower bounds on $\mathcal{L}(\mathcal{A})$; in that sense, they lead to sub-optimal methods. Indeed, consider the following information/assumptions $\mathcal{A}_2$: the space $\mathcal{X}$ is a product space with $X = (X_1, X_2, \ldots, X_m)$, the mean performance satisfies $\mathbb{E}[G(X)] \leq b$, the m inputs $X_1, \ldots, X_m$ are independent, and

$$\sum_{i=1}^{m} \sup \left\{ \left| G(x) - G(x') \right|^2 \,\Big|\, x, x' \in \mathcal{X}, x_j = x'_j \text{ for all } j \neq i \right\} \leq D^2.$$

It follows from McDiarmid's inequality [31, 32] that, under the assumptions $\mathcal{A}_2$,

$$\mathcal{U}(\mathcal{A}_2) \leq \exp\left(-2 \frac{(a-b)_+^2}{D^2} \right).$$

This example shows how existing techniques such as concentration-of-measure inequalities can be incorporated into OUQ. In Section 3, we will show how to reduce $\mathcal{U}(\mathcal{A}_2)$ to a finite dimensional optimization problem. Based on this reduction, in Section 4, we provide analytic solutions to the optimization problem $\mathcal{U}(\mathcal{A}_2)$ for $m = 1, 2, 3$.

2.2 The Optimal UQ Loop

In the previous subsection we discussed how the basic inequality

$$\mathcal{L}(\mathcal{A}) \leq \mathbb{P}[G \geq a] \leq \mathcal{U}(\mathcal{A})$$

provides rigorous optimal certification criteria. The certification process should not be confused with its three possible outcomes (see Table 2.1) which we call "certify" (we assert that the system is safe), "de-certify" (we assert that the system is unsafe) and "cannot decide" (the safety or un-safety of the system is undecidable given the information/assumption set $\mathcal{A}$). Indeed, in the case

$$\mathcal{L}(\mathcal{A}) \leq \epsilon < \mathcal{U}(\mathcal{A})$$

there exist admissible scenarios under which the system is safe, and other admissible scenarios under which it is unsafe. Consequently, it follows that we can make no definite certification statement for $(G, \mathbb{P})$ without introducing further information/assumptions. If no further information can be obtained, we conclude that we "cannot decide;" (this state could also be called "do not decide," because we could (arbitrarily) decide that the system is unsafe due to lack of information, for instance, but do not).

However, if sufficient resources exist to gather additional information, then we enter what may be called the *optimal uncertainty quantification loop*, illustrated in Figure 2.1.

9
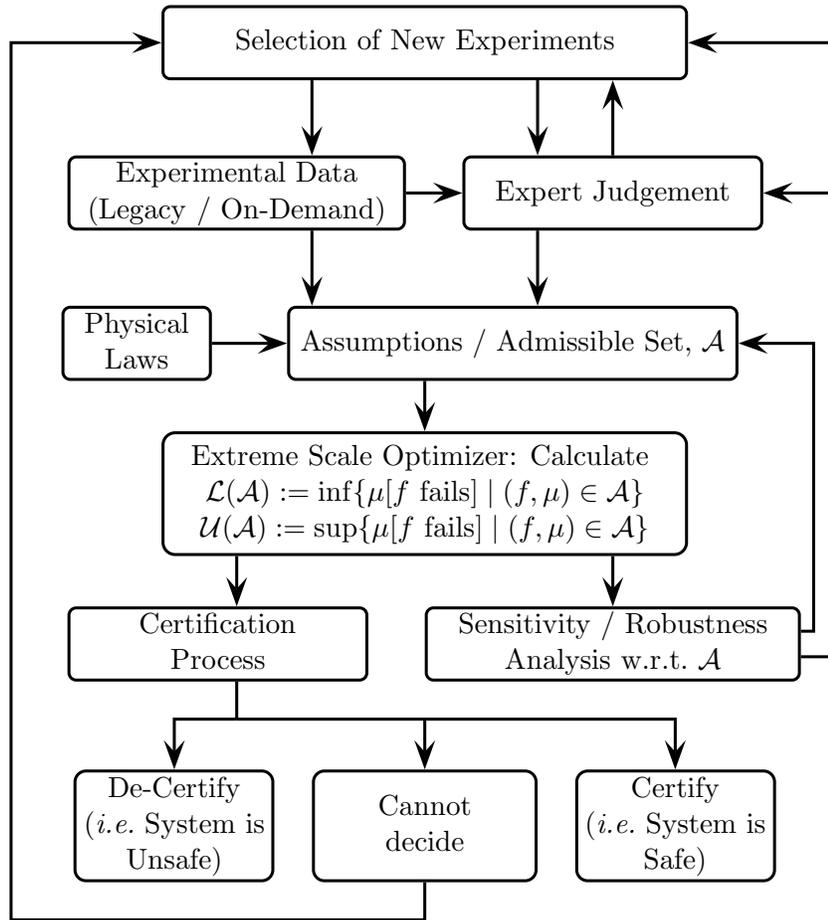
Figure 2.1: Optimal Uncertainty Quantification Loop.

The admissible set $\mathcal{A}$ draws on three principal sources of information: known physical laws, expert opinion, and experimental data. Once the set $\mathcal{A}$ has been constructed, the calculation of the lower and upper bounds $\mathcal{L}(\mathcal{A})$ and $\mathcal{U}(\mathcal{A})$ is a well-posed optimization problem. If the results of this optimization problem lead to certification or de-certification, then we are done; if not, then new experiments should be designed and expert opinion sought in order to validate or invalidate the extremal scenarios that cause the inequality

$$\mathcal{L}(\mathcal{A}) \leq \epsilon < \mathcal{U}(\mathcal{A})$$

to hold. The addition of information means further constraints on the collection of admissible scenarios; that is, the original admissible set $\mathcal{A}$ is reduced to a smaller one $\mathcal{A}' \subset \mathcal{A}$, thereby providing sharper bounds on the probability of failure:

$$\mathcal{L}(\mathcal{A}) \leq \mathcal{L}(\mathcal{A}') \leq \mathbb{P}[G(X) \geq a] \leq \mathcal{U}(\mathcal{A}') \leq \mathcal{U}(\mathcal{A}).$$

The sharper bounds may meet the certification criteria of Table 2.1. If not, and there are still resources for gathering additional information, then the loop should be repeated. This process is the feedback arrow on the left-hand side of Figure 2.1.

The right-hand side of Figure 2.1 constitutes another aspect of the OUQ loop. The bounds $\mathcal{L}(\mathcal{A})$ and $\mathcal{U}(\mathcal{A})$ are only useful insofar as the assumptions $\mathcal{A}$ are accurate. It is possible that the sources of information that informed $\mathcal{A}$ may have been in error: physical laws may have been extended outside their range of validity (*e.g.* Newtonian physics may have been applied in the relativistic regime), there may have been difficulties with the experiments or the results misinterpreted, or expert opinion may have been erroneous. Therefore, a vital part of OUQ is to examine the sensitivity and robustness of the bounds $\mathcal{L}(\mathcal{A})$ and $\mathcal{U}(\mathcal{A})$ with respect to the assumption set $\mathcal{A}$. If the bounds $\mathcal{L}(\mathcal{A})$ and $\mathcal{U}(\mathcal{A})$ are found to depend sensitively on one particular assumption (say, the mean performance of one component of the system), then it would be advisable to expend resources investigating this assumption.

Experimental selection and design An important aspect of the OUQ loop is the selection of new experiments. Suppose that a number of possible experiments E_i are proposed, each of which will determine some functional $\Phi_i(G, \mathbb{P})$ of G and $\mathbb{P}$. For example, $\Phi_1(G, \mathbb{P})$ could be $\mathbb{E}_{\mathbb{P}}[G]$, $\Phi_2(G, \mathbb{P})$ could be $\mathbb{P}[X \in A]$ for some subset $A \subseteq \mathcal{X}$ of the input parameter space, and so on. Suppose that there are insufficient experimental resources to run all of these proposed experiments. Let us now consider which experiment should be run for the certification problem. Recall that the admissible set $\mathcal{A}$ is partitioned into safe and unsafe subsets as in (2.3). Define $J_{\text{safe},\epsilon}(\Phi_i)$ to be the closed interval spanned by the possible values for the functional Φ_i over the safe admissible

scenarios (*i.e.* the closed convex hull of the range of Φ_i on $\mathcal{A}_{\text{safe},\epsilon}$): that is, let

$$J_{\text{safe},\epsilon}(\Phi_i) := \left[\inf_{(f,\mu)\in\mathcal{A}_{\text{safe},\epsilon}} \Phi_i(f,\mu), \sup_{(f,\mu)\in\mathcal{A}_{\text{safe},\epsilon}} \Phi_i(f,\mu) \right] \quad (2.6a)$$

$$J_{\text{unsafe},\epsilon}(\Phi_i) := \left[\inf_{(f,\mu)\in\mathcal{A}_{\text{unsafe},\epsilon}} \Phi_i(f,\mu), \sup_{(f,\mu)\in\mathcal{A}_{\text{unsafe},\epsilon}} \Phi_i(f,\mu) \right]. \quad (2.6b)$$

Note that, in general, these two intervals may be disjoint or may have non-empty intersection; the size of their intersection provides a measure of usefulness of the proposed experiment E_i. Observe that if experiment E_i were run, yielding the values $\Phi_i(G,\mathbb{P})$, then the following conclusions could be drawn:

$$\Phi_i(G,\mathbb{P}) \in J_{\text{safe},\epsilon}(\Phi_i) \cap J_{\text{unsafe},\epsilon}(\Phi_i) \implies \text{no conclusion,}$$
$$\Phi_i(G,\mathbb{P}) \in J_{\text{safe},\epsilon}(\Phi_i) \setminus J_{\text{unsafe},\epsilon}(\Phi_i) \implies \text{the system is safe,}$$
$$\Phi_i(G,\mathbb{P}) \in J_{\text{unsafe},\epsilon}(\Phi_i) \setminus J_{\text{safe},\epsilon}(\Phi_i) \implies \text{the system is unsafe,}$$
$$\Phi_i(G,\mathbb{P}) \notin J_{\text{safe},\epsilon}(\Phi_i) \cup J_{\text{unsafe},\epsilon}(\Phi_i) \implies \text{faulty assumptions,}$$

where the last assertion (*faulty assumptions*) means that $(G,\mathbb{P}) \notin \mathcal{A}$ and follows from the fact that $\Phi_i(G,\mathbb{P}) \notin J_{\text{safe},\epsilon}(\Phi_i) \cup J_{\text{unsafe},\epsilon}(\Phi_i)$ is a contradiction. The validity of the first three assertions is based on the supposition that $(G,\mathbb{P}) \in \mathcal{A}$.

In this way, the computational optimization exercise of finding $J_{\text{safe},\epsilon}(\Phi_i)$ and $J_{\text{unsafe},\epsilon}(\Phi_i)$ for each proposed experiment E_i provides an objective assessment of which experiments are worth performing: those for which $J_{\text{safe},\epsilon}(\Phi_i)$ and $J_{\text{unsafe},\epsilon}(\Phi_i)$ are nearly disjoint intervals are worth performing since they are likely to yield conclusive results *vis-à-vis* (de-)certification and conversely, if the intervals $J_{\text{safe},\epsilon}(\Phi_i)$ and $J_{\text{unsafe},\epsilon}(\Phi_i)$ have a large overlap, then experiment E_i is not worth performing since it is unlikely to yield conclusive results. Furthermore, the fourth possibility above shows how experiments can rigorously establish that one's assumptions $\mathcal{A}$ are incorrect. See Figure 2.2 for an illustration.

Remark 2.2. For the sake of clarity, we have started this description by defining *experiments* as *functionals* Φ_i of $\mathbb{P}$ and G. In practice, some experiments may not be functionals of $\mathbb{P}$ and G but of related objects. Consider, for instance, the situation where (X_1, X_2) is a two-dimensional Gaussian vector with zero mean and covariance matrix C, $\mathbb{P}$ is the probability distribution of X_1, the experiment E_2 determines the variance of X_2 and the information set $\mathcal{A}$ is $C \in B$, where B is a subset of symmetric positive definite 2×2 matrices. The outcome of the experiment E_2 is not a function of the probability distribution $\mathbb{P}$; however, the knowledge of $\mathbb{P}$ restricts the range of possible outcomes of E_2. Hence, for some experiments E_i, the knowledge of $(G,\mathbb{P})$ does not determine the outcome of the experiment, but only the set of possible outcomes. For those experiments, the description given above can be generalized to situations where Φ_i is a *multivalued* functional of $(G,\mathbb{P})$ determining the set of possible outcomes of the experiment E_i. This picture can be generalized further by introducing measurement noise, in which case $(G,\mathbb{P})$ may not determine a deterministic set of possible outcomes, but instead a measure of probability on a set of possible outcomes.
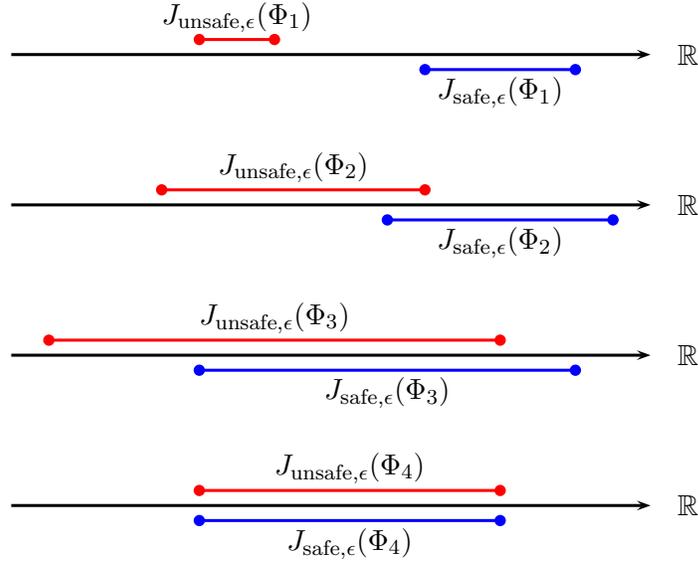
Figure 2.2: A schematic representation of the intervals $J_{\mathrm{unsafe},\epsilon}(\Phi_i)$ (in red) and $J_{\mathrm{safe},\epsilon}(\Phi_i)$ (in blue) as defined by (2.6) for four functionals Φ_i that might be the subject of an experiment. Φ_1 is a good candidate for experiment effort, since the intervals do not overlap and hence experimental determination of $\Phi_1(G, \mathbb{P})$ will certify or de-certify the system; Φ_4 is not worth investigating, since it cannot distinguish safe scenarios from unsafe ones; Φ_2 and Φ_3 are intermediate cases, with Φ_2 a better prospect than Φ_3.

3   Reduction of OUQ Optimization Problems

In general, the lower and upper values

$$\mathcal{L}(\mathcal{A}) \quad := \quad \inf_{(f,\mu)\in\mathcal{A}} \mu[f(X) \geq a]$$
$$\mathcal{U}(\mathcal{A}) \quad := \quad \sup_{(f,\mu)\in\mathcal{A}} \mu[f(X) \geq a]$$

are each defined by a non-convex and infinite-dimensional optimization problem, the solution of which poses significant computational challenges. These optimization problems can be considered to be a generalization of Chebyshev inequalities. The history of the classical inequalities can be found in [24], and some generalizations in [8] and [52]; in the latter works, the connection between Chebyshev inequalities and optimization theory is developed based on the work of Mulholland and Rogers [35], Godwin [15], Isii [19, 20, 21], Olkin and Pratt [36], Marshall and Olkin [29], and the classical Markov–Krein Theorem [24, pages 82 & 157], among others. The Chebyshev-type inequalities defined by $\mathcal{L}(\mathcal{A})$ and $\mathcal{U}(\mathcal{A})$ are a further generalization to independence assumptions, more general domains, more general systems of moments, and the inclusion of classes of functions, in addition to the probability measures, in the optimization problem. Moreover, although our goal is the computation of these values, and not an analytic expression for them, the study of probability inequalities should be useful in the reduction and approximation of these values. Without providing a survey of this large body of work, we mention the field of majorization, as discussed in Marshall and Olkin [30], the inequalities of Anderson [2], Hoeffding [16], Joe [22], Bentkus et al. [7], Bentkus [5, 6], Pinelis [40, 41], and Boucheron, Lugosi and Massart [9]. Moreover, the solution of the resulting nonconvex optimization problems should benefit from duality theories for nonconvex optimization problems such as Rockafellar [44] and the development of convex envelopes for them, as can be found, for example, in Rikun [43] and Sherali [48]. Finally, since Pardalos and Vavasis [38] show that quadratic programming with one negative eigenvalue is NP-hard, we expect that some OUQ problems may be difficult to solve.

Let us now return to the earlier simple example of an admissible set $\mathcal{A}_1$ in (2.1): the (non-unique) extremizers of the OUQ problem with the admissible set $\mathcal{A}_1$ all have the property that the support of the push-forward measure $f_*\mu$ on $\mathbb{R}$ contains at most two points, i.e., $f_*\mu$ is a convex combination of at most two Dirac delta measures:

$$\sup_{(f,\mu)\in\mathcal{A}_1} \mu[f(X) \geq a] = \sup_{\substack{(f,\mu)\in\mathcal{A}_1 \\ \#\mathrm{supp}(f_*\mu)\leq 2}} \mu[f(X) \geq a].$$

The optimization problem on the left-hand side is an infinite-dimensional one, whereas the optimization problem on the right-hand side is amenable to finite-dimensional parametrization for each f. The aim of this section is to show that a large class of OUQ problems — those governed by independence and linear inequality constraints on the moments, — are amenable to a similar finite-dimensional reduction, and that *a priori* upper bounds can be given on the number of Dirac delta masses that the reduction requires.

14

To begin with, we first show that an important class of optimization problems over the space of m-fold product measures can be reduced to optimization over products of finite convex combinations of Dirac masses. Consequently, we then show in Corollary 3.4 that OUQ optimization problems where the admissible set is defined as a subset of function-measure pairs (f, μ) that satisfy generalized moment constraints $\mathbf{G}_f(\mu) \leq 0$ can also be reduced from the space of measures to the products of finite convex combinations of Dirac masses. Theorem 3.6 shows that, when all the constraints are generalized moments of functions of f, the search space $\mathcal{G}$ of functions can be further reduced to a search over functions on an m-fold product of finite discrete spaces, and the search over m-fold products of finite convex combinations of Dirac masses can be reduced to a search over the products of probability measures on this m-fold product of finite discrete spaces. This latter reduction completely eliminates dependency on the coordinate positions in $\mathcal{X}$. Theorem 3.6 is then used in Proposition 3.7 to obtain an optimal McDiarmid inequality through the formulation of an appropriate OUQ optimization problem followed by the above-mentioned reductions to an optimization problem on the product of functions on $\{0, 1\}^m$ with the m-fold products of measures on $\{0, 1\}^m$. This problem is then further reduced, by Theorem 3.8, to an optimization problem on the product of the space of subsets of $\{0, 1\}^m$ with the product measures on $\{0, 1\}^m$. Finally, we obtain analytic solutions to this last problem for $m = 1, 2, 3$, thereby obtaining an optimal McDiarmid inequality in these cases. We also obtain an asymptotic formula for general m. Moreover, the solution for $m = 2$ indicates important information regarding the diameter parameters D_1 and D_2. For example, if D_2 is sufficiently smaller than D_1, then the optimal bound only depends on D_1 and therefore, any decrease in D_2 does not improve the inequality.

3.1 Reduction of OUQ

For a topological space $\mathcal{X}$, let $\mathcal{F}_{\mathcal{X}}$ (or simply $\mathcal{F}$) denote the space of real-valued measurable functions on $\mathcal{X}$, and let $\mathcal{M}(\mathcal{X})$ denote the set of Borel probability measures on $\mathcal{X}$. Denote the process of integration with respect to a measure μ by $\mathbb{E}_\mu$, and let

$$\Delta_k(\mathcal{X}) := \left\{ \sum_{j=0}^{k} \alpha_j \delta_{x^j} \,\middle|\, x^j \in \mathcal{X}, \alpha_j \geq 0 \text{ for } j = 0, \ldots, k \text{ and } \sum_{j=0}^{k} \alpha_j = 1 \right\}$$

denote the set of $(k+1)$-fold convex combinations of Dirac masses. When $\mathcal{X} = \prod_{i=1}^{m} \mathcal{X}_i$ is a product of topological spaces, and we speak of measurable functions on the product $\mathcal{X}$, we mean measurable with respect to the product σ-algebra and not the Borel σ-algebra of the product. For more discussion of this delicate topic, see *e.g.* [23]. The linear equality and inequality constraints on our optimization problems will be encoded in the following measurable functions:

$$g'_j \colon \mathcal{X} \to \mathbb{R} \text{ for } j = 1, \ldots, n',$$

and, for each $i = 1, \ldots, m$,

$$g^i_j \colon \mathcal{X}_i \to \mathbb{R} \text{ for } j = 1, \ldots, n_i.$$

15

Let $\mathcal{M}^{\mathbf{G}} \subseteq \mathcal{M}^m(\mathcal{X})$ denote the set of products of Borel measures for which these all these functions are integrable with finite integrals. We use the compact notation $\mathbf{G}(\mu) \leq 0$ to indicate that $\mu \in \mathcal{M}^{\mathbf{G}}$ and that

$$\mathbb{E}_\mu[g_j'] \leq 0, \quad \text{for all } j = 1, \ldots, n',$$
$$\mathbb{E}_\mu[g_j^i] \leq 0, \quad j = 1, \ldots, n_i, \text{ for all } i = 1, \ldots, m.$$

Moreover, let $r \colon \mathcal{X} \to \mathbb{R}$ be integrable for all $\mu \in \mathcal{M}^{\mathbf{G}}$ (possibly with values $+\infty$ or $-\infty$). For any set $\mathcal{M} \subseteq \mathcal{M}^{\mathbf{G}}$, let

$$\mathcal{U}(\mathcal{M}) := \sup_{\mu \in \mathcal{M}} \mathbb{E}_\mu[r],$$

with the convention that the supremum of the empty set is $-\infty$.

For a measurable function f, the map $\mu \mapsto \mathbb{E}_\mu[f]$ may not be defined, since f may not be absolutely integrable with respect to μ. If it is defined, then it is continuous in the strong topology on measures; however, this topology is too strong to provide any compactness. Moreover, although [1, Theorem 14.5] shows that if f is a bounded upper semi-continuous function on a metric space, then integration is upper semi-continous in the weak-$*$ topology, we consider the case in which $\mathcal{X}$ may not be metric or compact, and the functions f may be unbounded and lack continuity properties. The following results heavily use results of Winkler [54] — which follow from an extension of Choquet Theory (see *e.g.* [39]) by von Weizsäcker and Winkler [53, Corollary 3] to sets of probability measures with generalized moment constraints — and a result of Kendall [26] characterizing cones, which are lattice cones in their own order. These results generalize a result of Karr [25] that requires $\mathcal{X}$ to be compact, the constraint functions be bounded and continuous, and the constraints to be equalities. The results that follow are remarkable in that they make extremely weak assumptions on $\mathcal{X}$ and no assumptions on the functions f. Recall that a Suslin space is the continuous image of a Polish space.

Theorem 3.1. *Let $\mathcal{X} = \prod_{i=1}^m \mathcal{X}_i$ be a product of Suslin spaces and let*

$$\mathcal{M}^m(\mathcal{X}) := \bigotimes_{i=1}^m \mathcal{M}(\mathcal{X}_i)$$

denote the set of products of Borel probability measures on the spaces $\mathcal{X}_i$. As above, consider the generalized moment functions $\mathbf{G}$ and the corresponding finite moment set $\mathcal{M}^{\mathbf{G}}$. Suppose that $r \colon \mathcal{X} \to \mathbb{R}$ is integrable for all $\mu \in \mathcal{M}^{\mathbf{G}}$ (possibly with values $+\infty$ or $-\infty$). Define the reduced admissible set

$$\mathcal{M}_\Delta := \left\{ \mu \in \bigotimes_{i=1}^m \Delta_{n_i+n'}(\mathcal{X}_i) \,\middle|\, \mathbf{G}(\mu) \leq 0 \right\}.$$

Then, it holds that

$$\mathcal{U}(\mathcal{M}^{\mathbf{G}}) = \mathcal{U}(\mathcal{M}_\Delta).$$

Theorem 3.1 says that, on a product $\mathcal{X}$ of very general spaces $\mathcal{X}_i$, optimization problems constrained by n' linear moment constraints on $\mathcal{X}$ and n_i linear moment constraints on each factor space $\mathcal{X}_i$ achieve their optima among those product measures whose i^{th} marginal has support on at most $n' + n_i + 1$ points of $\mathcal{X}_i$.

Remark 3.2. Using [53, Corollary 3], this theorem and its consequences below easily generalize from the situation where $g_k(\mu) \leq 0$ for each k to that in which $g_k(\mu) \in I_k$ for each k, where k indexes the constraint functions, and where each I_k is a closed interval. Consequently, such pairs of linear constraints introduce a requirement for only one Dirac mass, not the two masses that one might expect. Moreover, observe that the condition that the function r is integrable (possibly with values $+\infty$ or $-\infty$) for all $\mu \in \mathcal{M}^{\mathbf{G}}$ is satisfied if r is non-negative. In particular, this holds when r is an indicator function of a set, which is our main application in this paper.

Remark 3.3. Theorem 3.1 and its consequents below can be expressed more generally in terms of extreme points of sets of measures, whereas in the above case, the extreme points are the Dirac masses. To that end, Dynkin [13] describes more general sets of measures and their extreme points, which can be useful in applications. In particular, one could consider

1. sets of measures that are invariant under a transformation (the extreme points are the ergodic measures);
2. symmetric measures on an infinite product space (the extreme points are the simple product measures);
3. the set of stationary distributions for a given Markov transition function;
4. the set of all Markov processes with a given transition function.

We now apply Theorem 3.1 to obtain the same type of reduction for an admissible set $\mathcal{A} \subseteq \mathcal{F} \times \mathcal{M}^m(\mathcal{X})$ consisting of pairs of functions and product measures — this is the case for the OUQ optimization problems $\mathcal{L}(\mathcal{A})$ and $\mathcal{U}(\mathcal{A})$. Let $\mathcal{G} \subseteq \mathcal{F}$ denote a subset of real-valued measurable functions on $\mathcal{X}$ and consider an admissible set $\mathcal{A} \subseteq \mathcal{G} \times \mathcal{M}^m(\mathcal{X})$ defined in the following way. For each $f \in \mathcal{G}$, let $\mathbf{G}(f, \cdot)$ denote a family of constraints as in Theorem 3.1 and Remark 3.2. For each $f \in \mathcal{G}$, let $\mathcal{M}^{\mathbf{G}_f} \subseteq \mathcal{M}^m(\mathcal{X})$ denote those product probability measures μ such that the moments $\mathbf{G}(f, \mu)$ are well-defined. Moreover, for each $f \in \mathcal{G}$, let $r_f \colon \mathcal{X} \to \mathbb{R}$ be integrable for all $\mu \in \mathcal{M}^{\mathbf{G}_f}$ (possibly with values $+\infty$ or $-\infty$). Define the admissible set

$$\mathcal{A} := \{(f, \mu) \in \mathcal{G} \times \mathcal{M}^m(\mathcal{X}) \,|\, \mathbf{G}(f, \mu) \leq 0\}$$

and define the OUQ optimization problem to be

$$\mathcal{U}(\mathcal{A}) := \sup_{(f,\mu) \in \mathcal{A}} \mathbb{E}_\mu[r_f]. \tag{3.1}$$

Corollary 3.4. *Consider the OUQ optimization problem* (3.1) *and define the reduced admissible set* $\mathcal{A}_\Delta \subseteq \mathcal{A}$ *by*

$$\mathcal{A}_\Delta := \left\{ (f, \mu) \in \mathcal{G} \times \bigotimes_{i=1}^m \Delta_{n_i + n'}(\mathcal{X}_i) \,\middle|\, \mathbf{G}(f, \mu) \leq 0 \right\}.$$

Then, it holds that

$$\mathcal{U}(\mathcal{A}) = \mathcal{U}(\mathcal{A}_\Delta).$$

Remark 3.5. Corollary 3.4 is easily generalized to the case where for each $f \in \mathcal{G}$, i, and fixed $\mu_j, j \neq i$, $\mathbf{G}(f, \mu_1, .., \mu_i, .., \mu_m)$ has affine dimension at most m_i as μ_i varies. In this case

$$\mathcal{A}_\Delta := \left\{ (f, \mu) \in \mathcal{G} \times \bigotimes_{i=1}^{m} \Delta_{m_i}(\mathcal{X}_i) \,\middle|\, \mathbf{G}(f, \mu) \leq 0 \right\}.$$

3.2 Simple Generalized Moments

We now consider the case where the function $r_f := r \circ f$ is defined through composition with a measurable function r, and all n constraints are determined by compositions $g'_j := g_j \circ f$, with $j = 1, \ldots, n$, of the function f. Hence, the symbol $\mathbf{G}(f, \mu)$ will mean that all functions $g_j \circ f$ are μ integrable and will represent the values $\mathbb{E}_\mu[g_j \circ f]$ for $j = 1, \ldots, n$. That is, we have the admissible set

$$\mathcal{A} := \{ (f, \mu) \in \mathcal{G} \times \mathcal{M}^m(\mathcal{X}) \,|\, \mathbf{G}(f, \mu) \leq 0 \} \tag{3.2}$$

and the optimization problem

$$\mathcal{U}(\mathcal{A}) := \sup_{(f, \mu) \in \mathcal{A}} \mathbb{E}_\mu[r \circ f] \tag{3.3}$$

as in (3.1). However, in this case, the fact that the criterion function $r \circ f$ and the constraint functions $g_j \circ f$ are compositions of the function f permits a finite-dimensional reduction of the space of functions $\mathcal{G}$ to a space of functions on $\{0, \ldots, n\}^m$ and a reduction of the space of m-fold products of finite convex combinations of Dirac masses to the space of product measures on $\{0, \ldots, n\}^m$. This reduction completely eliminates dependency on the coordinate positions in $\mathcal{X}$.

Formulating this result precisely will require some additional notation. By the Well-Ordering Theorem, there exists a well-ordering of each $\mathcal{X}_i$. Suppose that a total ordering of the elements of the spaces $\mathcal{X}_i$ for $i = 1, \ldots, m$ is specified. Let $\mathcal{N} := \{0, \ldots, n\}$ and $\mathcal{D} := \{0, \ldots, n\}^m = \mathcal{N}^m$. Every element $\mu \in \bigotimes_{i=1}^{m} \Delta_n(\mathcal{X}_i)$ is a product $\mu = \bigotimes_{i=1}^{m} \mu_i$ where each factor μ_i is a convex sum of $n + 1$ Dirac masses indexed according to the ordering; that is,

$$\mu_i = \sum_{k=0}^{n} \alpha_k^i \delta_{x_i^k}$$

for some $\alpha_1^i, \ldots, \alpha_n^i \geq 0$ with unit sum and some $x_i^1, \ldots, x_i^n \in \mathcal{X}_i$ such that, with respect to the given ordering of $\mathcal{X}_i$,

$$x_i^1 \leq x_i^2 \leq \cdots \leq x_i^n.$$

Let $\mathcal{F}_\mathcal{D}$ denote the real linear space of real functions on $\mathcal{D} = \{0, \ldots, n\}^m$ and consider the mapping

$$\mathbb{F} \colon \mathcal{F} \times \bigotimes_{i=1}^{m} \Delta_n(\mathcal{X}_i) \to \mathcal{F}_\mathcal{D}$$

18

defined by

$$(\mathbb{F}(f, \mu))(i_1, i_2, \ldots, i_m) = f(x_1^{i_1}, x_2^{i_2}, \ldots, x_m^{i_m}), \quad i_k \in \mathcal{N}, k = 1, \ldots, m.$$

$\mathbb{F}$ represents the values of the function f at the Dirac masses in μ, but does not carry information regarding the positions of the Dirac masses or their weights.

Theorem 3.6. *Consider the admissible set $\mathcal{A}$ and optimization problem $\mathcal{U}(\mathcal{A})$ defined in (3.2) and (3.3) where $r \circ f$ is integrable (possibly with values $+\infty$ or $-\infty$) for all product measures. For a subset $\mathcal{G}_\mathcal{D} \subseteq \mathcal{F}_\mathcal{D}$, define the admissible set*

$$\mathcal{A}_\mathcal{D} = \{(h, \alpha) \in \mathcal{G}_\mathcal{D} \times \mathcal{M}^m(\mathcal{D}) \,|\, \mathbb{E}_\alpha[g_i \circ h] \le 0 \text{ for all } j = 1, \ldots, n\}$$

and the optimization problem

$$\mathcal{U}(\mathcal{A}_\mathcal{D}) := \sup_{(h, \alpha) \in \mathcal{A}_\mathcal{D}} \mathbb{E}_\alpha[r \circ h].$$

If

$$\mathbb{F}\left(\mathcal{G} \times \bigotimes_{i=1}^{m} \Delta_n(\mathcal{X}_i)\right) = \mathcal{G}_\mathcal{D},$$

then it holds that

$$\mathcal{U}(\mathcal{A}) = \mathcal{U}(\mathcal{A}_\mathcal{D}).$$

When the constraint set also includes functions which are not compositions with f, then Theorem 3.6 does not apply. Although it does appear that results similar to Theorem 3.6 can be obtained, we leave that as a topic for future work.

3.3   Application to McDiarmid's inequality

Theorem 3.6 can be applied to the situation of McDiarmid's inequality in order to obtain an optimal solution for that problem. Let $D_i \ge 0$ for $i = 1, \ldots, m$ and define

$$\mathcal{G} := \{f \in \mathcal{F} \,|\, \mathrm{Osc}_i(f) \le D_i \text{ for each } i = 1, \ldots, m\} \tag{3.4}$$

where

$$\mathrm{Osc}_i(f) := \sup_{(x_1, \ldots, x_m) \in \mathcal{X}} \sup_{x_i' \in \mathcal{X}_i} \left| f(\ldots, x_i, \ldots) - f(\ldots, x_i', \ldots) \right|.$$

We have a product probability measure $\mathbb{P}$ on $\mathcal{X}$ and a measurable function $H \colon \mathcal{X} \to \mathbb{R}$ such that $H \in \mathcal{G}$. Suppose that we have an upper bound

$$\mathbb{P}[H - \mathbb{E}_\mathbb{P}[H] \ge a] \le \mathbb{H}(a, \mathcal{G}) \text{ for all } H \in \mathcal{G}. \tag{3.5}$$

It follows that if $H \in \mathcal{G}$ and $\mathbb{E}_\mathbb{P}[H] \le 0$, then

$$\mathbb{P}[H \ge a] \le \mathbb{P}[H - \mathbb{E}_\mathbb{P}[H] \ge a] \le \mathbb{H}(a, \mathcal{G}) \text{ for all } H \in \mathcal{G} \text{ with } \mathbb{E}_\mathbb{P}[H] \le 0.$$

19

On the other hand, suppose that

$$\mathbb{P}[H \geq a] \leq \mathbb{H}'(a, \mathcal{G}) \text{ for all } H \in \mathcal{G} \text{ with } \mathbb{E}_{\mathbb{P}}[H] \leq 0. \tag{3.6}$$

It follows that

$$\mathbb{P}[H \geq a] \leq \mathbb{H}'(a, \mathcal{G}) \text{ for all } H \in \mathcal{G} \text{ with } \mathbb{E}_{\mathbb{P}}[H] = 0.$$

Since the constraints $\mathcal{G}$ and the event $H - \mathbb{E}_{\mathbb{P}}[H] \geq a$ are invariant under scalar translation $H \mapsto H + c$ it follows that

$$\mathbb{P}[H - \mathbb{E}_{\mathbb{P}}[H] \geq a] \leq \mathbb{H}'(a, \mathcal{G}) \text{ for all } H \in \mathcal{G}.$$

That is, the inequalities (3.5) and (3.6) are equivalent.

McDiarmid's inequality [31, 32] provides the bound $\mathbb{H}(a, \mathcal{G}) := e^{-\frac{2a^2}{D^2}}$ for (3.5) and its equivalent (3.6), with

$$D^2 := \sum_{i=1}^{m} D_i^2. \tag{3.7}$$

Define the admissible set corresponding to the assumptions of McDiarmid's inequality:

$$\mathcal{A}_{\mathrm{McD}} = \{(f, \mu) \in \mathcal{G} \times \mathcal{M}^m(\mathcal{X}) \,|\, \mathbb{E}_\mu[f] \leq 0\}, \tag{3.8}$$

and define the optimization problem

$$\mathcal{U}(\mathcal{A}_{\mathrm{McD}}) := \sup_{(f, \mu) \in \mathcal{A}_{\mathrm{McD}}} \mu[f \geq a].$$

Since $(H, \mathbb{P}) \in \mathcal{A}_{\mathrm{McD}}$ and McDiarmid's inequality $\mu[f \geq a] \leq e^{-\frac{2a^2}{D^2}}$ is satisfied for all $(f, \mu) \in \mathcal{A}_{\mathrm{McD}}$, it follows that

$$\mathbb{P}[H \geq a] \leq \mathcal{U}(\mathcal{A}_{\mathrm{McD}}) \leq e^{-\frac{2a^2}{D^2}}.$$

Moreover, the inequality on the left is optimal in the sense that, for every $\varepsilon > 0$, there exists a McDiarmid-admissible scenario (f, μ) satisfying the same assumptions as $(H, \mathbb{P})$ such that $\mu[f \geq a] \geq \mathcal{U}(\mathcal{A}_{\mathrm{McD}}) - \varepsilon$.

To apply the previous results to computing $\mathcal{U}(\mathcal{A}_{\mathrm{McD}})$, let $\mathcal{D} := \{0, 1\}^m$ and define

$$\mathcal{G}_{\mathcal{D}} := \{h \in \mathcal{F}_{\mathcal{D}} \mid \mathrm{Osc}_k(h) \leq D_k \text{ for each } k = 1, \ldots, m\},$$

where the inequality $\mathrm{Osc}_k(h) \leq D_k$ for $h \in \mathcal{F}_{\mathcal{D}}$ means that

$$|h(s_1, \ldots, s_k, \ldots, s_m) - h(s_1, \ldots, s_{k'}, \ldots, s_m)| \leq D_i, \tag{3.9}$$

for all $s_j \in \{0, 1\}$, $j = 1, \ldots, m$, and all $s_{k'} \in \{0, 1\}$. Define the corresponding admissible set

$$\mathcal{A}_{\mathcal{D}} = \{(h, \alpha) \in \mathcal{G}_{\mathcal{D}} \times \mathcal{M}(\{0, 1\})^m \,|\, \mathbb{E}_\alpha[h] \leq 0\}$$

and the optimization problem

$$\mathcal{U}(\mathcal{A}_{\mathcal{D}}) := \sup_{(h, \alpha) \in \mathcal{A}_{\mathcal{D}}} \alpha[h \geq a]. \tag{3.10}$$

20

Proposition 3.7. *It holds that*

$$\mathcal{U}(\mathcal{A}_{\text{McD}}) = \mathcal{U}(\mathcal{A}_{\mathcal{D}}). \tag{3.11}$$

We now provide a further reduction of $\mathcal{U}(\mathcal{A}_{\text{McD}})$ by reducing $\mathcal{U}(\mathcal{A}_{\mathcal{D}})$. To that end, for two vertices s and t of $\mathcal{D} = \{0,1\}^m$, let $I(s,t)$ be the set of indices i such that $s_i \neq t_i$. For $s \in \mathcal{D}$, define the function $h^s \in \mathcal{F}_{\mathcal{D}}$ by

$$h^s(t) = a - \sum_{i \in I(s,t)} D_i.$$

For $C \subseteq \mathcal{D}$, define $h^C \in \mathcal{F}_{\mathcal{D}}$ by

$$h^C(t) := \max_{s \in C} h^s(t) = a - \min_{s \in C} \sum_{i \in I(s,t)} D_i. \tag{3.12}$$

Let $\mathcal{C} := \{C \mid C \subseteq \mathcal{D}\}$ be the power set of $\mathcal{D}$ (the set of all subsets of $\mathcal{D}$), define the admissible set $\mathcal{A}_{\mathcal{C}}$ by

$$\mathcal{A}_{\mathcal{C}} := \left\{ (C,\alpha) \in \mathcal{C} \times \mathcal{M}(\{0,1\})^m \mid \mathbb{E}_\alpha[h^C] \leq 0 \right\}$$

and consider the optimization problem

$$\mathcal{U}(\mathcal{A}_{\mathcal{C}}) := \sup_{(C,\alpha) \in \mathcal{A}_{\mathcal{C}}} \alpha(h^C \geq a). \tag{3.13}$$

Theorem 3.8. *It holds that*

$$\mathcal{U}(\mathcal{A}_{\mathcal{D}}) = \mathcal{U}(\mathcal{A}_{\mathcal{C}}). \tag{3.14}$$

Remark 3.9. The proof of this reduction theorem utilizes the standard lattice structure of the space of functions $\mathcal{F}_{\mathcal{D}}$ in a substantial way. To begin with, the reduction to $\max h = a$ is attained through lattice invariance. Moreover, we have a lattice $\mathcal{F}_{\mathcal{D}}$, with sub-lattice $\mathcal{G}_{\mathcal{D}}$, and for each $C \in \mathcal{C}$, the set $C_{\mathcal{D}} := \{h \in \mathcal{F}_{\mathcal{D}} \mid \{s \mid h(s) = a\} = C\}\}$ of functions with value a precisely on the set C is a sub-lattice. For a clipped h, let $C(h) := \{s \in \mathcal{D} \mid h(s) = a\}$ be the set where h has the value a. Then, if for each C the set

$$\bigcap_{h:C(h)=C} \{f \leq h\} \cap C_{\mathcal{D}} \cap \mathcal{G}_{\mathcal{D}}$$

is nonempty, then we obtain a reduction. However, not only is the set nonempty, but the map $C \mapsto h^C$ is a simple algorithm that produces a point in this intersection, and therefore an explicit reduction. We suspect that the existence of a simple reduction algorithm in this case is due to the lattice structures, and that such structures may be useful in the more general case. Indeed, the condition $f \leq h$ implies that $\mathbb{E}_\alpha[f] \leq \mathbb{E}_\alpha[h]$ for any α, and the the condition that $\mathbb{E}_\alpha[f] \leq \mathbb{E}_\alpha[h]$ for all α implies that $f \leq h$, so that the above condition may be stated as

$$\bigcap_{h:C(h)=C} \left\{ \bigcap_\alpha \{\mathbb{E}_\alpha[f] \leq \mathbb{E}_\alpha[h]\} \right\} \cap C_{\mathcal{D}} \cap \mathcal{G}_{\mathcal{D}}.$$

21

For the more general constraints, we would instead have to solve (*i.e.*, find an element of)

$$\bigcap_{h:C(h)=C}\left\{\bigcap_{\alpha}\{\mathbf{G}(f,\alpha)\leq\mathbf{G}(h,\alpha)\}\right\}\cap C_{\mathcal{D}}\cap\mathcal{G}_{\mathcal{D}}.$$

4 Optimal Concentration Inequalities

In this Section, the results of Section 3 will be applied to obtain optimal concentration inequalities under the assumptions of McDiarmid's inequality and Hoeffding's inequality. The following subsection gives explicit concentration results under the assumptions of McDiarmid's inequality, and Subsection 4.2 gives explicit concentration results under the assumptions of Hoeffding's inequality. Surprisingly, these explicit results show that uncertainties in input parameters do not necessarily propagate to global (output) uncertainties. We refer to Subsection 6.2 for the proofs.

4.1 Explicit solutions under the assumptions of McDiarmid's inequality

In this subsection, we will apply Theorem 3.8 to obtain explicit formulae for the OUQ problem $\mathcal{U}(\mathcal{A}_{\mathrm{McD}})$ (defined in equation (3.10)) under the assumptions of McDiarmid's inequality (3.8). More precisely, we will compute $\mathcal{U}(\mathcal{A}_{\mathcal{C}})$ defined by equation (3.13) and use equalities (3.14) and (3.11) to obtain $\mathcal{U}(\mathcal{A}_{\mathrm{McD}})=\mathcal{U}(\mathcal{A}_{\mathcal{C}})$. It is important to observe that all the following optimization problems possess solutions because they involve the optimization of a continuous function (with respect to α) in a compact space.

Since the inequalities (3.5) and (3.6) are equivalent, it follows that

$$\mathcal{U}(\mathcal{A}_{\mathrm{McD}})=\sup_{(f,\mu)\in\mathcal{G}\times\mathcal{M}_m}\mu\big[f\geq a+\mathbb{E}_\mu[f]\big].$$

In particular, if $\mathbb{E}_\mu[f]\leq 0$ is replaced by $\mathbb{E}_\mu[f]\leq b$ or $\mathbb{E}_\mu[f]=b$ in McDiarmid's inequality assumptions (3.8), then the results given in this section remain valid by replacing a by $M:=a-b$ (observe that M plays the role of a margin).

Those results should be compared with McDiarmid's inequality [31, 32], which provides the bound

$$\sup_{(f,\mu)\in\mathcal{G}\times\mathcal{M}_m}\mu\big[f\geq a+\mathbb{E}_\mu[f]\big]\leq e^{-\frac{2a^2}{\sum_{i=1}^m D_i^2}}. \tag{4.1}$$

The statements of the theorem will be given assuming that $a\geq 0$; in the complementary case of $a<0$, the solution is simply $\mathcal{U}(\mathcal{A}_{\mathrm{McD}})=1$.

To the best of the authors' knowledge, the optimal bounds given here are new. There is a substantial literature relating to optimization of concentration bounds and de-randomization algorithms (see for instance [49] and references therein) but, as far as the authors know, those bounds were suboptimal because they were obtained through the moment generating function technique.

4.1.1 Explicit solutions in dimensions one and two

Theorem 4.1 (Explicit solution for $m = 1$). *For $m = 1$, $\mathcal{U}(\mathcal{A}_{\mathrm{McD}})$ is given by*

$$\mathcal{U}(\mathcal{A}_{\mathrm{McD}}) = \begin{cases} 0, & \text{if } D_1 \leq a, \\ 1 - \dfrac{a}{D_1}, & \text{if } 0 \leq a \leq D_1. \end{cases} \tag{4.2}$$

Theorem 4.2 (Explicit solution for $m = 2$). *For $m = 2$, $\mathcal{U}(\mathcal{A}_{\mathrm{McD}})$ is given by*

$$\mathcal{U}(\mathcal{A}_{\mathrm{McD}}) = \begin{cases} 0, & \text{if } D_1 + D_2 \leq a, \\ \dfrac{(D_1 + D_2 - a)^2}{4 D_1 D_2}, & \text{if } |D_1 - D_2| \leq a \leq D_1 + D_2, \\ 1 - \dfrac{a}{\max(D_1, D_2)}, & \text{if } 0 \leq a \leq |D_1 - D_2|. \end{cases} \tag{4.3}$$

See Figure 4.1 for an illustration comparing the McDiarmid and OUQ bounds for $m = 2$. Observe that

- If $a \leq D_1 - D_2$, then a decrease in D_2 does not lead to a decrease in the OUQ bound $\mathcal{U}(\mathcal{A}_{\mathrm{McD}})$. In other words, if most of the uncertainty is contained in the first variable $(a + D_2 \leq D_1)$, then the uncertainty associated with the second variable does not affect the global uncertainty; a reduction of the global uncertainty requires a reduction in D_1.

- For $D_1 + D_2 = 2a$, the ratio between the OUQ bound and the McDiarmid bound is minimized near the diagonal.

Remark 4.3. The maximum of (4.3) over D_1, D_2 under the constraints $D_1 + D_2 = D$ and $D_1 \geq D_2$ is achieved at $D_2 = 0$ and is equal to $1 - a/D$. The minimum of (4.3) over D_1, D_2 under the constraints $D_1 + D_2 = D$ and $D_1 \geq D_2$ is achieved on the diagonal $D_1 = D_2$ and is equal to $(1 - a/D)^2$.

4.1.2 Explicit solution in dimension three

Assume that $D_1 \geq D_2 \geq D_3$. Write

$$\mathcal{F}_1 := \begin{cases} 0, & \text{if } D_1 + D_2 + D_3 \leq a, \\ \dfrac{(D_1 + D_2 + D_3 - a)^3}{27 D_1 D_2 D_3}, & \text{if } D_1 + D_2 - 2D_3 \leq a \leq D_1 + D_2 + D_3, \\ \dfrac{(D_1 + D_2 - a)^2}{4 D_1 D_2}, & \text{if } D_1 - D_2 \leq a \leq D_1 + D_2 - 2D_3, \\ 1 - \dfrac{a}{D_1}, & \text{if } 0 \leq a \leq D_1 - D_2. \end{cases} \tag{4.4}$$

and

$$\mathcal{F}_2 := \max_{i \in \{1,2,3\}} \phi(\gamma_i) \psi(\gamma_i) \tag{4.5}$$
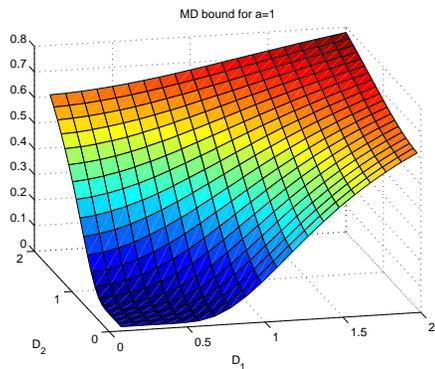
(a) McDiarmid upper bound
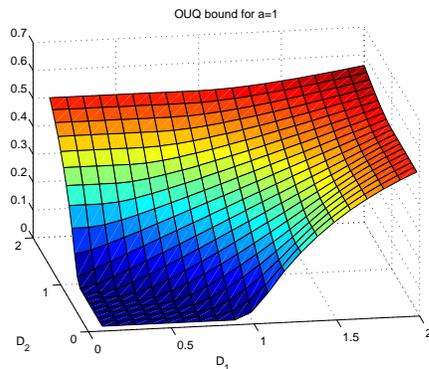
(b) OUQ upper bound
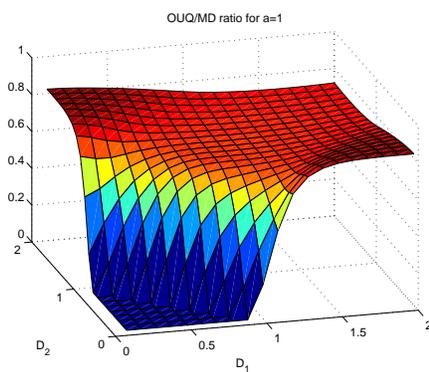
(c) Ratio of the two bounds: OUQ bound
divided by McDiarmid bound

Figure 4.1: Comparison of the McDiarmid and OUQ bounds for $m = 2$, with mean performance 0 and failure threshold $a = 1$. The optimal upper bound is calculated using the explicit solution (4.3) to the reduced optimization problem $\mathcal{U}(\mathcal{A}_{\mathcal{C}})$ defined by equation (3.13).

where
$$\psi(\gamma) := \gamma^2 \left(2\frac{D_2}{D_3} - 1\right) - 2\gamma \left(3\frac{D_2}{D_3} - 1\right) + \frac{\gamma}{1+\gamma}\left(8\frac{D_2}{D_3} - 2\frac{a}{D_3}\right)$$

and $\gamma_1, \gamma_2, \gamma_3$ are the roots (in γ) of the cubic polynomial

$$(1+\gamma)^3 - A(1+\gamma)^2 + B = 0, \tag{4.6}$$

where

$$A := \frac{5D_2 - 2D_3}{2D_2 - D_3},$$

$$B := \frac{4D_2 - a}{2D_2 - D_3}.$$

Define a function ϕ by

$$\phi(\gamma) := \begin{cases} 1, & \text{if } \gamma \in (0,1) \text{ and } \theta(\gamma) \in (0,1), \\ 0, & \text{otherwise,} \end{cases}$$

where

$$\theta(\gamma) := 1 - \frac{a}{D_3(1-\gamma^2)} + \frac{D_2}{D_3}\frac{1-\gamma}{1+\gamma}.$$

By the standard formula for the roots of a cubic polynomial, the roots of (4.6) are given by

$$\gamma_1 := -1 - \frac{1}{3}\left(-A + \kappa_1 + \kappa_2\right),$$

$$\gamma_2 := -1 - \frac{1}{3}\left(-A + \omega_2\kappa_1 + \omega_1\kappa_2\right),$$

$$\gamma_3 := -1 - \frac{1}{3}\left(-A + \omega_1\kappa_1 + \omega_2\kappa_2\right),$$

where

$$\omega_1 := -\frac{1}{2} + \frac{\sqrt{3}}{2}i, \quad \omega_2 := -\frac{1}{2} - \frac{\sqrt{3}}{2}i,$$

$$\kappa_1 := \left(\frac{\beta_1 + \sqrt{\beta_2}}{2}\right)^{\frac{1}{3}}, \quad \kappa_2 := \left(\frac{\beta_1 - \sqrt{\beta_2}}{2}\right)^{\frac{1}{3}},$$

$$\beta_1 := -2A^3 + 27B, \text{ and } \beta_2 := \beta_1^2 - 4A^6.$$

Since there are 3 possible values for each cube root, κ_1 and κ_2 must be taken so that they satisfy $\kappa_1\kappa_2 = A^2$.

Theorem 4.4 (Explicit solution for $m = 3$). *For $m = 3$ with $D_1 \geq D_2 \geq D_3$, $\mathcal{U}(\mathcal{A}_{\mathrm{McD}})$ is given by*

$$\mathcal{U}(\mathcal{A}_{\mathrm{McD}}) = \max(\mathcal{F}_1, \mathcal{F}_2). \tag{4.7}$$

Remark 4.5. Figures 4.2 and 4.3 suggest that the inequality $\mathcal{F}_2 > \mathcal{F}_1$ holds only if $D_3 \approx D_2$, and D_2 is large enough relative to D_1.

25

4.1.3 Explicit solution in dimension m

For $C_0 \in \mathcal{C}$, write

$$\mathcal{U}(\mathcal{A}_{C_0}) = \sup_{\alpha \, : \, (C_0, \alpha) \in \mathcal{A}_{\mathcal{C}}} \alpha[h^{C_0} \geq a], \qquad (4.8)$$

where h^{C_0} is defined by equation (3.12).

Proposition 4.6. *Assume that $D_1 \geq \cdots \geq D_{m-1} \geq D_m$. For $C_0 := \{(1, 1, \ldots, 1, 1)\}$, it holds that*

$$\mathcal{U}(\mathcal{A}_{C_0}) = \begin{cases} 0, & \text{if } \sum_{j=1}^{m} D_j \leq a, \\[2mm] \dfrac{(\sum_{j=1}^{m} D_j - a)^m}{m^m \prod_{j=1}^{m} D_j}, & \text{if } \sum_{j=1}^{m} D_j - mD_m \leq a \leq \sum_{j=1}^{m} D_j, \\[2mm] \dfrac{(\sum_{j=1}^{k} D_j - a)^k}{k^k \prod_{j=1}^{k} D_j}, & \text{if, for } k \in \{1, \ldots, m-1\}, \\[2mm] & \sum_{j=1}^{k} D_j - kD_k \leq a \leq \sum_{j=1}^{k+1} D_j - (k+1)D_{k+1}. \end{cases} \qquad (4.9)$$

Remark 4.7. The maximum of (4.9) over $D_1, \ldots, D_m$ under the constraints $D_1 + \cdots + D_m = D$ and $D_1 \geq \cdots \geq D_m$ is achieved at $D_1 = D$ and is equal to $1 - a/D$.
The minimum of (4.9) over $D_1, \ldots, D_m$ under the constraints $D_1 + \cdots + D_m = D$ and $D_1 \geq \cdots \geq D_m$ is achieved on the diagonal $D_1 = \cdots = D_m$ and is equal to $(1 - a/D)^m$.

Proposition 4.8. *Assume that $D_1 \geq \cdots \geq D_{m-1} \geq D_m$. If $a \geq \sum_{j=1}^{m-2} D_j + D_m$, then $\mathcal{U}(\mathcal{A}_{\mathrm{McD}})$ is given by equation (4.9).*

4.2 Explicit solutions under the assumptions of Hoeffding's inequality

This subsection treats a further special case of OUQ, where the assumptions correspond to Hoeffding's inequality [18]. Define the admissible set

$$\mathcal{A}_{\mathrm{Hfd}} := \left\{ (f, \mu) \, \middle| \, \begin{matrix} f = X_1 + \cdots + X_m, \\ \mu \in \bigotimes_{i=1}^{m} \mathcal{M}([b_i - D_i, b_i]), \\ \mathbb{E}_\mu[f] \leq 0 \end{matrix} \right\}, \qquad (4.10)$$

and define the optimization problem

$$\mathcal{U}(\mathcal{A}_{\mathrm{Hfd}}) := \sup_{(f, \mu) \in \mathcal{A}_{\mathrm{Hfd}}} \mu[f \geq a].$$

By Hoeffding's inequality, for $a \geq 0$,

$$\mathcal{U}(\mathcal{A}_{\mathrm{Hfd}}) \leq \exp\left(-2 \frac{a^2}{\sum_i^m D_i^2} \right).$$

Theorem 4.9. *If $m = 2$, then*

$$\mathcal{U}(\mathcal{A}_{\mathrm{Hfd}}) = \mathcal{U}(\mathcal{A}_{\mathrm{McD}}). \qquad (4.11)$$
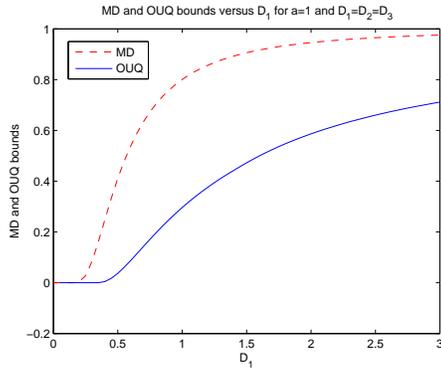
(a) McDiarmid vs OUQ bound
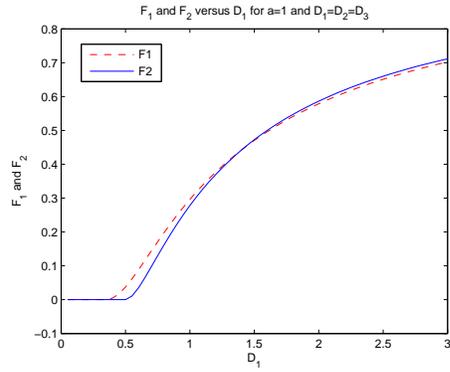
(b) $\mathcal{F}_1$ vs $\mathcal{F}_2$

Figure 4.2: Comparison of the McDiarmid and OUQ bounds for $m = 3$, with mean performance $m = 0$, $D_1 = D_2 = D_3$, and failure threshold $a = 1$. Observe that $\mathcal{F}_2 > \mathcal{F}_1$ for D_1 large enough.
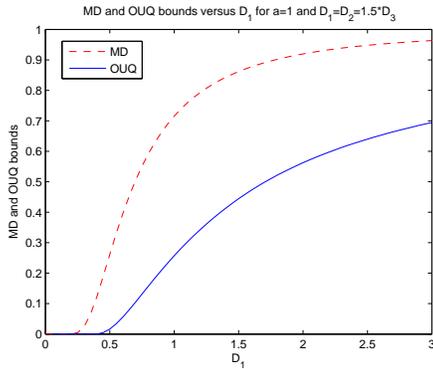
(a) McDiarmid vs OUQ bound
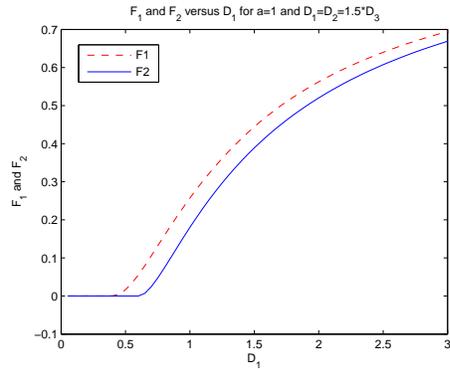
(b) $\mathcal{F}_1$ vs $\mathcal{F}_2$

Figure 4.3: Comparison of the McDiarmid and OUQ bounds for $m = 3$, with mean performance 0, $D_1 = D_2 = \frac{3}{2}D_3$, and failure threshold $a = 1$. Observe that $\mathcal{F}_2 < \mathcal{F}_1$ for all D_1.

Remark 4.10. Another proof of Theorem 4.9 can be obtained using entirely different methods than presented in Section 6.2. Although omitted for brevity, this method may be useful in higher dimensions, so we describe an outline of it here. We begin at the reduction obtained through Proposition 3.7 to the hypercube. Whereas the proof of Theorem 4.9 first applies the reduction of Theorem 3.8 to subsets of the hypercube, here we instead fix the oscillations in each direction to be $0 \leq d_i \leq D_i$, and solve the fixed $d := (d_1, d_2)$ case, not using a Langrangian type analysis but a type of spectral reduction. We then show that the resulting value $\mathcal{U}(d)$ is increasing in d with respect to the standard (lexicographic) partial order on vectors. The result then easily follows by taking the supremum over all vectors $0 \leq d \leq D$.

Theorem 4.11. *Let* $m = 3$*, and define* $\mathcal{F}_1$ *and* $\mathcal{F}_2$ *as in Theorem 4.4. If* $\mathcal{F}_1 \geq \mathcal{F}_2$*, then*

$$\mathcal{U}(\mathcal{A}_{\mathrm{Hfd}}) = \mathcal{U}(\mathcal{A}_{\mathrm{McD}}). \tag{4.12}$$

If $\mathcal{F}_1 < \mathcal{F}_2$*, then*

$$\mathcal{U}(\mathcal{A}_{\mathrm{Hfd}}) < \mathcal{U}(\mathcal{A}_{\mathrm{McD}}). \tag{4.13}$$

Under the assumptions of Hoeffding's inequality, each variable X_i is bounded from below and from above. Without the upper bounds on the variables X_i, it is possible to leverage additional reduction properties and conjecture an explicit form for the optimal inequality on $\mu[X_1 + \cdots + X_m \geq a]$. Here we refer to the work and conjecture of Samuels [45] (see also page 542 of [24]), which has been proven true for $m = 1, 2, 3$.

5 Computational Implementation and Examples

In this section, we discuss the results of numerical implementation of OUQ algorithms for an analytical surrogate model for hypervelocity impact. In particular, we compute the upper ($\mathcal{U}(\mathcal{A})$) and lower ($\mathcal{L}(\mathcal{A})$) bounds on the probability of failure for various sets of assumptions $\mathcal{A}$, and also consider an example of the experimental selection problem.

The physical system of interest is one in which a 400C steel ball of diameter $D_{\mathrm{p}} = 1.778 \, \mathrm{mm}$ impacts a 440C steel plate of thickness h (expressed in mm) at speed v (expressed in $\mathrm{km} \cdot \mathrm{s}^{-1}$) at obliquity θ from the plate normal. The physical experiments are performed at the California Institute of Technology SPHIR (Small Particle Hypervelocity Impact Range) facility. An analytical surrogate model was developed to calculate the perforation area (in mm^2) caused by this impact scenario. The surrogate response function is as follows:

$$H(h, \theta, v) = K \left(\frac{h}{D_{\mathrm{p}}} \right)^p (\cos \theta)^u \left(\tanh \left(\frac{v}{v_{\mathrm{bl}}} - 1 \right) \right)_+^m, \tag{5.1}$$

where the *ballistic limit velocity* (the speed below which no perforation area occurs) is given by

$$v_{\mathrm{bl}} := H_0 \left(\frac{h}{(\cos \theta)^n} \right)^s. \tag{5.2}$$

28

The seven quantities H_0, s, n, K, p, u and m are fitting parameters that have been chosen to minimize the least-squares error between the surrogate and a set of 56 experimental data points; they take the values

$$H_0 = 0.5794\,\mathrm{km\cdot s^{-1}}, \qquad s = 1.4004, \qquad n = 0.4482, \qquad K = 10.3936\,\mathrm{mm^2},$$
$$p = 0.4757, \qquad\qquad u = 1.0275, \qquad m = 0.4682.$$

For the remainder of this section, the surrogate response function, H, will be taken as given and fixed, and its behaviour will be investigated on the input parameter range

$$(h, \theta, v) \in \mathcal{X} := \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3, \tag{5.3a}$$
$$h \in \mathcal{X}_1 := [1.524, 2.667]\,\mathrm{mm} = [60, 105]\,\mathrm{mils}, \tag{5.3b}$$
$$\theta \in \mathcal{X}_2 := [0, \tfrac{\pi}{6}], \tag{5.3c}$$
$$v \in \mathcal{X}_3 := [2.1, 2.8]\,\mathrm{km\cdot s^{-1}}. \tag{5.3d}$$

We will measure lengths in both mm and mils (recall that $1\,\mathrm{mm} = 39.4\,\mathrm{mils}$). Since H has been fixed, the optimizations that follow are all optimizations with respect to the unknown random distribution $\mathbb{P} \in \mathcal{M}(\mathcal{X})$ of the inputs of H. We adopt the "gunner's perspective" that the failure event is non-perforation, *i.e.*, the event $[H = 0]$.

5.1 Computational Examples

Example 5.1 (Computation of optimal bounds on the probability of non-perforation)**.** We assume that the impact velocity, impact obliquity and plate thickness are independent random variables, and that the mean perforation area must lie in a prescribed range $[m_1, m_2] := [5.5, 7.5]\,\mathrm{mm^2}$. Therefore, the corresponding admissible set for the OUQ problem is

$$\mathcal{A} := \left\{ (H, \mu) \,\middle|\, \begin{array}{l} H \text{ given by (5.1)}, \\ \mu = \mu_1 \otimes \mu_2 \otimes \mu_3, \\ m_1 \le \mathbb{E}_\mu[H] \le m_2 \end{array} \right\}. \tag{5.4}$$

By way of contrast, an admissible set that corresponds to this problem and the assumptions of McDiarmid's inequality is

$$\mathcal{A}_{\mathrm{McD}} := \left\{ (f, \mu) \,\middle|\, \begin{array}{c} \mu = \mu_1 \otimes \mu_2 \otimes \mu_3, \\ m_1 \le \mathbb{E}_\mu[f] \le m_2, \\ \mathrm{Osc}_i(f) \le \mathrm{Osc}_i(H) \text{ for } i = 1, 2, 3 \end{array} \right\}. \tag{5.5}$$

We refer to equation (3.9) for the definition of $\mathrm{Osc}_i(f)$ and to equation (3.8) for measurability issues. This example serves to illustrate the extent to which McDiarmid's inequality is not sharp. Application of McDiarmid's inequality to the above situation yields the bound

$$\mathbb{P}[H = 0] \le \mathcal{U}(\mathcal{A}_{\mathrm{McD}}) \le \exp\left(-\frac{2m_1^2}{\sum_{i=1}^3 \mathrm{Osc}_i(H)^2} \right) = 66.4\%.$$

The "optimal McDiarmid inequality" of Theorem 4.4 and Remark 3.2 instead provide the upper bound

$$\mathbb{P}[H = 0] \leq \mathcal{U}(\mathcal{A}_{\mathrm{McD}}) = 43.7\%.$$

With access to H, not just its componentwise oscillations, even sharper bounds on the probability of non-perforation can be calculated; these calculations are greatly simplified by the reduction results of Section 3. Theorem 3.1 implies that in order to find $\mathcal{U}(\mathcal{A})$ it is sufficient to search among those measures μ whose marginal distributions on each of the three input parameter ranges have support on at most two points. That is, $\mathcal{U}(\mathcal{A}) = \mathcal{U}(\mathcal{A}_\Delta)$, where the reduced feasible set $\mathcal{A}_\Delta$ is given by

$$\mathcal{A}_\Delta := \left\{ (H, \mu) \,\middle|\, \begin{array}{c} H \text{ given by (5.1)}, \\ \mu = \mu_1 \otimes \mu_2 \otimes \mu_3, \\ \mu_i \in \Delta_1(\mathcal{X}_i) \text{ for } i = 1, 2, 3, \\ m_1 \leq \mathbb{E}_\mu[H] \leq m_2 \end{array} \right\}. \tag{5.6}$$

A numerical optimization over this finite-dimensional reduced feasible set $\mathcal{A}_\Delta$ using a genetic algorithm global optimization routine in the *mystic* framework [34] yields the following optimal upper bound on the probability of non-perforation:

$$\mathbb{P}[H = 0] \leq \mathcal{U}(\mathcal{A}) = \mathcal{U}(\mathcal{A}_\Delta) \overset{\mathrm{num}}{=} 37.9\%.$$

Observe that $\mathbb{P}[H = 0] \leq \mathcal{U}(\mathcal{A}) = \mathcal{U}(\mathcal{A}_\Delta)$ is a theorem, whereas the $\mathcal{U}(\mathcal{A}_\Delta) \overset{\mathrm{num}}{=} 37.9\%$ is the output of an algorithm (in this case, a Differential Evolution algorithm implemented in the *mystic framework*, see Subsection 5.3). In particular, its validity is correlated with the efficiency of the specific algorithm. We have introduced the symbol $\overset{\mathrm{num}}{=}$ to emphasize the distinction between mathematical (in)equalities and numerical outputs.

Although we do not have a theorem associated with the convergence of the numerical optimization algorithm, we have a robust control over its efficiency because it is applied to the finite dimensional problem $\mathcal{U}(\mathcal{A}_\Delta)$ instead of the infinite optimization problem associated with $\mathcal{U}(\mathcal{A})$ (thanks to the reduction theorems obtained in Section 3).

For $\#\mathrm{supp}(\mu_i) \leq 2$, $i = 1, 2, 3$ (where $\#\mathrm{supp}(\mu_i)$ is the number of points forming the support of μ_i), Figure 5.1 shows that numerical simulations collapse to two-point support. Indeed, even when a wider search is performed (*i.e.*, over measures $\mu \in \bigotimes_{i=1}^3 \Delta_k(\mathcal{X}_i)$ for $k > 1$), it is observed that the calculated maximizers for these problems maintain two-point support: the velocity and obliquity marginals each collapse to a single Dirac mass, and the plate thickness marginal collapses to have support on the two extremes of its range. As expected, optimization over a larger search space is more computationally intensive and takes longer to perform. This observation suggests that the extreme points of the reduced OUQ problems are, in some sense, attractors — this point will be revisited in the next subsection.

We also refer to figures 5.4, 5.5 and 5.6 for plots of the locations and weights of the Dirac masses forming each marginal μ_i as functions of the number of iterations. Note that the lines for *thickness* and *thickness weight* are of the same color if they correspond to the same support point for the measure. In particular, Figure 5.6 shows that at

30

Admissible scenarios, $\mathcal{A}$	$\mathcal{U}(\mathcal{A})$	Method
$\mathcal{A}_{\mathrm{McD}}$: independence, oscillation and mean constraints as given by (5.5)	$\leq 66.4\%$ $= 43.7\%$	McDiarmid's inequality Theorem 4.4
$\mathcal{A}$ as given by (5.4)	$\overset{\mathrm{num}}{=} 37.9\%$	*mystic*, H known
$\mathcal{A} \cap \left\{ (H, \mu) \left\| \begin{array}{c} \mu\text{-median velocity} \\ = 2.45\,\mathrm{km} \cdot \mathrm{s}^{-1} \end{array} \right. \right\}$	$\overset{\mathrm{num}}{=} 30.0\%$	*mystic*, H known
$\mathcal{A} \cap \left\{ (H, \mu) \,\middle|\, \mu\text{-median obliquity} = \frac{\pi}{12} \right\}$	$\overset{\mathrm{num}}{=} 36.5\%$	*mystic*, H known
$\mathcal{A} \cap \left\{ (H, \mu) \,\middle|\, \text{obliquity} = \frac{\pi}{6} \; \mu\text{-a.s.} \right\}$	$\overset{\mathrm{num}}{=} 28.0\%$	*mystic*, H known

Table 5.1: Summary of the upper bounds on the probability of non-perforation for Example 5.1 as obtained by various methods. Note that OUQ calculations using *mystic* involve evaluations of the function H, whereas McDiarmid's inequality and the optimal bound given the assumptions of McDiarmid's inequality use only the mean of H and its McDiarmid subdiameters, not H itself. Note also that the incorporation of additional information (assumptions), *e.g.* on impact obliquity, always reduces the OUQ upper bound on the probability of failure, since this corresponds to a restriction to a subset of the original feasible set $\mathcal{A}$ for the optimization problem.

iteration number 3500 the *thickness* support point at $62.5\,\mathrm{mils}$ (shown in Figure 5.3) has zero weight.

The inclusion of additional information further reduces the upper bound (the addition of assumptions lead to a smaller admissible set $\mathcal{A} \mapsto \mathcal{A}' \subset \mathcal{A}$, therefore $\mathcal{U}$ decreases and $\mathcal{L}$ increases). For example, if the median of the third input parameter (velocity) is known to lie at the midpoint of its range, and this information is used to provide an additional constraint, then the least upper bound on the probability of non-perforation drops to 30.0%.

See Table 5.1 for a summary of the bounds presented in this example, and for further examples of the effect of additional information/constraints.

On the selectiveness of the information set $\mathcal{A}$. Observe that, except for the bound obtained from McDiarmid's inequality, the bounds obtained in Table 5.1 are the best possible given the information contained in $\mathcal{A}$. If the unknown distribution $\mathbb{P}$ is completely specified, say by restricting to the feasible set $\mathcal{A}_{\mathrm{unif}}$ for which the only admissible measure is the uniform probability measure on the cube $\mathcal{X}$ (in which case the mean perforation area is $\mathbb{E}[H] = 6.58\,\mathrm{mm}^2$), then the probability of non-perforation is $\mathcal{U}(\mathcal{A}_{\mathrm{unif}}) = \mathcal{L}(\mathcal{A}_{\mathrm{unif}}) \overset{\mathrm{num}}{=} 3.8\%$. One may argue that there is a large gap between the fifth (28%) row of Table 5.1 and 3.8% but observe that 3.8% relies on the exact knowledge of G (called H here) and $\mathbb{P}$ whereas 28% relies on the limited knowledge contained in $\mathcal{A} \cap$

$\{(H, \mu) \,|\, \text{obliquity} = \frac{\pi}{6} \, \mu\text{-a.s.}\}$ with respect to which 28% is optimal. In particular, the gap between those two values is not caused by a lack of tightness of the method, but by a lack of selectiveness of the information contained in $\mathcal{A} \cap \{(H, \mu) \,|\, \text{obliquity} = \frac{\pi}{6} \, \mu\text{-a.s.}\}$. The (mis)use of the terms "tight" and "sharp" without reference to available information (and in presence of asymmetric information) can be the source of much confusion. Given prior knowledge of G and $\mathbb{P}$, it would be an easy task to construct a set $\mathcal{A}_{\mathbb{P},G}$ containing $(G, \mathbb{P})$ and such that $\mathcal{U}(\mathcal{A}_{\mathbb{P},G}) \approx 4\%$, but doing so would be delaying a honest discussion on one of the issues at the core of UQ: *How to construct $\mathcal{A}$ without prior knowledge of G and $\mathbb{P}$?* In other words, how to improve the "selectiveness" or $\mathcal{A}$ or how to design experiments leading to "narrow" $\mathcal{A}$s? We have shown in Section 2 that this question is well-posed within the OUQ framework and has an optimal solution. The following example shows that the notion of "best" experiment depends on the admissible safety threshold ϵ for $\mathbb{P}[G \geq a]$.

Example 5.2 (Computational solution of the experimental selection problem)**.** Now consider the problem of experimental selection as discussed in Section 2, and consider again the admissible set $\mathcal{A}$ as given by (5.4):

$$\mathcal{A} := \left\{ (H, \mu) \,\middle|\, \begin{array}{c} H \text{ given by } (5.1), \\ \mu = \mu_1 \otimes \mu_2 \otimes \mu_3 \in \mathcal{M}(\mathcal{X}), \\ m_1 \leq \mathbb{E}_\mu[H] \leq m_2 \end{array} \right\}.$$

Suppose that an experiment E is proposed that will determine the probability mass of the upper half of the velocity range, $[2.45, 2.8] \, \text{km} \cdot \text{s}^{-1}$; the corresponding functional Φ of study is

$$\Phi(\mu) := \mu[v \geq 2.45 \, \text{km} \cdot \text{s}^{-1}],$$

and the proposed experiment E will determine $\Phi(\mathbb{P})$ (approximate determinations including measurement and sampling errors can also be handled, but the exact determination is done here for simplicity). The corresponding intervals $J_{\text{safe},\epsilon}(\Phi)$ and $J_{\text{unsafe},\epsilon}(\Phi)$ as defined by (2.6) and (2.3) are reported in Table 5.2 for various acceptable probabilities of failure ϵ. Note that, for larger values of ϵ, E is a "better" experiment in the sense that it can distinguish safe scenarios from unsafe ones (see also Figure 2.2); for smaller values of ϵ, E is a poor experiment. In any case, since the intersection $J_{\text{safe},\epsilon}(\Phi) \cap J_{\text{unsafe},\epsilon}(\Phi)$ is not empty, E is not an ideal experiment.

It is important to note that the optimization calculations necessary to compute the intervals $J_{\text{safe},\epsilon}(\Phi)$ and $J_{\text{unsafe},\epsilon}(\Phi)$ are simplified by the application of Theorem 3.1: in this case, the objective function is $\mu[v \geq 2.45]$ instead of $\mu[H = 0]$, but the constraints are once again linear inequalities on generalized moments of the optimization variable μ.

On the number of total evaluations on H. Recall that, for simplicity, we have assumed the response function G to be known and given by H. A large number of evaluations of H has been used in Table 5.2 to ensure convergence towards the global optimum. It is important to observe that those evaluations of H are not (actual, costly)
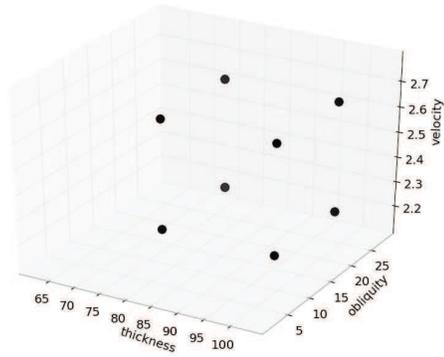
(a) support points at iteration 0
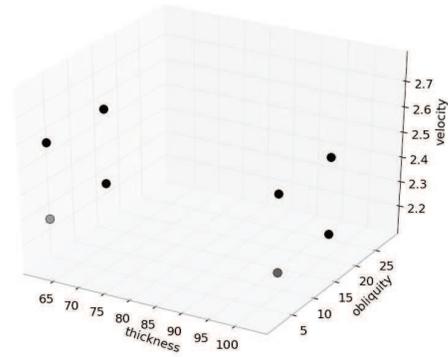
(b) support points at iteration 150
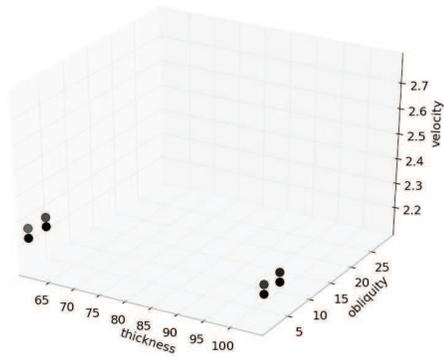
(c) support points at iteration 200
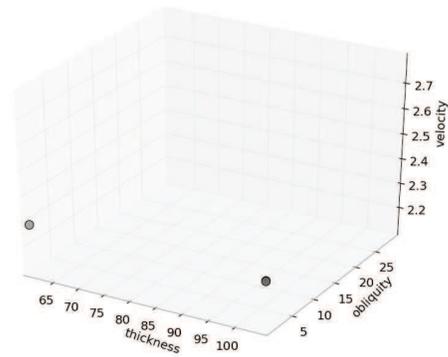
(d) support points at iteration 1000

Figure 5.1: For $\#\mathrm{supp}(\mu_i) \leq 2$, $i = 1, 2, 3$, the maximizers of Example 5.1 collapse to two-point support. Velocity and obliquity marginals each collapse to a single Dirac mass, while the plate thickness marginal collapses to have support on the extremes of its range. Note the perhaps surprising result that the probability of non-perforation is maximized by a distribution supported on the minimal, not maximal, impact obliquity.

33
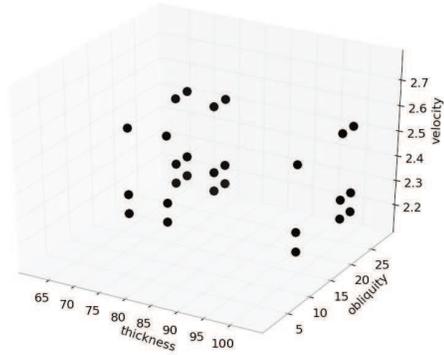
(a) support points at iteration 0
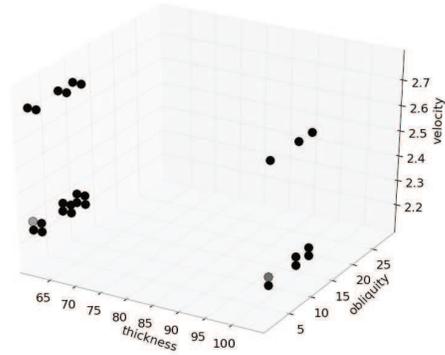
(b) support points at iteration 500
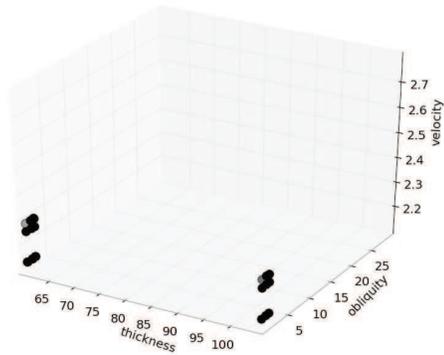
(c) support points at iteration 1000
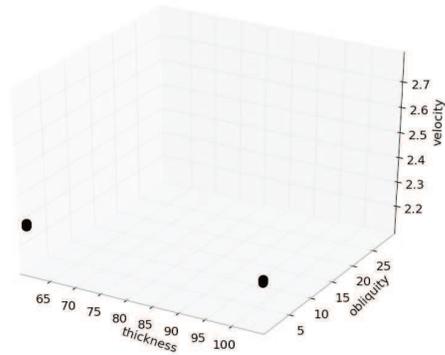

(d) support points at iteration 2155

Figure 5.2: For $\#\mathrm{supp}(\mu_i) \leq 3$, $i = 1, 2, 3$, the maximizers of Example 5.1 collapse to two-point support. Velocity, obliquity and plate thickness marginals collapse as in Figure 5.1.
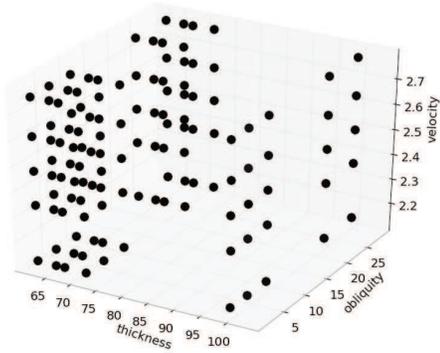
(a) support points at iteration 0
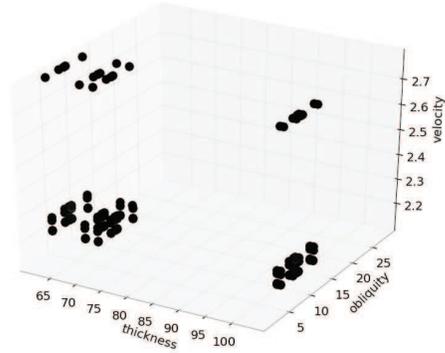
(b) support points at iteration 1000
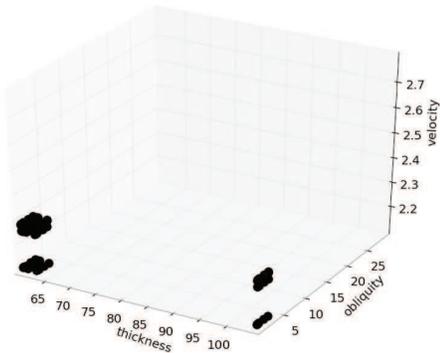
(c) support points at iteration 3000
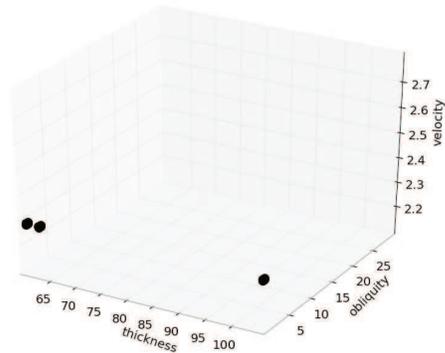
(d) support points at iteration 7100

Figure 5.3: For $\#\mathrm{supp}(\mu_i) \leq 5$, $i = 1, 2, 3$, the maximizers of Example 5.1 collapse to two-point support. Velocity, obliquity and plate thickness marginals collapse as in Figure 5.1. At iteration 7100, the thickness support point at 62.5 mils has zero weight, as can be seen in Figure 5.6.
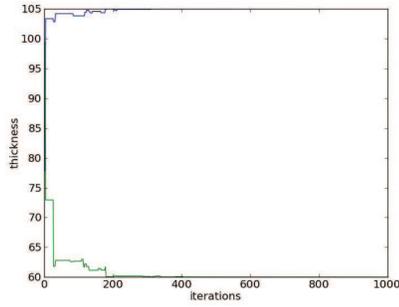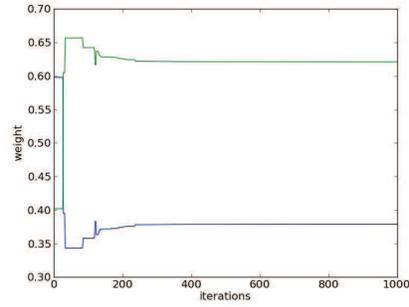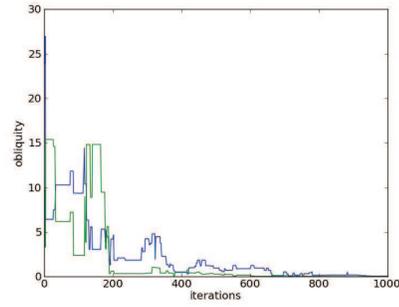
35

(a) convergence for thickness

(b) convergence for thickness weight

(c) convergence for obliquity
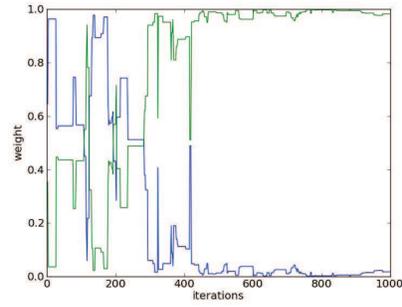
(d) convergence for obliquity weight
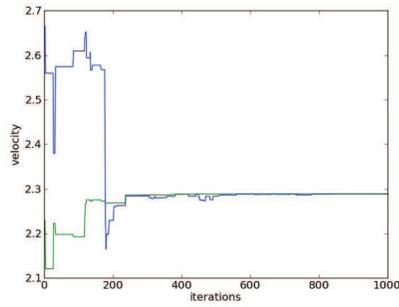
(e) convergence for velocity
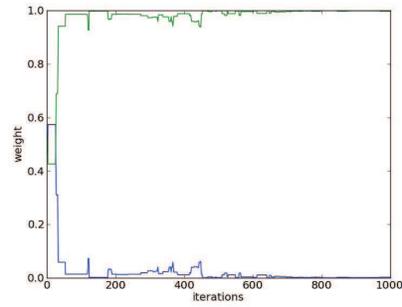
(f) convergence for velocity weight

Figure 5.4: Time evolution of the maximizers for Example 5.1 for $\#\mathrm{supp}(\mu_i) \leq 2$ for $i = 1, 2, 3$, as optimized by *mystic*. Thickness quickly converges to the extremes of its range, with a weight of 0.621 at 60 mils and a weight of 0.379 at 105 mils. The degeneracy in obliquity at 0 causes the fluctuations seen in the convergence of obliquity weight. Similarly, velocity converges to a single support point at 2.289 km$\cdot$s^{-1}, the ballistic limit velocity for thickness 105 mils and obliquity 0 (see (5.2)).
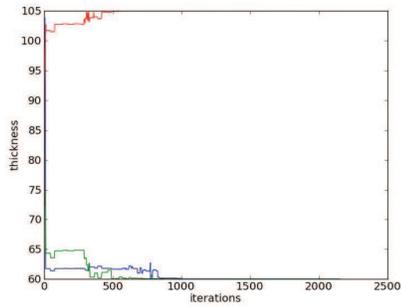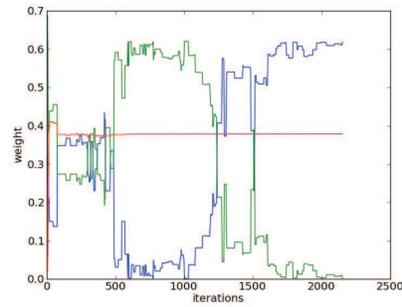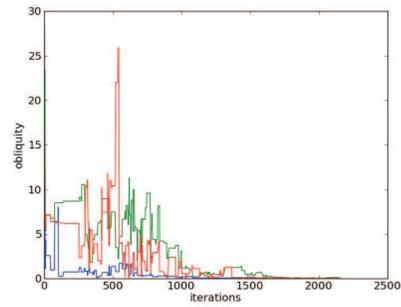
36

(a) convergence for thickness

(b) convergence for thickness weight

(c) convergence for obliquity
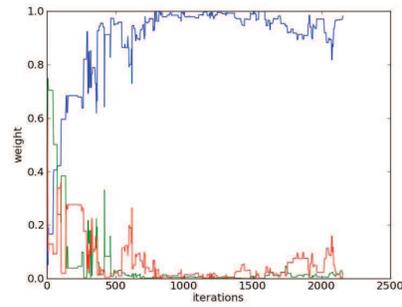
(d) convergence for obliquity weight
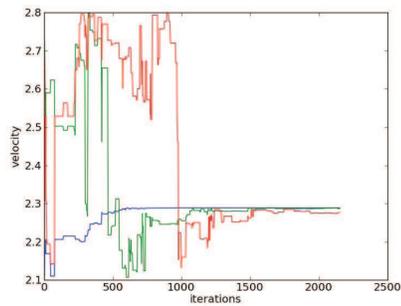
(e) convergence for velocity
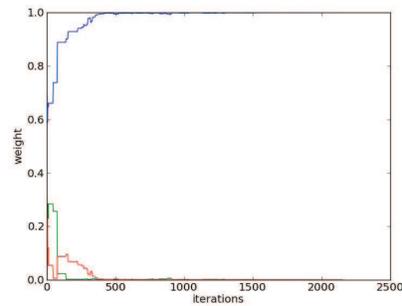
(f) convergence for velocity weight

Figure 5.5: Time evolution of the maximizers for Example 5.1 for $\#\text{supp}(\mu_i) \leq 3$ for $i = 1, 2, 3$, as optimized by *mystic*. Thickness quickly converges to the extremes of its range, with a weight of 0.621 at 60 mils and a weight of 0.379 at 105 mils. The degeneracy in thickness at 60 mils causes the fluctuations seen in the convergence of thickness weight. Obliquity and velocity each converge to a single support point, while the convergence of obliquity weight also demonstrates small fluctuations due to degeneracy.
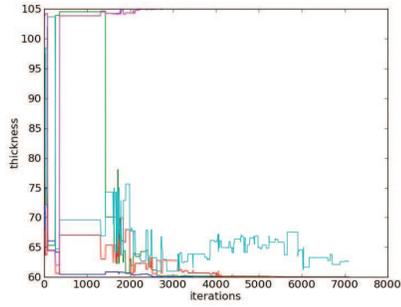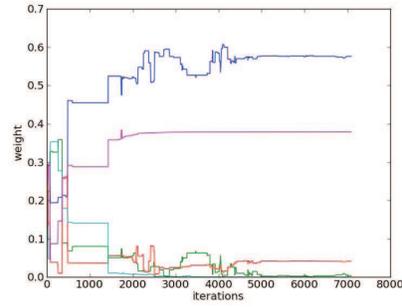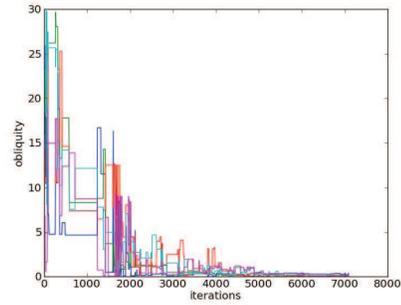
37

(a) convergence for thickness

(b) convergence for thickness weight

(c) convergence for obliquity
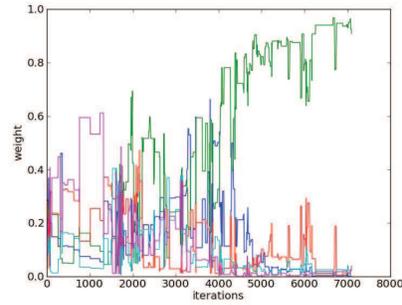
(d) convergence for obliquity weight
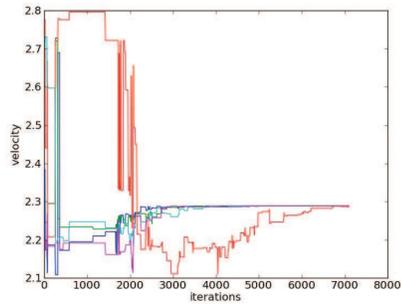
(e) convergence for velocity
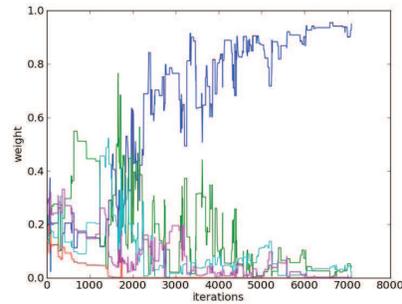
(f) convergence for velocity weight

Figure 5.6: Time evolution of the maximizers for Example 5.1 for #supp(μ_i) $\leq$ 5 for $i = 1, 2, 3$, as optimized by *mystic*. Four of the five thickness support points quickly converge to the extremes of its range, with weights 0.024, 0.058, and 0.539 at 60 mils and weight 0.379 at 105 mils. The thickness support point that does not converge to an extremum has zero weight. Obliquity and velocity each collapse to a single support point, again with the corresponding weights demonstrating fluctuations due to degeneracies.

38

	$J_{\mathrm{safe},\epsilon}(\Phi)$		$J_{\mathrm{unsafe},\epsilon}(\Phi)$	
	inf	sup	inf	sup
$\epsilon = 0.100$	0.000	1.000	0.000	0.900
iterations until numerical convergence	40	40	40	300
total evaluations of H	1,000	1,000	1,000	8,000
$\epsilon = 0.200$	0.000	1.000	0.000	0.800
iterations until numerical convergence	40	40	40	400
total evaluations of H	1,000	1,000	1,000	12,000
$\epsilon = 0.300$	0.000	1.000	0.000	0.599
iterations until numerical convergence	40	40	40	1000
total evaluations of H	1,000	1,000	1,000	33,000

Table 5.2: The results of the calculation of the intervals $J_{\mathrm{safe},\epsilon}(\Phi)$ and $J_{\mathrm{unsafe},\epsilon}(\Phi)$ for the functional $\Phi(\mu) := \mu[v \geq 2.45\,\mathrm{km} \cdot \mathrm{s}^{-1}]$. Note that, as the acceptable probability of system failure, ϵ, increases, the two intervals overlap less, so experimental determination of $\Phi(\mathbb{P})$ would be more likely to yield a decisive conservative certification of the system as safe or unsafe; the computational cost of this increased decisiveness is a greater number of function evaluations in the optimization calculations. All computational cost figures are approximate.

experiments but (cheap) numerical evaluations of equation (5.1). More precisely, the method for selecting new best experiments does not require new experiments; *i.e.*, it relies entirely on the information set $\mathcal{A}$ (which contains the information gathered from previous experiments). Hence those evaluations should not be *viewed as* "information gained from Monte Carlo samples" but as "pure CPU processing time". In situations where the numerical evaluation of H is expensive, one can introduce its cost in the optimization loop. An investigation of the best algorithm to perform the numerical optimization with the least number of function evaluations is a worthwhile subject but is beyond the scope of the present paper. Observe also that the method proposed in Section 2 does not rely on the exact knowledge of the response function G. More precisely, in situations where the response function is unknown, the selection of next best experiments is still entirely computational, and based upon previous data/information gathered on G enforced as constraints in a numerical optimization algorithm. More precisely, in those situations, the optimization algorithm may require the numerical evaluation of a large number of admissible functions f (compatible with the prior information available on G) but it does not require any new evaluation of G.

5.2   Coagulation–Fragmentation Algorithm for OUQ

The results of Sections 3 and 4 give explicit *a priori* bounds on the number of Dirac masses sufficient to find the lower and upper bounds $\mathcal{L}(\mathcal{A})$ and $\mathcal{U}(\mathcal{A})$ when the admissible

set $\mathcal{A}$ is given by independence and linear inequality constraints. However, it is possible that reduction properties are present for more general admissible sets $\mathcal{A}$. Can such "hidden" reduction properties be detected by computational means, even in the absence of theorems that prove their existence?

Consider again the results of the previous subsection. Theorem 3.1 provides an *a priori* guarantee that, to find $\mathcal{U}(\mathcal{A})$, it is sufficient to search the reduced feasible set $\mathcal{A}_\Delta$, which consists of those $\mu \in \mathcal{A}$ whose marginal distributions each have support on at most two points. However, Figures 5.2 and 5.3 provide numerical evidence that something much stronger is true: even if we search among measures $\mu \in \bigotimes_{i=1}^3 \Delta_k(\mathcal{X}_i)$ for $k \geq 1$, the measures collapse to an optimizer

$$\mu^* \in \Delta_1(\mathcal{X}_1) \otimes \Delta_0(\mathcal{X}_2) \otimes \Delta_0(\mathcal{X}_3)$$

(that is, two-point support on the thickness axis, and one-point support on the obliquity and velocity axes). In some sense, the (small support) optimizers are attractors for the optimization process even when the optimization routine is allowed to search over measures with larger support than that asserted by Theorem 3.1.

Therefore, we propose the following general algorithm for the detection of *hidden* reduction properties. Let an admissible set $\mathcal{A}$ be given; for $k \in \mathbb{N}$, let

$$\mathcal{A}_k := \{(f, \mu) \in \mathcal{A} \mid \mu \in \Delta_k(\mathcal{X})\}$$

be the collection of admissible scenarios such that μ has support on at most $k+1$ points of $\mathcal{X}$.

1. Fix any initial value of $k \in \mathbb{N}$.
2. Numerically calculate $\mathcal{U}(\mathcal{A}_k)$ and obtain an approximate optimizer $\mu^* \in \mathcal{A}_k$.
3. Calculate $\#\operatorname{supp}(\mu^*)$ and proceed as follows:
   - If $\#\operatorname{supp}(\mu^*) < k+1$, then the measure has *coagulated* to have less-than-maximally-sized support and we terminate the algorithm.
   - If $\#\operatorname{supp}(\mu^*) = k+1$, then no coagulation/reduction has yet been observed. We enter a *fragmentation* phase: replace k by any $k' > k$ and return to step 2.

Remark 5.1. It would be more accurate to say that the above algorithm is a sketch of an algorithm, and that its details should be adjusted to fit the circumstances of application. For example, if the admissible set $\mathcal{A}$ includes an independence constraint, then it would be appropriate to base decisions upon the cardinality of the support of the *marginal* distributions of μ^*, not on the cardinality of the support of μ^* itself. The termination of the algorithm if $\#\operatorname{supp}(\mu^*) < k+1$ is motivated by supposition that a *hidden* reduction property has been found and that $\mathcal{U}(\mathcal{A})$ has an (approximate) optimizer in $\mathcal{A}_k$.

Remark 5.2. We reiterate the point made in Remark 3.3 that these methods apply to more general situations than finite convex combinations of Dirac measures; Dirac measures are simply a well-known class of geometrically extreme probability measures, and can be replaced by the extremal points of any class of probability measures as

required by the situation of study. For example, if the OUQ problem of interest involved the invariant measures for some transformation $T \colon \mathcal{X} \to \mathcal{X}$, then each occurence of $\Delta_k(\mathcal{X})$ above would be replaced by

$$\mathcal{E}_k^T(\mathcal{X}) := \left\{ \sum_{j=0}^{k} \alpha_j \mu_j \ \middle| \ \begin{array}{c} \text{for each } j = 0, \ldots, k, \alpha_j \geq 0, \\ \mu_j \in \mathcal{M}(\mathcal{X}) \text{ is ergodic with respect to } T, \\ \text{and } \sum_{j=0}^{k} \alpha_j = 1 \end{array} \right\}.$$

5.3 The OUQ Algorithm in the *mystic* Framework

As posed above, OUQ at the high-level is a global optimization of a cost function that satisfies a set of constraints. This optimization is performed in *mystic* using the differential evolution algorithm of Price & Storn [42, 50], with constraints satisfied through a modified Lagrange multiplier method [34].

The *mystic* optimization framework [33] provides a collection of optimization algorithms and tools that lowers the barrier to solving complex optimization problems. Specifically, *mystic* provides flexibility in specifying the optimization algorithm, constraints, and termination conditions. For example, *mystic* classifies constraints as either "bounds constraints" (linear inequality constraints that involve precisely one input variable) or "non-bounds constraints" (constraints between two or more parameters), where either class of constraint modifies the cost function accordingly in attempt to maximize algorithm accuracy and efficiency. Every *mystic* optimizer provides the ability to apply bounds constraints generically and directly to the cost function, so that the difference in the speed of bounds-constrained optimization and unconstrained optimization is minimized. *Mystic* also enables the user to impose an arbitrary parameter constraint function on the input of the cost function, allowing non-bounds constraints to be generically applied in any optimization problem.

The *mystic* framework was extended for the OUQ algorithm. A modified Lagrange multiplier method was added to *mystic*, where an internal optimization is used to satisfy the constraints at each iteration over the cost function [34]. Since evaluation of the cost function is commonly the most expensive part of the optimization, our implementation of OUQ in *mystic* attempts to minimize the number of cost function evaluations required to find an acceptable solution. By satisfying the constraints within some tolerance at each iteration, our OUQ algorithm will (likely) numerically converge much more quickly than if we were to apply constraints by invalidating generated results (*i.e.* filtering) at each iteration. In this way, we can use *mystic* to efficiently solve for rare events, because the set of input variables produced by the optimizer at each iteration will also be an admissible point in problem space — this feature is critical in solving OUQ problems, as tens of thousands of function evaluations may be required to produce a solution. We refer to [34] for a detailed description of the implementation of the OUQ algorithm in the *mystic* framework.

Measures as data objects. Theorem 3.1 states that a solution to an OUQ problem, with linear constraints on marginal distributions, can be expressed in terms of products

41

of convex linear combinations of Dirac masses. In our OUQ algorithm, the optimizer's parameter generator produces new parameters each iteration, and hence produces new product measures to be evaluated within the cost function. For instance, the response function H, as defined by $H(h, \theta, v)$ in (5.1), requires a product measure of dimension $n = 3$ for support. In Example 5.1, the mean perforation area is limited to $[m_1, m_2] = [5.5, 7.5] \, \mathrm{mm}^2$, the parameters h, θ, v are bounded by the range provided by (5.3a), and products of convex combinations of Dirac masses are used as the basis for support. The corresponding OUQ code parameterizes the Dirac masses by their weights and positions.

More generally, it is worth noting that our computational implementation of OUQ is expressed in terms of methods that act on a hierarchy of parameterized measure data objects. Information is thus passed between the different elements of the OUQ algorithm code as a list of parameters (as required by the optimizer) or as a parameterized measure object. *Mystic* includes methods to automate the conversion of measure objects to parameter lists and vice versa, hence the bulk of the OUQ algorithm code (*i.e.* an optimization on a product measure) is independent of the selection of basis of the product measure itself. In particular, since the measure data objects can be decoupled from the rest of the algorithm, the product measure representation can be chosen to best provide support for the model, whether it be a convex combination of Dirac masses as required by Example 5.1, or measures composed of another basis such as Gaussians. More precisely, this framework can naturally be extended to Gaussians merely by adding covariance matrices as data object variables and by estimating integral moments equations (with a Monte Carlo method for instance).

6   Proofs

6.1   Proofs of Section 3

Proof of Theorem 3.1. For $\mu = \bigotimes_{i=1}^{m} \mu_i \in \mathcal{M}^{\mathbf{G}}$, consider the optimization problem

$$\begin{aligned} \text{maximize: } & \mathbb{E}_{(\mu'_1, \mu_2, \ldots, \mu_m)}[r], \\ \text{subject to: } & \mu'_1 \in \mathcal{M}(\mathcal{X}_1), \\ & \mathbf{G}(\mu'_1, \mu_2, \ldots, \mu_m) \le 0. \end{aligned}$$

By Fubini's Theorem,

$$\mathbb{E}_{(\mu'_1, \mu_2, \ldots, \mu_m)}[r] = \mathbb{E}_{\mu'_1}\left[\mathbb{E}_{(\mu_2, \ldots, \mu_m)}[r] \right],$$

where $\mathbb{E}_{(\mu_2, \ldots, \mu_m)}[r]$ is a Borel-measurable function on $\mathcal{X}_1$ and, for $j = 1, \ldots, n$, it holds that

$$\mathbb{E}_{(\mu'_1, \mu_2, \ldots, \mu_m)}[g'_j] = \mathbb{E}_{\mu'_1}\left[\mathbb{E}_{(\mu_2, \ldots, \mu_m)}[g'_j] \right]$$

where $\mathbb{E}_{(\mu_2, \ldots, \mu_m)}g'_j$ is a Borel-measurable function on $\mathcal{X}_1$. In the same way, we see that

$$\mathbb{E}_{(\mu'_1, \mu_2, \ldots, \mu_m)}[g^1_j] = \mathbb{E}_{\mu'_1}[g^1_j],$$

42

and, for $k = 2, \ldots, m$ and $j = 1, \ldots, n_k$, it holds that

$$\mathbb{E}_{(\mu_1', \mu_2, \ldots, \mu_m)}[g_j^k] = \mathbb{E}_{\mu_k}[g_j^k],$$

which are constant in μ_1'.

Since each $\mathcal{X}_i$ is Suslin, it follows that all the measures in $\mathcal{M}(\mathcal{X}_i)$ are regular. Consequently, by [51, Theorem 11.1], the extreme set of $\mathcal{M}(\mathcal{X}_i)$ is the set of Dirac masses. For fixed $(\mu_2, \ldots, \mu_m)$, let $\mathbf{G}_1 \subseteq \mathcal{M}(\mathcal{X}_1)$ denote those measures that satisfy the constraints $\mathbf{G}(\mu_1', \mu_2, \ldots, \mu_m) \leq 0$. Consequently, since for $k = 2, \ldots, m$ and $j = 1, \ldots, n_k$, $\mathbb{E}_{(\mu_1', \mu_2, \ldots, \mu_m)}[g_j^k]$ is constant in μ_1', it follows from [54, Theorem 2.1] that the extreme set $\mathrm{ex}(\mathbf{G}_1) \subseteq \mathcal{M}(\mathcal{X}_1)$ of the constraint set consists only of elements of $\Delta_{n_1 + n'}(\mathcal{X}_1)$. In addition, von Weizsäcker and Winkler [53, Corollary 3] show that a Choquet theorem holds: let μ' satisfy the constraints. Then

$$\mu'(B) = \int_{\mathrm{ex}(\mathbf{G}_1)} \nu(B) \, \mathrm{d}p(\nu)$$

for all Borel sets $B \subseteq \mathcal{X}_1$, where p is a probability measure on the extreme set $\mathrm{ex}(\mathbf{G}_1)$.

According to Winkler, an extended-real-valued function K on $\mathbf{G}_1$ is called *measure affine* if it satisfies the barycentric formula

$$K(\mu') = \int_{\mathrm{ex}(\mathbf{G}_1)} K(\nu) \, \mathrm{d}p(\nu).$$

When K is measure affine, [54, Theorem 3.2] asserts that

$$\sup_{\mu' \in \mathbf{G}_1} K(\mu') = \sup_{\nu \in \mathrm{ex}(\mathbf{G}_1)} K(\nu),$$

and so we conclude that

$$\sup_{\mu' \in \mathbf{G}_1} K(\mu') = \sup_{\nu \in \mathrm{ex}(\mathbf{G}_1)} K(\nu) \leq \sup_{\nu \in \Delta_{n_1 + n'}(\mathcal{X}_1) \cap \mathbf{G}_1} K(\nu).$$

However, since

$$\sup_{\nu \in \Delta_{n_1 + n'}(\mathcal{X}_1) \cap \mathbf{G}_1} K(\nu) \leq \sup_{\nu \in \mathbf{G}_1} K(\nu),$$

it follows that

$$\sup_{\mu' \in \mathbf{G}_1} K(\mu') = \sup_{\nu \in \Delta_{n_1 + n'}(\mathcal{X}_1) \cap \mathbf{G}_1} K(\nu).$$

To apply this result, observe that [54, proposition 3.1] asserts that the evaluation function

$$\mu_1' \mapsto \mathbb{E}_{\mu_1'} \left[\mathbb{E}_{(\mu_2, \ldots, \mu_m)}[r] \right]$$

is measure affine. Therefore,

$$\sup_{\substack{\mu_1' \in \mathcal{M}(\mathcal{X}_1) \\ \mathbf{G}(\mu_1', \mu_2, \ldots, \mu_m) \leq 0}} \mathbb{E}_{\mu_1', \mu_2, \ldots, \mu_m}[r] = \sup_{\substack{\mu_1' \in \Delta_{n_1 + n'}(\mathcal{X}_i) \\ \mathbf{G}(\mu_1', \mu_2, \ldots, \mu_m) \leq 0}} \mathbb{E}_{(\mu_1', \mu_2, \ldots, \mu_m)}[r]. \tag{6.1}$$

43

Now let $\varepsilon > 0$ and let $\mu_1^* \in \Delta_{n_1+n'}(\mathcal{X}_1)$ be ε-suboptimal for the right-hand side of (6.1): that is, $\mathbf{G}(\mu_1^*, \mu_2, \ldots, \mu_m) \leq 0$, and

$$\mathbb{E}_{(\mu_1^*, \mu_2, \ldots, \mu_m)}[r] \geq \sup_{\substack{\mu_1' \in \Delta_{n_1+n'}(\mathcal{X}_i) \\ \mathbf{G}(\mu_1', \mu_2, \ldots, \mu_m) \leq 0}} \mathbb{E}_{(\mu_1', \mu_2, \ldots, \mu_m)}[r] - \varepsilon.$$

Hence, by (6.1),

$$\mathbb{E}_{(\mu_1^*, \mu_2, \ldots, \mu_m)}[r] \geq \sup_{\substack{\mu_1' \in \mathcal{M}(\mathcal{X}_1) \\ \mathbf{G}(\mu_1', \mu_2, \ldots, \mu_m) \leq 0}} \mathbb{E}_{(\mu_1', \mu_2, \ldots, \mu_m)}[r] - \varepsilon \geq \mathbb{E}_{(\mu_1, \mu_2, \ldots, \mu_m)}[r] - \varepsilon.$$

Consequently, the first component of μ by can be replaced some element of $\Delta_{n_1+n'}(\mathcal{X}_1)$ to produce a feasible point $\mu' \in \mathcal{M}^{\mathbf{G}}$ without decreasing $\mathbb{E}[r]$ by more than ε. By repeating this argument, it follows that for every point $\mu \in \mathcal{M}^{\mathbf{G}}$ there exists a $\mu' \in \mathcal{M}_\Delta$ such that

$$\mathbb{E}_{\mu'}[r] \geq \mathbb{E}_\mu[r] - m\varepsilon.$$

Since ε was arbitrary the result follows. $\qquad\square$

Proof of Corollary 3.4. Simply use the identity

$$\mathcal{U}(\mathcal{A}) = \sup_{(f, \mu) \in \mathcal{A}} \mathbb{E}_\mu[r_f] = \sup_{f \in \mathcal{G}} \sup_{\substack{\mu \in \mathcal{M}^m(\mathcal{X}) \\ \mathbf{G}(f, \mu) \leq 0}} \mathbb{E}_\mu[r_f]$$

and then apply Theorem 3.1 to the inner supremum. $\qquad\square$

Proof of Theorem 3.6. Corollary 3.4 implies that $\mathcal{U}(\mathcal{A}) = \mathcal{U}(\mathcal{A}_\Delta)$ where,

$$\mathcal{A}_\Delta := \left\{ (f, \mu) \in \mathcal{G} \times \bigotimes_{i=1}^m \Delta_n(\mathcal{X}_i) \,\middle|\, \mathbb{E}_\mu[g_i \circ f] \leq 0 \text{ for all } j = 1, \ldots, n \right\}.$$

For each $i = 1, \ldots, m$, the indexing of the Dirac masses pushes forward the measure μ^i with weights $\alpha_k^i, k = 0, \ldots, n$ to a measure α^i on $\mathcal{N}$ with weights $\alpha_k^i, k = 0, \ldots, n$. Let $\alpha := \bigotimes_{i=1}^m \alpha^i$ denote the corresponding product measure on $\mathcal{D} = \mathcal{N}^m$. That is, we have a map

$$\mathbb{A} \colon \bigotimes_{i=1}^m \Delta_n(\mathcal{X}_i) \to \mathcal{M}^m(\mathcal{D})$$

and the product map

$$\mathbb{F} \times \mathbb{A} \colon \mathcal{G} \times \bigotimes_{i=1}^m \Delta_n(\mathcal{X}_i) \to \mathcal{F}_\mathcal{D} \times \mathcal{M}^m(\mathcal{D}).$$

Since for any function $g \colon \mathbb{R} \to \mathbb{R}$, we have $\mathbb{F}(g \circ f, \mu) = g \circ \mathbb{F}(f, \mu)$, it follows that for any $(f, \mu) \in \mathcal{F} \times \bigotimes_{i=1}^m \Delta_n(\mathcal{X}_i)$ that

$$\mathbb{E}_\mu[g \circ f] = \mathbb{E}_{\alpha_\mu}[\mathbb{F}(g \circ f, \mu)] = \mathbb{E}_{\alpha_\mu}[g \circ \mathbb{F}(f, \mu)].$$

44

Consequently, with the function $\mathcal{R}^{\mathcal{D}}\colon \mathcal{F}_{\mathcal{D}} \times \mathcal{M}^m(\mathcal{D}) \to \mathbb{R}$ defined by

$$\mathcal{R}^{\mathcal{D}}(h, \alpha) := \mathbb{E}_\alpha[r \circ h],$$

and for each $j = 1, \ldots, n$, the functions $\mathbf{G}_j^{\mathcal{D}}\colon \mathcal{F}_{\mathcal{D}} \times \mathcal{M}^m(\mathcal{D}) \to \mathbb{R}$ defined by

$$\mathbf{G}_j^{\mathcal{N}}(h, \alpha) := E_\alpha[g_i \circ h],$$

we have that, for all $(f, \mu) \in \mathcal{F} \times \bigotimes_{i=1}^m \Delta_n(\mathcal{X}_i)$,

$$\mathcal{R}(f, \mu) = \mathcal{R}^{\mathcal{D}}(\mathbb{F}(f, \mu), \alpha_\mu), \tag{6.2}$$

and, for all $j = 1, \ldots, n$ and all $(f, \mu) \in \mathcal{F}_{\mathcal{D}} \times \bigotimes_{i=1}^m \Delta_n(\mathcal{X}_i)$,

$$\mathbf{G}_j(f, \mu) = \mathbf{G}_j^{\mathcal{D}}(\mathbb{F}(f, \mu), \alpha_\mu), \tag{6.3}$$

where $\alpha_\mu := \mathbb{A}(\mu)$.

That is,

$$\mathcal{R} = \mathcal{R}^{\mathcal{D}} \circ (\mathbb{F} \times \mathbb{A}),$$

$$\mathbf{G}_j = \mathbf{G}_j^{\mathcal{D}} \circ (\mathbb{F} \times \mathbb{A}) \text{ for each } j = 1, \ldots, n.$$

Consequently, any $(f, \mu) \in \mathcal{A}_\Delta$ is mapped by $\mathbb{F} \times \mathbb{A}$ to a point in $\mathcal{F} \times \mathcal{M}^n(\mathcal{D})$ that preserves the criterion value and the constraint, and by the assumption must lie in $\mathcal{G}_{\mathcal{D}} \times \mathcal{M}^n(\mathcal{D})$. This establishes $\mathcal{U}(\mathcal{A}_\Delta) \leq \mathcal{U}(\mathcal{A}_{\mathcal{D}})$.

To obtain equality, consider $(h, \alpha) \in \mathcal{A}_{\mathcal{D}}$. By assumption, there exists an $(f, \mu) \in \mathcal{G} \times \bigotimes_{i=1}^m \Delta_n(\mathcal{X}_i)$ such that $\mathbb{F}(f, \mu) = h$. If we adjust the weights on μ so that $\mathbb{A}(\mu) = \alpha$, we still maintain $\mathbb{F}(f, \mu) = h$. By (6.2) and (6.3), this point has the same criterion value and satisfies the integral constraints of $\mathcal{A}_\Delta$. The proof is finished. $\square$

Proof of Proposition 3.7. Let $I := \mathbb{1}_{[a,\infty)}$ be the indicator function and consider $r_f := I \circ f$ so that $\mu[f \geq a] = \mathbb{E}_\mu[I \circ f]$. Since $I \circ f$ is integrable for all $\mu \in \mathcal{M}^m(\mathcal{X})$ and we have one constraint $\mathbb{E}_\mu[f] \leq 0$, the result follows from Theorem 3.6, provided that we have

$$\mathbb{F}\left(\mathcal{G} \times \bigotimes_{i=1}^m \Delta_1(\mathcal{X}_i)\right) = \mathcal{G}_{\mathcal{D}}.$$

To establish this, consider $f \in \mathcal{G}$ and observe that for all $\mu \in \bigotimes_{i=1}^m \Delta_1(\mathcal{X}_i)$ it holds that $\mathbb{F}(f, \mu) \in \mathcal{G}_D$. Therefore, we conclude that $\mathbb{F}(\mathcal{G} \times \bigotimes_{i=1}^m \Delta_1(\mathcal{X}_i)) \subseteq \mathcal{G}_{\mathcal{D}}$. On the other hand, for any $h \in \mathcal{G}_{\mathcal{D}}$, we can choose a measurable product partition of $\mathcal{X}$ dividing each $\mathcal{X}_i$ into 2 cells. We pull back the function h to a function $f \in \mathcal{F}$ that has the correct constant values in the partition cells, and place the Dirac masses into the correct cells. Set the weights to any nonzero values. It is easy to see that $f \in \mathcal{G}$. Moreover, we have a measure μ which satisfies $\mathbb{F}(f, \mu) = h$. Therefore, we conclude that $\mathbb{F}(\mathcal{G} \times \bigotimes_{i=1}^m \Delta_1(\mathcal{X}_i)) \supseteq \mathcal{G}_{\mathcal{D}}$. This completes the proof. $\square$

45

Proof of Theorem 3.8. First, observe that $\mathcal{G}_\mathcal{D}$ is a sub-lattice of $\mathcal{F}_\mathcal{D}$ in the usual lattice structure on functions. That is, if $h_1, h_2 \in \mathcal{G}_\mathcal{D}$, then it follows that both $\min(h_1, h_2) \in \mathcal{G}_\mathcal{D}$ and $\max(h_1, h_2) \in \mathcal{G}_\mathcal{D}$. Therefore, for any admissible $(h, \alpha) \in \mathcal{A}_\mathcal{D}$, it follows that clipping h at a to $\min(h, a)$ produces an admissible $(\min(h, a), \alpha)$ and does not change the criterion value $\alpha[h \geq a]$. Consequently, we can reduce to functions with maximum value a. Moreover, since each function h^s is in the sub-lattice $\mathcal{G}_\mathcal{D}$, it follows that $h^C \in \mathcal{G}_\mathcal{D}, C \in \mathcal{C}$. For $C \in \mathcal{C}$, define the sub-lattice

$$C_\mathcal{D} := \{h \in \mathcal{F}_\mathcal{D} \mid \{s \mid h(s) = a\} = C\}\}$$

of functions with value a precisely on the set C. Now, consider a function $h \in \mathcal{G}_\mathcal{D}$ such that $h \leq a$ and let C be the set where $h = a$. It follows that $h^C \leq h$, $h^C \in \mathcal{G}_\mathcal{D}$, and $h^C \in C_\mathcal{D}$. Since $h^C \leq h$ implies that $\mathbb{E}_\alpha[h^C] \leq \mathbb{E}_\alpha[h]$ for all α, it follows that replacing (h, α) by (h^C, α) keeps it admissible, and $\alpha[h^C \geq a] = \alpha[h \geq a]$. The proof is finished. $\qquad\square$

6.2   Proofs of Section 4

The proofs given in this subsection are direct applications of Theorem 3.8. In particular, they are based on an analytical calculation of (3.13). Because Proposition 4.6 is fundamental to all the other results of the section, its proof will be given first.

Proof of Proposition 4.6. When non-ambiguous, we will use the notation $\mathbb{E}[h^{C_0}]$ for $\mathbb{E}_\alpha[h^{C_0}]$ and $\mathbb{P}[h^{C_0} \geq a]$ for $\alpha[h^{C_0} \geq a]$. First, observe that, if $\sum_{j=1}^m D_j \leq a$, then $\min(h^{C_0}) \geq 0$, and the constraint $\mathbb{E}[h^{C_0}] \leq 0$ imply $\mathbb{P}[h^{C_0} = 0] = 1$. This proves the first equation of (4.9). Now, assume $a < \sum_{j=1}^m D_j$ and observe that

$$h^{C_0}(s) = a - \sum_{j=1}^m (1 - s_j) D_j.$$

It follows that

$$\mathbb{E}_\alpha[h^{C_0}] = a - \sum_{j=1}^m (1 - \alpha_j) D_j. \tag{6.4}$$

If $D_m = 0$, then the optimum is achieved on boundary of $[0, 1]^m$ (*i.e.* by taking $\alpha_m = 1$ since $C_0 = \{(1, \ldots, 1)\}$ and since h^{C_0} does not depend on s_m) and the optimization reduces to an $(m - 1)$-dimensional problem. For that reason, we will assume in all of the proofs of the results given in this section that all the D_is are strictly positive. The statements of those results remain valid even if one or more of the D_is are equal to zero.

The condition $D_m > 0$ implies that $\min(D_1, \ldots, D_m) > 0$ and that

$$\alpha[h^{C_0} \geq a] = \prod_{j=1}^m \alpha_j. \tag{6.5}$$

46

If the optimum in α is achieved in the interior of the hypercube $[0,1]^m$, then at that optimum the gradients of (6.4) and (6.5) are collinear. Hence, in that case, there exists $\lambda \in \mathbb{R}$ such that for all $i \in \{1,\dots,m\}$,

$$\frac{\prod_{j=1}^m \alpha_j}{\alpha_i} = \lambda D_i. \tag{6.6}$$

Since $\alpha[h^{C_0} \geq a]$ is increasing in each α_j, the optimum is achieved at $\mathbb{E}_\alpha[h^{C_0}] = 0$. Substitution of (6.6) into the equation $\mathbb{E}_\alpha[h^{C_0}] = 0$ yields that

$$\lambda = \frac{m \prod_{j=1}^m \alpha_j}{\sum_{j=1}^m D_j - a}$$

and, hence,

$$\alpha_i = \frac{\sum_{j=1}^m D_j - a}{m D_i}. \tag{6.7}$$

For all $i \in \{1,\dots,m\}$, the condition $0 < \alpha_i < 1$ is equivalent to $a < \sum_{j=1}^m D_j$ and

$$\sum_{j=1}^m D_j < a + m D_i. \tag{6.8}$$

It follows that for $\sum_{j=1}^m D_j - m D_m < a < \sum_{j=1}^m D_j$, the α defined by (6.7) lies in the interior of $[0,1]^m$ and satisfies

$$\alpha[h^{C_0} \geq a] = \frac{\left(\sum_{j=1}^m D_j - a\right)^m}{m^m \prod_{j=1}^m D_j}.$$

If $a \leq \sum_{j=1}^m D_j - m D_m$, then the optimum is achieved on boundary of $[0,1]^m$ (*i.e.* by taking $\alpha_m = 1$, since $C_0 = \{(1,\dots,1)\}$), and the optimization reduces to an $(m-1)$-dimensional problem.

To complete the proof, we use an induction. Observe in particular that, for $k \leq m-1$,

$$\frac{(\sum_{j=1}^k D_j - a)^k}{k^k \prod_{j=1}^k D_j} = \frac{(\sum_{j=1}^{k+1} D_j - a)^{k+1}}{(k+1)^{k+1} \prod_{j=1}^{k+1} D_j}$$

for $a = \sum_{j=1}^{k+1} D_j - (k+1) D_{k+1}$, and that

$$\frac{(\sum_{j=1}^k D_j - a)^k}{k^k \prod_{j=1}^k D_j} \leq \frac{(\sum_{j=1}^{k+1} D_j - a)^{k+1}}{(k+1)^{k+1} \prod_{j=1}^{k+1} D_j} \tag{6.9}$$

is equivalent to $a \geq \sum_{j=1}^{k+1} D_j - (k+1) D_{k+1}$. Indeed, writing $a = \sum_{j=1}^{k+1} D_j - (k+1) D_{k+1} + b$, equation (6.9) is equivalent to

$$\left(1 - \frac{b}{k D_{k+1}}\right)^k \leq \left(1 - \frac{b}{(k+1) D_{k+1}}\right)^{k+1}.$$

47

The function f given by $f(x) := \left(1 - \frac{y}{x}\right)^x$ is increasing in x (for $0 < y < x$): simply examine the derivative of $\log f$, and use the elementary inequality

$$\log(1 - z) + \frac{z}{1 - z} \geq 0 \text{ for } 0 < z < 1.$$

We will now give the outline of the induction. It is trivial to obtain that equation (4.9) is true for $m = 1$. Assume that it is true for $m = q-1$ and consider the case $m = q$. Equation (6.7) isolates the only potential optimizer α^q, which is not on the boundary of $[0,1]^q$ (not $(q-1)$-dimensional).

One obtains that equation (4.9) holds for $m = q$ by comparing the value of $\alpha[h^{C_0} \geq a]$ at locations α isolated by equations (6.7) and (6.8) with those isolated at step $q-1$. This comparison is performed via equation (6.9).

More precisely, if α^q (the candidate for the optimizer in α isolated by the previous paragraph) is not an optimum, then the optimum must lie in the boundary of $[0,1]^q$. Hence, the optimum must be achieved by taking $\alpha_i = 1$ for some $i \in \{1, \ldots, q\}$. Observing that $\mathcal{U}(\mathcal{A}_{C_0})$ is increasing in each D_i, and since $D_q = \min_{i \in \{1, \ldots, q\}} D_i$, that optimum can be achieved by taking $i = q$, which leads to computing $\mathcal{U}(\mathcal{A}_{C_0})$ with $(D_1, \ldots, D_{q-1})$, where we can use the $(q-1)$-step of the induction. Using equation (6.9) for $k = q-1$, we obtain that α^q is an optimum for $a \geq \sum_{j=1}^{q} D_j - qD_q$, and that, for $a \leq \sum_{j=1}^{q} D_j - qD_q$, the optimum is achieved by calculating $\mathcal{U}(\mathcal{A}_{C_0})$ with $q-1$ variables and $(D_1, \ldots, D_{q-1})$. This finishes the proof by using the induction assumption (see formula (4.9)). $\qquad\square$

The following two lemmas illustrate simplifications that can be made using the symmetries of the hypercube:

Lemma 6.1. *Let $C_0 \in \mathcal{C}$. If C_0 is symmetric with respect to the hyperplane containing the center of the hypercube and normal to the direction i, then the optimum of (4.8) can be achieved by taking $\alpha_i = 1$.*

Proof. The proof follows by observing that if C_0 is symmetric with respect to the hyperplane containing the center of the hypercube and normal to the direction i, then $h^{C_0}(s)$ does not depend on the variable s_i. $\qquad\square$

The following lemma is trivial:

Lemma 6.2. *Let (α, C) be an optimizer of (3.13). Then, the images of (α, C) by reflections with respect to hyperplanes containing the center of the hypercube and normal to its faces are also optimizers of (3.13).*

The proofs of the remaining theorems now follow in the order that the results were stated in the main part of the paper.

Proof of Theorem 4.1. The calculation of $\mathcal{U}(\mathcal{A}_C)$ for $m = 1$ is trivial and also follows from Proposition 4.6. $\qquad\square$
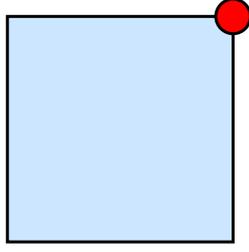
Figure 6.1: For $m = 2$, the optimum associated with $\mathcal{U}(\mathcal{A_C})$ can be achieved with $C = \{(1,1)\}$. For that specific value of C, the linearity of $h^C(s) = a - D_1(1-s_1) - D_2(1-s_2)$ implies $\mathcal{U}(\mathcal{A}_{\text{Hfd}}) = \mathcal{U}(\mathcal{A}_{\text{McD}})$.

Proof of Theorem 4.2. Write $C_1 = \{(1,1)\}$ (see Figure 6.1). Theorem 4.2 is a consequence of the following inequality:

$$\max_{C_0 \in \mathcal{C}} \mathcal{U}(\mathcal{A}_{C_0}) \leq \mathcal{U}(\mathcal{A}_{C_1}) \tag{6.10}$$

Assuming equation (6.10) to be true, equation (4.3) is obtained by calculating $\mathcal{U}(\mathcal{A}_{C_1})$ from Proposition 4.6 with $m = 2$. Let us now prove equation (6.10). Let $C_0 \in \mathcal{C}$; we need to prove that

$$\mathcal{U}(\mathcal{A}_{C_0}) \leq \mathcal{U}(\mathcal{A}_{C_1}). \tag{6.11}$$

By symmetry (using Lemma 6.2), it is no loss of generality to assume that $(1,1) \in C_0$. By Lemma 6.1 the optima for $C_0 = \{(1,1),(1,0)\}$ and $C_0 = \{(1,1),(0,1)\}$ can be achieved with C_1 by taking $\alpha_1 = 1$ or $\alpha_2 = 1$.

The case $C_0 = \{(1,1);(1,0);(0,1);(0,0)\}$ is infeasible.

For $C_0 = \{(1,1),(1,0),(0,1)\}$, we have $\mathbb{P}[h^{C_0} = a] = \beta$ and $\mathbb{E}[h^{C_0}] = a - (1 - \beta)\min(D_1, D_2)$ with $\beta = 1 - (1 - \alpha_1)(1 - \alpha_2)$ (recall that h^{C_0} is defined by equation (3.12)). Hence, at the optimum (in α),

$$\mathbb{P}[h^{C_0} = a] = \begin{cases} 1 - a/\min(D_1, D_2), & \text{if } a < \min(D_1, D_2), \\ 0, & \text{if } a \geq \min(D_1, D_2). \end{cases} \tag{6.12}$$

Equation (6.11) then holds by observing that one always has both

$$1 - \frac{a}{\min(D_1, D_2)} \leq 1 - \frac{a}{\max(D_1, D_2)}$$

and

$$1 - \frac{a}{\min(D_1, D_2)} \leq \frac{(D_1 + D_2 - a)^2}{4D_1 D_2}.$$

The last inequality is equivalent to $(D_1 - D_2 + a)^2 \geq 0$, which is always true. The case $C_0 = \{(1,1),(0,0)\}$ is bounded by the previous one since $\mathbb{P}[h^{C_0} = a] = \beta$ and $\mathbb{E}[h^{C_0}] = a\beta - (1 - \beta)\min(D_1, D_2)$ with $\beta = \alpha_1\alpha_2 + (1 - \alpha_1)(1 - \alpha_2)$. This finishes the proof. $\square$

(a) C_1 (b) C_2
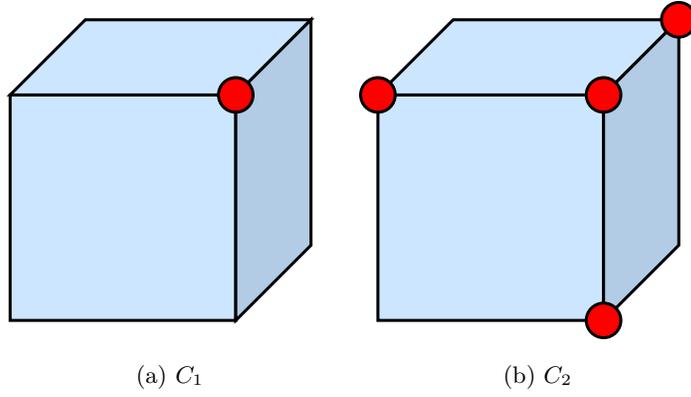
Figure 6.2: For $m = 3$, the optimum associated with $\mathcal{U}(\mathcal{A}_{\mathcal{C}})$ can be achieved with $C_1 = \{(1,1,1)\}$ (leading to $\mathcal{F}_1$) or $C_2 = \{(1,1,1),(0,1,1),(1,0,1),(1,1,0)\}$ (leading to $\mathcal{F}_2$). The linearity of $h^{C_1}(s) = a - D_1(1-s_1) - D_2(1-s_2) - D_3(1-s_3)$ implies that $\mathcal{U}(\mathcal{A}_{\mathrm{Hfd}}) = \mathcal{U}(\mathcal{A}_{\mathrm{McD}})$ when $\mathcal{F}_1 \geq \mathcal{F}_2$. Similarly, the nonlinearity of h^{C_2} leads to $\mathcal{U}(\mathcal{A}_{\mathrm{Hfd}}) < \mathcal{U}(\mathcal{A}_{\mathrm{McD}})$ when $\mathcal{F}_1 < \mathcal{F}_2$.

Proof of Theorem 4.4. Recall that

$$\mathcal{U}(\mathcal{A}_{\mathrm{McD}}) = \max_{C_0 \in \mathcal{C}} \mathcal{U}(\mathcal{A}_{C_0}).$$

It follows from Proposition 4.6 that $\mathcal{F}_1$ corresponds to $\mathcal{U}(\mathcal{A}_{C_1})$ with $C_1 = \{(1,1,1)\}$. Write $C_2 = \{(1,1,1),(0,1,1),(1,0,1),(1,1,0)\}$ (see Figure 6.2). Let us now calculate $\mathcal{U}(\mathcal{A}_{C_2})$ ($\mathcal{F}_2$ corresponds to $\mathcal{U}(\mathcal{A}_{C_2})$, which is the optimum, and it is achieved in the interior of $[0,1]^3$). We have $\mathbb{P}[h^{C_2} = a] = \alpha_2\alpha_3 + \alpha_1\alpha_3 + \alpha_1\alpha_2 - 2\alpha_1\alpha_2\alpha_3$, and

$$\mathbb{E}[h^{C_2}] = a - D_2(1-\alpha_1)(1-\alpha_2) - D_3\left((1-\alpha_2)(1-\alpha_3) + (1-\alpha_1)\alpha_2(1-\alpha_3)\right).$$

An internal optimal point α must satisfy, for some $\lambda \in \mathbb{R}$,

$$\alpha_2 + \alpha_3 - 2\alpha_2\alpha_3 = \lambda\left(D_2(1-\alpha_2) + D_3\alpha_2(1-\alpha_3)\right), \tag{6.13a}$$
$$\alpha_1 + \alpha_3 - 2\alpha_1\alpha_3 = \lambda\left(D_2(1-\alpha_1) + D_3\alpha_1(1-\alpha_3)\right), \tag{6.13b}$$
$$\alpha_1 + \alpha_2 - 2\alpha_1\alpha_2 = \lambda\left(D_3(1-\alpha_1\alpha_2)\right). \tag{6.13c}$$

If we multiply the first equation by α_1 and subtract the second equation multiplied by α_2, the we obtain that

$$(\alpha_1 - \alpha_2)\alpha_3 = \lambda D_2(\alpha_1 - \alpha_2),$$

which implies that either $\alpha_1 = \alpha_2$ or $\alpha_3 = \lambda D_2$.

Suppose that $\alpha_1 \neq \alpha_2$, so that $\alpha_3 = \lambda D_2$. Subtraction of the second equation in (6.13) from the first yields

$$(\alpha_2 - \alpha_1)(1 - 2\alpha_3) = \lambda\left(-D_2(\alpha_2 - \alpha_1) + D_3(\alpha_2 - \alpha_1)(1 - \alpha_3)\right),$$

50

which implies that either $\alpha_1 = \alpha_2$ or

$$1 - 2\alpha_3 = \lambda\left(-D_2 + D_3(1 - \alpha_3)\right).$$

Since $\alpha_3 = \lambda D_2$, this becomes

$$1 - \alpha_3 = \frac{\alpha_3 D_3}{D_2}(1 - \alpha_3),$$

which implies, since $\alpha_3 \neq 1$, that $\alpha_3 = \frac{D_2}{D_3}$. Therefore, $\lambda = \frac{1}{D_3}$. Therefore, the third equation in (6.13) becomes

$$\alpha_1 + \alpha_2 - 2\alpha_1\alpha_2 = \lambda\left(D_3(1 - \alpha_1\alpha_2)\right) = 1 - \alpha_1\alpha_2,$$

which amounts to

$$\alpha_1 + \alpha_2 - \alpha_1\alpha_2 = 1,$$

which in turn amounts to $\alpha_1(1 - \alpha_2) = 1 - \alpha_2$. Since $\alpha_2 \neq 1$, we conclude that $\alpha_1 = 1$, contradicting the supposition that α is an interior point. Therefore, $\alpha_1 = \alpha_2$ and equations (6.13), with $\alpha := \alpha_1 = \alpha_2$, become

$$\alpha + \alpha_3 - 2\alpha\alpha_3 = \lambda\left(D_2(1 - \alpha) + D_3\alpha(1 - \alpha_3)\right) \tag{6.14a}$$

$$2\alpha - 2\alpha^2 = \lambda\left(D_3(1 - \alpha^2)\right). \tag{6.14b}$$

Hence,

$$\mathbb{P}[h^{C_2} = a] = 2\alpha\alpha_3 + \alpha^2 - 2\alpha^2\alpha_3 \tag{6.15}$$

and

$$\mathbb{E}[h^{C_2}] = a - D_2(1 - \alpha)^2 - D_3\left((1 - \alpha^2)(1 - \alpha_3)\right).$$

The hypothesis that the optimum is not achieved on the boundary requires that

$$D_3, 0 < \alpha < 1, D_2 + D_3 > a \text{ and } E[h^{C_2}] = 0.$$

The condition $\mathbb{E}[h^{C_2}] = 0$ is required because equation (6.15) is strictly increasing along the direction $\alpha = \alpha_3$.

Suppose that those conditions are satisfied. The condition $\mathbb{E}[h^{C_2}] = 0$ implies that

$$1 - \alpha_3 = \frac{a}{D_3(1 - \alpha^2)} - \frac{D_2(1 - \alpha)}{D_3(1 + \alpha)},$$

which in turns implies that

$$\alpha_3 = 1 - \frac{a}{D_3(1 - \alpha^2)} + \frac{D_2(1 - \alpha)}{D_3(1 + \alpha)}. \tag{6.16}$$

Substitution of (6.16) into (6.15) yields that $\mathbb{P}[h^{C_2} = a] = \Psi(\alpha)$, with

$$\Psi(\alpha) = \alpha^2 + 2(\alpha - \alpha^2)\left(1 - \frac{a}{D_3}\frac{1}{(1 - \alpha^2)} + \frac{D_2}{D_3}\frac{1 - \alpha}{1 + \alpha}\right).$$

51

Hence,
$$\Psi(\alpha) = 2\alpha - \alpha^2 - 2\frac{a}{D_3}\frac{\alpha}{1+\alpha} + 2\frac{D_2}{D_3}\alpha\frac{(1-\alpha)^2}{1+\alpha}.$$

$\Psi(\alpha)$ can be simplified using polynomial division. In particular, using
$$\alpha\frac{(1-\alpha)^2}{1+\alpha} = (1-\alpha)^2 - \frac{(1-\alpha)^2}{1+\alpha},$$

$$\alpha\frac{(1-\alpha)^2}{1+\alpha} = \alpha^2 + 1 - 2\alpha - (1+\alpha) + 4 - \frac{4}{1+\alpha},$$

where the last step is obtained from
$$(1-\alpha)^2 = (\alpha + 1 - 2)^2 = (\alpha+1)^2 - 4(1+\alpha) + 4,$$

we obtain that
$$\Psi(\alpha) = 2\alpha - \alpha^2 - 2\frac{a}{D_3}\frac{\alpha}{1+\alpha} + 2\frac{D_2}{D_3}\left(4 + \alpha^2 - 3\alpha - \frac{4}{1+\alpha}\right).$$

Therefore,
$$\Psi(\alpha) = \alpha^2\left(2\frac{D_2}{D_3} - 1\right) - \frac{2a}{D_3}\frac{\alpha}{1+\alpha} + 2\alpha\left(1 - 3\frac{D_2}{D_3}\right) + 8\frac{D_2}{D_3}\frac{\alpha}{1+\alpha}$$

and
$$\Psi(\alpha) = \alpha^2\left(2\frac{D_2}{D_3} - 1\right) - 2\alpha\left(3\frac{D_2}{D_3} - 1\right) + \frac{\alpha}{1+\alpha}\left(8\frac{D_2}{D_3} - 2\frac{a}{D_3}\right). \qquad (6.17)$$

Equation (6.17) implies that
$$D_3\Psi'(\alpha) = 2\alpha(2D_2 - D_3) + 2(D_3 - 3D_2) - \frac{1}{(1+\alpha)^2}(2a - 8D_2).$$

The equation $\Psi'(\alpha) = 0$ is equivalent to equation (4.6). An interior optimum requires the existence of an $\alpha \in (0,1)$ such that $\Psi'(\alpha) = 0$ and $\alpha_3 \in (0,1)$, which leads to the definition of $\mathcal{F}_2$. This establishes the theorem for the $\mathcal{F}_2$ case.

Next, using symmetries of the hypercube and through direct computation (as in the $m = 2$ case), we obtain that
$$C_0 \neq C_2 \implies \mathcal{U}(\mathcal{A}_{C_0}) \leq \mathcal{U}(\mathcal{A}_{C_1}). \qquad (6.18)$$

For the sake of concision, we will give the detailed proof of (6.18) only for
$$C_3 = \{(1,1,1),(0,1,1),(1,0,1)\}.$$

This proof will give an illustration of generic reduction properties used in other cases. To address all the symmetric transformations of C_3, we will give the proof without assuming

that D_1, D_2 and D_3 are ordered. Let us now consider the C_3 scenario. If the optimum in α is achieved on the boundary of $[0,1]^3$, then equation (6.10) implies $\mathcal{U}(\mathcal{A}_{C_3}) \leq \mathcal{U}(\mathcal{A}_{C_1})$. Let us assume that the optimum is not achieved on the boundary of $[0,1]^3$. Observe that

$$h^{C_3}(s_1, s_2, 0) = h^{C_3}(s_1, s_2, 1) - D_3. \tag{6.19}$$

Combining (6.19) with

$$\mathbb{E}[h^{C_3}] = \alpha_3 \mathbb{E}[h^{C_3}(s_1, s_2, 1)] + (1 - \alpha_3)\mathbb{E}[h^{C_3}(s_1, s_2, 0)]$$

implies that

$$\mathbb{E}[h^{C_3}] = \mathbb{E}[h^{C_3}(s_1, s_2, 1)] - (1 - \alpha_3)D_3.$$

Furthermore,

$$\mathbb{P}[h^{C_3} = a] = \alpha_3 \mathbb{P}[h^{C_3}(s_1, s_2, 1) = a]. \tag{6.20}$$

Maximizing (6.20) with respect to α_3 under the constraint $\mathbb{E}[h^{C_3}] \leq 0$ leads to $\mathbb{E}[h^{C_3}] = 0$ (because $\mathbb{P}[h^{C_3} = a]$ and $\mathbb{E}[h^{C_3}]$ are linear in α_3) and

$$\alpha_3 = 1 - \frac{\mathbb{E}[h^{C_3}(s_1, s_2, 1)]}{D_3}. \tag{6.21}$$

Observe that the condition $\alpha_3 < 1$ requires $\mathbb{E}[h^{C_3}(s_1, s_2, 1)] > 0$. If $\mathbb{E}[h^{C_3}(s_1, s_2, 1)] \leq 0$ then $\alpha_3 = 1$, and the optimum is achieved on the boundary of $[0,1]^3$.

The maximization of $\mathbb{P}[h^{C_3}(s_1, s_2, 1) = a]$ under the constraint $\mathbb{E}[h^{C_3}(s_1, s_2, 1)] \leq E$ (where E is a slack optimization variable) leads to (using the $m = 2$ result)

$$\mathbb{P}[h^{C_3}(s_1, s_2, 1) = a] = 1 - \frac{(a - E)}{\min(D_1, D_2)}$$

if $a - E \leq \min(D_1, D_2)$, and $\mathbb{P}[h^{C_3}(s_1, s_2, 1) = a] = 0$ otherwise. It follows from (6.21) and (6.20) that if the optimum is achieved at an interior point, then the optimal value of $\mathbb{P}[h^{C_3} = a]$ is achieved by taking the maximum of

$$\mathbb{P}[h^{C_3} = a] = \left(1 - \frac{E}{D_3}\right)\left(1 - \frac{a - E}{\min(D_1, D_2)}\right)$$

with respect to E with the constraints $0 \leq E \leq D_3$ and $a - \min(D_1, D_2) \leq E$. If the optimum is not achieved on the boundary of $[0,1]^3$ then one must have

$$E = \frac{D_3 + \min(D_1, D_2) - a}{2},$$

which leads to

$$\mathbb{P}[h^{C_3} = a] = \frac{\left(D_3 + \min(D_1, D_2) - a\right)^2}{4D_3 \min(D_1, D_2)}. \tag{6.22}$$

Comparison of (6.22) and (4.3) implies that $\mathcal{U}(\mathcal{A}_{C_3}) \leq \mathcal{U}(\mathcal{A}_{C_1})$, by Proposition 4.6. $\quad\square$

Proof of Proposition 4.8. The idea of the proof is to show that h^C can be chosen so that C contains only one vertex of the hypercube, in which case we have the explicit formula obtained in Proposition 4.6.

First, observe that if $a > \sum_{j=1}^{m-1} D_j$, then it is not possible to satisfy the constraint $\mathbb{E}_\alpha[h^C] \leq 0$ whenever C contains two or more vertices of the hypercube. Next, if C contains two vertices s^1, s^2 of the hypercube, and the Hamming distance between those points is 1, then C is symmetric with respect to a hyperplane containing the center of the hypercube and normal to one of its faces, and the problem reduces to dimension $m-1$. It follows from Lemma 6.1 that the optimum of (4.8) can be achieved by a C that has only one element. If C contains two vertices of the hypercube, and the Hamming distance between those points is greater than or equal to 2, then the constraint $\mathbb{E}_\alpha[h^C] \leq 0$ is infeasible if $a > \sum_{j=1}^{m-2} D_j + D_m$ (because $h^C > 0$ in that case). Therefore, we conclude using Proposition 4.6. $\qquad\square$

Proof of Theorem 4.9. First, we observe that we always have

$$\mathcal{U}(\mathcal{A}_{\mathrm{Hfd}}) \leq \mathcal{U}(\mathcal{A}_{\mathrm{McD}}). \qquad (6.23)$$

We observe from equation (6.10) that the maximizer (h^C) of $\mathcal{U}(\mathcal{A}_{\mathrm{McD}})$ is linear (see Figure 6.1), and hence is also an admissible function under $\mathcal{U}(\mathcal{A}_{\mathrm{Hfd}})$. This finishes the proof. $\qquad\square$

Proof of Theorem 4.11. Just as for $m = 2$, equation (6.23) is always satisfied. Next, observing that $\mathcal{F}_1$, in Theorem 4.4, is associated with a linear maximizer h^C (see Figure 6.2), we deduce that

$$\mathcal{F}_1 \leq \mathcal{U}(\mathcal{A}_{\mathrm{Hfd}}) \leq \max(\mathcal{F}_1, \mathcal{F}_2).$$

This finishes the proof for equation (4.12). Let us now prove equation (4.13). Assuming that $\mathcal{U}(\mathcal{A}_{\mathrm{Hfd}}) = \mathcal{U}(\mathcal{A}_{\mathrm{McD}})$, it follows that $\mathcal{U}(\mathcal{A}_{\mathrm{McD}})$ can be achieved by a linear function h_0. Since at the optimum we must have $\mathbb{E}[h_0] = 0$, and since $\min(h_0, a)$ is also a maximizer of $\mathcal{U}(\mathcal{A}_{\mathrm{McD}})$, it follows that $\min(h_0, a) = h_0$. Using the linearity of h_0, and the lattice structure of the set of functions in $\mathcal{U}(\mathcal{A}_{\mathrm{McD}})$, we deduce that $h_0 = h^C$, where C contains only one vertex of the cube. It follows that $\mathcal{F}_1 \geq \mathcal{F}_2$. This finishes the proof. $\qquad\square$

7 Optimal Uncertainty Quantification with Sample Data

For the sake of clarity, we have started the description of OUQ with deterministic information and assumptions (when $\mathcal{A}$ is a deterministic set of functions and measures of probability). In this section, we will see that the addition of sample data to the available information and assumptions leads to non-trivial theoretical questions. As will be shown below, these questions are of practical importance beyond their fundamental connections with information theory and nonparametric statistics. In particular, we will see

that while the notion of an optimal bound (2.4) is transparent and unambiguous, the notion of an optimal bound on $\mathbb{P}[G(X) \geq a]$ in presence of sample data is not immediately obvious and should be treated with care.

By *sample data* we mean observations of a related random vector $\mathcal{D}$; for instance, samples of $(X, G(X))$ or $(X_1, X_4, G(X))$, or a subset of the entries of X and $G(X)$, or (possibly non-injective) functions of X and $G(X)$. The measurement of these samples may also be corrupted with noise, in which case the information set $\mathcal{A}$ must be modified to incorporate the information (or lack of information) on measurement noise.

Estimators. Let $\mathbb{D}$ be the sample data space and let $\mathcal{D}$ denote a random variable with values in $\mathbb{D}$. Consider the set of functions $\Psi : \mathbb{D} \to [0, 1]$, and fix $\delta > 0$. Let $\mathcal{C}_\delta$ denote the set of such functions Ψ such that for all $(f, \mu) \in \mathcal{A}$, with probability at least $1 - \delta$ on the sample data, we have $\mu[f \geq a] \leq \Psi(\mathcal{D})$. That is,

$$\mathcal{C}_\delta := \left\{ \Psi \;\middle|\; \inf_{(f,\mu)\in\mathcal{A}} \mu_{\mathcal{D}} \big[\mu[f(X) \geq a] \leq \Psi(\mathcal{D}) \big] \geq 1 - \delta \right\}. \tag{7.1}$$

For simplicity, we start our description where the law of the random vector $\mathcal{D}$ is determined by the measure $\mathbb{P}$. Thus, in this case, we write $\mu_{\mathcal{D}}$ for the law of the random vector $\mathcal{D}$ under the admissible scenario that $\mathbb{P} = \mu$. The discussion at the end of the experimental selection paragraph shows that the knowledge of $\mathbb{P}$ may not determine the law of the random vector $\mathcal{D}$, but only a set of possible values for that law. Indeed, $\mathcal{D}$ may contain sample data of random variables that are not functions of X, but only correlated with X. For those situations, our description can be generalized by, for instance, taking infima and suprema with respect to set of laws of the random vector $\mathcal{D}$ under the admissible scenario that $\mathbb{P} = \mu$. For conciseness, we will not develop this generalization in this paper. Consequently, by applying any $\Psi \in \mathcal{C}_\delta$ to the sample data $\mathcal{D}$ one obtains, with probability at least $1 - \delta$, an upper bound on the probability of failure $\mathbb{P}[G \geq a]$. We call the elements of $\mathcal{C}_\delta$ δ-*feasible estimators* (or simply *estimators*). The collection $\mathcal{C}_\delta$ is always non-empty, since the constant estimator $\Psi \equiv 1$ lies in $\mathcal{C}_\delta$. Since the constant estimator $\Psi \equiv 1$ is useless for practical decision-making purposes, we seek a "least" element of $\mathcal{C}_\delta$. However, the notion of optimality in the presence of sample data is not a straightforward question, nor is it a philosophical irrelevance, but it is a subtle issue of great practical importance.

Supplier/client certification conflicts. Consider a situation with conflicting estimators, *i.e.* a supplier is using a function $\Psi_1 \in \mathcal{C}_\delta$ to certify the safety of a design while the client is using another function $\Psi_2 \in \mathcal{C}_\delta$. Suppose that they both apply their estimators Ψ_1, Ψ_2 — each of which the user believes to be "optimal" in some sense — to the same sample data $\mathcal{D}$, and it transpires that

$$\Psi_1(\mathcal{D}) < \epsilon < \Psi_2(\mathcal{D}).$$

Hence, the supplier states that the design is safe while the client states that it is not! How can this conflict be resolved? This is only a relatively tame example: what if the

certification process is delegated to 10 independent firms, 6 of which say that the system is safe while 4 decide that the system is unsafe? Which, if any, of those conclusions can be trusted? One may trust all 10 certification "processes" but not all 10 "conclusions"/"outcomes" if they are in conflict. The conflict cannot be resolved simply by taking the minimum value between $\Psi_1(\mathcal{D}), \Psi_2(\mathcal{D})$, since it is not necessarily true that $\min(\Psi_1, \Psi_2) \in \mathcal{C}_\delta$. Should one try to construct a shared notion of optimality that "respects" the preferences (notions of optimality) of each of the suppliers and clients? In voting theory, it is well known that such shared notions of optimality are, in general, impossible to construct: Arrow's Theorem, [3] the Gibbard–Satterthwaite Theorem [14, 46], and the Duggan–Schwartz Theorem [11, 12] are all manifestations of this phenomenon. One of the standard ways to avoid these paradoxes of voting theory is to allow for a "dictator," *i.e.* a single preference that overrides all others. These paradoxes can also be avoided using type I and type II errors as defined in hypothesis testing; see the discussion below equation (7.5). We now present an criterion function that can be used to formulate the OUQ with sample data problem.

An agreed-upon risk function. One possible solution to the problems outlined above is for the client and suppliers to agree upon $\delta > 0$, and on a single scalar-valued risk function E to be used to quantify the optimality of estimators Ψ. For example, take

$$E(\Psi) := \sup_{(f,\mu)\in\mathcal{A}} \mathbb{E}_{\mu_\mathcal{D}}[\Psi(\mathcal{D})]. \tag{7.2}$$

There is admittedly, no objective reason why this "$\sup L^1$ norm" should be preferred to, say, any other "$\sup L^p$ norm." However, the agreement upon a risk function turns the OUQ-with-sample-data problem into a well-posed constrained optimization problem. In this case, we obtain the minimax problem

$$\text{minimize: } E(\Psi) := \sup_{(f,\mu)\in\mathcal{A}} \mathbb{E}_{\mu_\mathcal{D}}[\Psi(\mathcal{D})],$$
$$\text{subject to: } \inf_{(f,\mu)\in\mathcal{A}} \mu_\mathcal{D}\big[\mu[f(X) \geq a] \leq \Psi(\mathcal{D})\big] \geq 1 - \delta.$$

Clearly, this minimax problem, consisting of an averaging with respect to the law of the sample data, has an additional layer of complexity compared to the OUQ problem without sample data (2.4). As before, approximate optimizers always exist, and will in general be both necessary and sufficient for real applications, so we do not have to concern ourselves with the existence of true optimizers.

Connection with information theory Let $\hat{\mathcal{D}}$ be the empirical distribution associated with the sample data $\mathcal{D}$. If the samples are i.i.d, then, due to the law of large numbers, $\hat{\mathcal{D}}$ converges almost surely towards μ as n, the number of samples, gets large. Let $\mathcal{H}(\nu|\mu)$ be the relative entropy (Kullback–Leibler divergence) between two measures ν and μ. For each probability measure μ, define τ_μ as the smallest $\tau \geq 0$ such that

$$\mu_\mathcal{D}\left[\mathcal{H}(\hat{\mathcal{D}}|\mu) \geq \tau_\mu\right] \leq \delta.$$

Write $\mathcal{A}_\delta$ as the subset of couples $(f, \mu) \in \mathcal{A}$ such that $f(X^i) = G(X^i)$ for each observation X^i of the system, and $\mathcal{H}(\hat{\mathcal{D}}, \mu) \leq \tau_\mu$. Define

$$\Psi(\mathcal{D}) := \max_{(f,\mu) \in \mathcal{A}_\delta} \mu[f(X) \geq a]. \tag{7.3}$$

Observe that Ψ is an element of $\mathcal{C}_\delta$ and that the estimation of $\Psi(\mathcal{D})$ involves the same optimization as the one employed by OUQ in the absence of sample data (2.4). Here, the information contained in $\mathcal{D}$ is used to shrink the admissible set $\mathcal{A}$ via the relative entropy. Moreover, by [10], $\hat{\mathcal{D}}$ satisfies a large deviation property with rate function given by the relative entropy distance $\mathcal{H}(\cdot|\mu)$. That is, for large n, we obtain that the probability that $\hat{\mathcal{D}}$ is close to μ is of the order of $e^{-n\mathcal{H}(\hat{\mathcal{D}}|\mu)}$. Hence, for large n, the set $\mathcal{A}_\delta$ can be obtained by taking the intersection of $\mathcal{A}$ with the set of measures μ such that

$$\mathcal{H}(\hat{\mathcal{D}}|\mu) \leq \frac{1}{n} \log \frac{1}{\delta}. \tag{7.4}$$

Hence, equation (7.4) establishes a natural connection between OUQ and Information Theory. Moreover, using the lower bound of the large deviation property, it is possible to show that (7.3) is asymptotically optimal in the limit of large sample size.

Connection with hypothesis testing In equation (7.2), the notion of optimality has been defined by using $\mathbb{E}_{\mu_\mathcal{D}}[\Psi(\mathcal{D})]$; one could also have used the median or other quantile of μ. Why should we choose one definition of optimality and not another? This definition involves, a priori, a subjective choice and, as shown in the supplier-client conflict example, it is important to eliminate subjective steps in the certification process. Is it possible to relate the notion of optimality in OUQ to that of a framework that is already universally accepted and used? In this paragraph, we will show that it is possible to connect the notion of optimality in OUQ-based-certification (in presence of sampled data) to the notion of optimal tests in hypothesis testing.

Recall that in statistical hypothesis testing [27], there are two types of incorrect conclusions. A *type I error* (α) occurs when the hypothesis is inappropriately rejected and a *type II error* (β) occurs when we inappropriately fail to reject the hypothesis. A *uniformly most powerful test* (UMP test) is a hypothesis test which has the greatest power $1 - \beta$ among all possible tests of a given size α. For example, according to the Neyman–Pearson Lemma, the likelihood-ratio test is UMP for testing simple (point) hypotheses.

Define $\mathcal{A}_{\text{safe},\epsilon}$ and $\mathcal{A}_{\text{unsafe},\epsilon}$ as in (2.3). In this certification context, the (null) hypothesis can be defined as "the system is unsafe, *i.e.* $\mathbb{P}[G \geq a] > \epsilon$". Obtaining a uniformly most powerful test is equivalent to finding a test of significance level α (the probability of declaring an unsafe system safe is at most α), and minimizing β (the probability of declaring a safe system unsafe). Note that tests are necessarily functions of $\mathcal{A}$ and $\mathcal{D}$, and can be defined as: "reject if $\Psi(\mathcal{D}) > \epsilon$" and "accept if $\Psi(\mathcal{D}) \leq \epsilon$".

The notion of optimality in OUQ with sample data can then be formulated by the

following optimization problem for Ψ:

$$\text{minimize with respect to } \Psi: \quad \sup_{(f,\mu)\in\mathcal{A}_{\text{safe},\epsilon}} \mu_{\mathcal{D}}\left[\Psi(\mathcal{D}) > \epsilon\right], \qquad (7.5)$$

$$\text{subject to:} \quad \sup_{(f,\mu)\in\mathcal{A}_{\text{unsafe},\epsilon}} \mu_{\mathcal{D}}\left[\Psi(\mathcal{D}) \leq \epsilon\right] \leq \alpha.$$

Observe that if Ψ belongs to $\mathcal{C}_\alpha$, as defined by (7.1), then it is admissible for the optimization problem (7.5) (but the converse is not necessarily true). Observe also that the paradoxes emerging from the supplier/client conflict can be avoided by letting the client choose α (the bound on the probability that an unsafe system is declared safe) and letting the supplier solve (7.5). Indeed, those paradoxes result from a conflict of interest: it is in the interest of the client to purchase a safe system, and it is in the interest of the supplier to minimize cost and not declare a safe system unsafe. Hence, once a minimum level of safety α has been agreed upon, it is in the interest of the supplier to provide the sharpest estimator under the constraint imposed by α. This is why the optimization problem (7.5) provides a natural solution to that conflict.

Observe that OUQ "always" has approximate optima for every degree of "approximate," whereas it is known that this is not true for the UMP test [27]. One interpretation of this difference is that the UMP test requires superiority over all other tests for every element in the alternative hypothesis set, whereas OUQ is a minimax criterion. In the statistical literature, tests related to the OUQ formulation are described as minimax tests for composite hypothesis testing, and are generally used to produce robust tests. This difference represents one of the fundamental contributions of the OUQ viewpoint: namely, instead of doing theoretical analysis to describe the optimal test, the OUQ approach provides an optimization problem which must be solved — with the help of optimization theory, computer science, and applied mathematics — to determine the test.

8   Conclusions

The UQ Problem — A Problem with UQ?   The 2003 *Columbia* space shuttle accident and the 2010 Icelandic volcanic ash cloud crisis have demonstrated two sides of the same problem: discarding information may lead to disaster, whereas over-conservative safety certification may result in unnecessary economic loss and supplier-client conflict. Furthermore, while everyone agrees that UQ is a fundamental component of objective science (because, for instance, objective assertions of the validity of a model or the certification of a physical system require UQ), it appears that not only is there no universally accepted notion of the objectives of UQ, there is also no universally accepted framework for the communication of UQ results. At present, the "UQ problem" appears to have all the symptoms of an ill-posed problem; at the very least, it lacks a coherent general presentation, much like the state of probability theory before its rigorous formulation by Kolmogorov in the 1930s.

- At present, UQ is an umbrella term that encompasses a large spectrum of methods: Bayesian methods, Latin hypercube sampling, polynomial chaos expansions, stochastic finite-element methods, Monte Carlo, *etc.* Most (if not all) of these methods are characterized by a list of assumptions required for their application or efficiency. For example, Monte Carlo methods require a large number of samples to estimate rare events; stochastic finite-element methods require the precise knowledge of probability density functions and some regularity (in terms of decays in spectrum) for their efficiency; and concentration-of-measure inequalities require uncorrelated (or weakly correlated) random variables.
- There is a disconnect between theoretical UQ methods and complex systems of importance requiring UQ in the sense that the assumptions of the methods do not match the assumption/information set of the application. This disconnect means that often a specific method adds inappropriate implicit or explicit assumptions (for instance, when the knowledge of probability density functions is required for polynomial chaos expansions, but is unavailable) and/or the repudiation of relevant information (for instance, the monotonicity of a response function in a given variable) that the method is not designed to incorporate.

OUQ as an opening gambit. OUQ is not the definitive answer to the UQ problem, but we hope that it will help to stimulate a discussion on the development of a rigorous and well-posed UQ framework analogous to that surrounding the development of probability theory. The reduction theorems of Section 3, the Optimal Concentration Inequalities and non-propagation of input uncertainties of Section 4, the possibility of the selection of optimal experiments described at the end of Section 2, and the numerical evidence of Section 5 that (singular, *i.e.* low dimensional) optimizers are also attractors, suggest that such a discussion may lead to non-trivial and worthwhile questions and results at the interface of optimization theory, probability theory, computer science and statistics.

In particular, many questions raised by the OUQ formulation remain to be investigated. A few of those questions are as follows:
- Any (possibly numerical) method that finds admissible states $(f, \mu) \in \mathcal{A}$ leads to rigorous lower bounds on $\mathcal{U}(\mathcal{A})$. It is known that duality techniques lead to upper bounds on $(f, \mu) \in \mathcal{A}$ provided that the associated Lagrangians can be computed. Are there interesting classes of problems for which those Lagrangians can rigorously be estimated or bounded from above?
- The reduction theorems of Section 3 are limited to linear constraints on probability distribution marginals and the introduction of sample data may lead to other situations of interest (for instance, relative-entropy type constraints).
- Although general in its range of application, the algorithmic framework introduced in Section 5 is still lacking general convergence theorems.
- The introduction of sample data appears to render the OUQ optimization problem even more complex. Can this optimization problem be made equivalent to applying the deterministic setting to an information set $\mathcal{A}$ randomized by the sample data?

- In the presence of sample data, instead of doing theoretical analysis to describe the optimal statistical test, one formulation of the OUQ approach provides an optimization problem that must be solved to determine the test. Is this optimization problem reducible?

Acknowledgements

References

[1] C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis*. Springer, Berlin, third edition, 2006. A Hitchhiker's Guide.

[2] T. W. Anderson. The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. *Proc. Amer. Math. Soc.*, 6(2):170–176, 1955. http://dx.doi.org/10.2307/2032333.

[3] K. J. Arrow. A difficulty in the concept of social welfare. *J. Pol. Econ.*, 58(4):328–346, August 1950. http://dx.doi.org/10.1086/256963.

[4] R. E. Barlow and F. Proschan. *Mathematical Theory of Reliability*. SIAM, Philadelphia, 1996.

[5] V. Bentkus. A remark on the inequalities of Bernstein, Prokhorov, Bennett, Hoeffding, and Talagrand. *Liet. Mat. Rink.*, 42(3):332–342, 2002. http://dx.doi.org/10.1023/A:1020221925664.

[6] V. Bentkus. On Hoeffding's inequalities. *Ann. Probab.*, 32(2):1650–1673, 2004. http://dx.doi.org/10.1214/009117904000000360.

[7] V. Bentkus, G. D. C. Geuze, and M. C. A. van Zuijlen. Optimal Hoeffding-like inequalities under a symmetry assumption. *Statistics*, 40(2):159–164, 2006. http://dx.doi.org/10.1080/02331880600619085.

[8] D. Bertsimas and I. Popescu. Optimal inequalities in probability theory: a convex optimization approach. *SIAM J. Optim.*, 15(3):780–804 (electronic), 2005. http://dx.doi.org/10.1137/S1052623401399903.

[9] S. Boucheron, G. Lugosi, and P. Massart. A sharp concentration inequality with applications. *Random Structures Algorithms*, 16(3):277–292, 2000. http://dx.doi.org/10.1002/(SICI)1098-2418(200005)16:3<277::AID-RSA4>3.0.CO;2-1.

[10] A. Dembo and O. Zeitouni. Large deviations and applications. In *Handbook of Stochastic Analysis and Applications*, volume 163 of *Statist. Textbooks Monogr.*, pages 361–416. Dekker, New York, 2002.

[11] J. Duggan and T. Schwartz. Strategic manipulability is inescapable: Gibbard–Satterthwaite without resoluteness, 1992. Working Papers 817, California Institute of Technology, Division of the Humanities and Social Sciences.

[12] J. Duggan and T. Schwartz. Strategic manipulability without resoluteness or shared beliefs: Gibbard–Satterthwaite generalized. *Soc. Choice Welf.*, 17(1):85–93, 2000. http://dx.doi.org/10.1007/PL00007177.

[13] E. B. Dynkin. Sufficient statistics and extreme points. *Ann. Prob.*, 6(5):705–730, 1978. http://dx.doi.org/10.1214/aop/1176995424.

[14] A. Gibbard. Manipulation of voting schemes: a general result. *Econometrica*, 41(4):587–601, 1973. http://dx.doi.org/10.2307/1914083.

[15] H. J. Godwin. On generalizations of Tchebychef's inequality. *J. Amer. Statist. Assoc.*, 50(271):923–945, 1955. http://dx.doi.org/10.2307/2281177.

[16] W. Hoeffding. On the distribution of the number of successes in independent trials. *Ann. Math. Statist.*, 27(3):713–721, 1956. http://dx.doi.org/10.1214/aoms/1177728178.

[17] W. Hoeffding. The role of assumptions in statistical decisions. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. I*, pages 105–114, Berkeley and Los Angeles, 1956. University of California Press.

[18] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58(301):13–30, 1963. http://dx.doi.org/10.2307/2282952.

[19] K. Isii. On a method for generalizations of Tchebycheff's inequality. *Ann. Inst. Statist. Math. Tokyo*, 10(2):65–88, 1959. http://dx.doi.org/10.1007/BF02883863.

[20] K. Isii. The extrema of probability determined by generalized moments. I. Bounded random variables. *Ann. Inst. Statist. Math.*, 12(2):119–134; errata, 280, 1960. http://dx.doi.org/10.1007/BF01733120.

[21] K. Isii. On sharpness of Tchebycheff-type inequalities. *Ann. Inst. Statist. Math.*, 14(1):185–197, 1962/1963. http://dx.doi.org/10.1007/BF02868641.

[22] H. Joe. Majorization, randomness and dependence for multivariate distributions. *Ann. Probab.*, 15(3):1217–1225, 1987. http://dx.doi.org/10.1214/aop/1176992093.

[23] R. A. Johnson. Products of two Borel measures. *Trans. Amer. Math. Soc.*, 269(2):611–625, 1982. http://dx.doi.org/10.2307/1998470.

[24] S. Karlin and W. J. Studden. *Tchebycheff Systems: With Applications in Analysis and Statistics.* Pure and Applied Mathematics, Vol. XV. Interscience Publishers John Wiley & Sons, New York-London-Sydney, 1966.

[25] A. F. Karr. Extreme points of certain sets of probability measures, with applications. *Math. Oper. Res.*, 8(1):74–85, 1983. http://dx.doi.org/10.1287/moor.8.1.74.

[26] D. G. Kendall. Simplexes and vector lattices. *J. London Math. Soc.*, 37(1):365–371, 1962. http://dx.doi.org/10.1112/jlms/s1-37.1.365.

[27] E. L. Lehmann and J. P. Romano. *Testing Statistical Hypotheses.* Springer Texts in Statistics. Springer, New York, third edition, 2005.

[28] L. J. Lucas, H. Owhadi, and M. Ortiz. Rigorous verification, validation, uncertainty quantification and certification through concentration-of-measure inequalities. *Comput. Methods Appl. Mech. Engrg.*, 197(51-52):4591–4609, 2008. http://dx.doi.org/10.1016/j.cma.2008.06.008.

[29] A. W. Marshall and I. Olkin. Multivariate Chebyshev inequalities. *Ann. Math. Statist.*, 31(4):1001–1014, 1960. http://dx.doi.org/10.1214/aoms/1177705673.

[30] A. W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and its Applications*, volume 143 of *Mathematics in Science and Engineering.* Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1979.

[31] C. McDiarmid. On the method of bounded differences. In *Surveys in combinatorics, 1989 (Norwich, 1989)*, volume 141 of *London Math. Soc. Lecture Note Ser.*, pages 148–188. Cambridge Univ. Press, Cambridge, 1989.

[32] C. McDiarmid. Concentration. In *Probabilistic methods for algorithmic discrete mathematics*, volume 16 of *Algorithms Combin.*, pages 195–248. Springer, Berlin, 1998.

[33] M. McKerns, P. Hung, and M. Aivazis. Mystic: A simple model-independent inversion framework, 2009. http://dev.danse.us/trac/mystic.

[34] M. McKerns, H. Owhadi, C. Scovel, T. J. Sullivan, and M. Ortiz. The Optimal Uncertainty Quantification Algorithm in the *Mystic* Framework. http://dev.danse.us/trac/mystic, 2010.

[35] H. P. Mulholland and C. A. Rogers. Representation theorems for distribution functions. *Proc. London Math. Soc. (3)*, 8(2):177–223, 1958. http://dx.doi.org/10.1112/plms/s3-8.2.177.

[36] I. Olkin and J. W. Pratt. A multivariate Tchebycheff inequality. *Ann. Math. Statist*, 29(1):226–234, 1958. http://dx.doi.org/10.1214/aoms/1177706720.

[37] Committee on the Evaluation of Quantification of Margins, Uncertainties Methodology for Assessing, and Certifying the Reliability of the Nuclear Stockpile. *Evaluation of Quantification of Margins and Uncertainties Methodology for Assessing and Certifying the Reliability of the Nuclear Stockpile*. National Academies Press, Washington, D.C., 2009.

[38] P. M. Pardalos and S. A. Vavasis. Quadratic programming with one negative eigenvalue is NP-hard. *Journal of Global Optimization*, 1:15–22, 1991.

[39] R. R. Phelps. *Lectures on Choquet's Theorem*, volume 1757 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, second edition, 2001. http://dx.doi.org/10.1007/b76887.

[40] I. Pinelis. Exact inequalities for sums of asymmetric random variables, with applications. *Probab. Theory Related Fields*, 139(3-4):605–635, 2007. http://dx.doi.org/10.1007/s00440-007-0055-4.

[41] I. Pinelis. On inequalities for sums of bounded random variables. *J. Math. Inequal.*, 2(1):1–7, 2008.

[42] K. V. Price, R. M. Storn, and J. A. Lampinen. *Differential Evolution*. Natural Computing Series. Springer-Verlag, Berlin, 2005. A practical approach to global optimization, With 1 CD-ROM (Windows, Macintosh and UNIX).

[43] A. D. Rikun. A convex envelope formula for multilinear functions. *J. Global Optim.*, 10(4):425–437, 1997. http://dx.doi.org/10.1023/A:1008217604285.

[44] R. T. Rockafellar. Augmented Lagrange multiplier functions and duality in nonconvex programming. *SIAM J. Control*, 12(2):268–285, 1974. http://dx.doi.org/10.1137/0312021.

[45] S. M. Samuels. On a Chebyshev-type inequality for sums of independent random variables. *Ann. Math. Statist.*, 37(1):248–259, 1966. http://dx.doi.org/10.1214/aoms/1177699614.

[46] M. A. Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *J. Econom. Theory*, 10(2):187–217, 1975. http://dx.doi.org/10.1016/0022-0531(75)90050-2.

[47] C. Scovel and I. Steinwart. Hypothesis testing for validation and certification. *Journal of Complexity*, Submitted, 2010. http://www.ccs3.lanl.gov/group/papers/la-ur-10-02355.pdf.

[48] H. D. Sherali. Convex envelopes of multilinear functions over a unit hypercube and over special discrete sets. *Acta Math. Vietnam.*, 22(1):245–270, 1997.

[49] A. Srivastav and P. Stangier. Algorithmic Chernoff–Hoeffding inequalities in integer programming. *Random Structures Algorithms*, 8(1):27–58, 1996. http://dx.doi.org/10.1002/(SICI)1098-2418(199601)8:1<27::AID-RSA2>3.0.CO;2-T.

[50] R. M. Storn and K. V. Price. Differential evolution — a simple and efficient heuristic for global optimization over continuous spaces. *J. Global Optim.*, 11(4):341–359, 1997. http://dx.doi.org/10.1023/A:1008202821328.

[51] F. Topsøe. *Topology and Measure*. Lecture Notes in Mathematics, Vol. 133. Springer-Verlag, Berlin, 1970.

[52] L. Vandenberghe, S. Boyd, and K. Comanor. Generalized Chebyshev bounds via semidefinite programming. *SIAM Rev.*, 49(1):52–64 (electronic), 2007. http://dx.doi.org/10.1137/S0036144504440543.

[53] H. von Weizsäcker and G. Winkler. Integral representation in the set of solutions of a generalized moment problem. *Math. Ann.*, 246(1):23–32, 1979/80. http://dx.doi.org/10.1007/BF01352023.

[54] G. Winkler. Extreme points of moment sets. *Math. Oper. Res.*, 13(4):581–587, 1988. http://dx.doi.org/10.1287/moor.13.4.581.