Uncertainty Quantification Using Concentration-of-Measure Inequalities

Thesis by

Leonard J. Lucas

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2009

(Defended May 4, 2009)

To my family — Mom, Dad, Matt, and Jon.

Abstract

This work introduces a rigorous uncertainty quantification framework that exploits *concentration–of–measure* inequalities to bound failure probabilities using a well-defined certification campaign regarding the performance of engineering systems. The framework is constructed to be used as a tool for deciding whether a system is likely to perform safely and reliably within design specifications. Concentration-of-measure inequalities rigorously bound probabilities-of-failure and thus supply conservative certification criteria, in addition to supplying unambiguous quantitative definitions of terms such as margins, epistemic and aleatoric uncertainties, verification and validation measures, and confidence factors. This methodology unveils clear procedures for computing the latter quantities by means of concerted simulation and experimental campaigns. Extensions to the theory include hierarchical uncertainty quantification, and validation with experimentally uncontrollable random variables.

Acknowledgments

The creation of this work and its preceding works certainly has credit due to numerous individuals and groups of people during my time in graduate school. Without the support, motivation, and creative thinking of these people nearest to me, this work would not have come into fruition. First and foremost I would like to acknowledge my advisor at Caltech, Professor Michael Ortiz, and my laboratory advisor for the Naval Nuclear Propulsion Fellowship, Dr. David Jones. Professor Ortiz welcomed me into his group without knowledge of my existence, my research strengths, or guarantee of funds. Being welcomed into Michael's group was a vital step in my academic journey. Michael immediately routed my research to probability estimation of computational systems, a field I was unsure I would succeed in with my limited experience. The production of this work testifies to Michael's confidence and motivational skills — without these I would not think myself capable. His constant support and enthusiasm for my project were reassuring in times of doubt and meager results. David Jones, my laboratory mentor, advisor, and friend, presented a golden opportunity to work in a research-oriented environment at Bettis Atomic Power Laboratory for two summers. His anecdotes and calming demeanor cultivated enjoyment and purpose for my work and unstressed me when I pushed myself too hard — he reminded me to live life beyond work. Also I'd like to acknowledge Ron Coffield at Bettis for believing in my capabilities and opening the door for me to two amazing summers. The Naval Nuclear

Propulsion Fellowship fostered these relationships and the assistance provided fiscal security that promoted focused research complemented by invaluable working experience — I am indebted to those associated with granting this fellowship. Houman Owhadi, like an advisor to me, always so jovial, and happy to explain the latest UQ theory. Lydia Suarez, a free spirit who always had an ear and put up with me on so many levels, cared so much about my personal life — my mom on the west coast. Marta Kahl for miraculously making my computer woes disappear and providing the inner construction-worker in me with heavy lifting.

Without the very deep or extremely shallow conversations with Tamer Elsayed, Marcial Gonzalez, Ben Hansen, Yashashree Kulkarni, Bo Li, Julian Rimoli, Celia Romo, Gabriela Venturini, and Matias Zielonka, "the lab" would just be a room. Feras Habbal provided essential comedic relief, and companioned the thrill of a romp down the mountain trails on our bikes, twisties on the motorcycles, or the slopes on our snowboards, and shared the experience of bleeding like hell out of our noses. Ufuk Topcu for pushing the UQ so hard it was easy. The CSE group, especially Sharon Brunett and Julian Cummings for their patience and friendliness towards my supercomputing ignorance.

Rory Perkins, my confidante, travel companion, and pursuer of odd achievements - from climbing up a waterfall in Angeles Forest to our grand idea of riding down 10 miles of Highway 2 on adult-sized Heelys — he was always there for a listening ear or to have a drink when I really needed it. Misha Kitslitsyn, the adventurist who took me to the hidden trails to ride mountain bikes, my bodyboarding partner, and a world-class butterflyer in the pool. My roommates, Matt Whited, John Matson, and Alon Gorodetsky, though we quibbled over small things, it's the late night talks and social gatherings that kept us together and distinguished us as the "Stupid House." John stuck it out with me, and we can't recall the

number of times we climbed up the mountains on our bikes, bodyboarded at Manhattan Beach or on Los Robles. JoAnn Boyd, who took me under wing when I hobbled on crutches at the gym, and put miles on our rollerblades at Santa Monica, watched the Clarks play live together, and rode our hogs to unfamiliar territories. Damien Gates, my study and surfing companion — our qualifying exams fell in the wake of surfing. Ricky Lim and Daniel Sutoyo, my study buddies during our first year, who taught me how Caltech homework must be done, and techniques to celebrate when it is. Jorge Palamara, who always had a smile especially while doing a concocted triathlon workout around the Rose Bowl. Mike Krout, who helped me to discover the treasures of trails to rip down in Angeles Forest and leave nothing behind. Habib Ahmad, our oceanic photographer and a force on the go-karting track. Playmakers flag football team, for giving something to look forward to each week, and squeezing out unheralded victories for no material gain. Jon Young, the skinniest man in history that could manipulate any turns on his motorcycle or go-kart, especially during our wild trip to Monterrey where we camped with three dudes to a two-man tent. Ryan Zeidan, my sports-watching rock, always there to watch the big football or hockey game. Matt Eichenfeld for being like a big brother, with training, girls, and beer. Russ Ernst and Will Ford, true Pittsburghers to keep this 'burgher feeling at home. Caltech Rugby Football Club, for giving me hairs on my chest and memories that can never be reproduced in my lifetime. Alfredo Cabrera, Brett Hannah, Tim Murphy, Norberto & Carlyn Oropez, Rob Shaw, and Allen Lin for inducing me into the beauty and serenity of night riding up above LA, participating in the pain of riding our mountain bikes for 12 hours straight, and singlespeeding on the steepest of trails. Jeff Block and Jesse Arian, who became the best of training partners that made the last year go by in a second. Jon McCune, for introducing me to mountain biking, doing two 24-hour races, and winning an adventure race and the

whole Grassroots series in 2007, despite my best efforts to sabotage the race with fallible directions. Justin Samuels, David Rice, Jared Deible, Paul Raff, and Paul Harvin, quality individuals, who remain the heart of CMU crew and cultivating that unbreakable bond between hardened athletes. Neal Turbow, my senior buddy and cross-country partner. The hockey crew, especially Adrian Yuen, Damon Landau, Nando Duran, Brian Arguijo, and Rob Briceño, for making every Wednesday night an anticipated one, and jumpstarting my fruitless career as a goalie. The Pasadena Triathlon Club, for rocking my guts out on Saturdays with epic road rides and having the wackiest bike, run, and social outings imaginable. David Piper, for doing 126 miles on the road bike together in one day as two Pittsburghers against the world, and my CA injection into the Pittsburgh Triathlon. Amy Gouley for feeding a poor graduate student at every opportunity. Mud, Blood, and Beer, my mountain biking and beer drinking crew in Pittsburgh that kept the trails fresh, where falling was applauded, and the beer was cold. Patrick Barlas, Eric Parrish, and Cameron Dolbee for being dual-sport bikers and expert jack-of-all trades, who helped me pimp my motorcycle and find places to ride it, and who helped me explore my own home state on two wheels at 135 mph and flirt with death twice. Nick Krizan, Jeff Weiss, John Collinger, and Pasha Petite for sanity checks at work and for long lunches, and the longest lunch ever. Tom Black, Tom Damiani, Dave Hutula, and Jim Holliday for jovial ad-hoc conversations that kept the working days short. Richard Duong, my canyon carving buddy in the early years. The three of the fearsome foursome, Paul Richelmi, Paul Shatlock, and Lev Pinelis for always being there for me, and staying friends no matter the distance. Jon Weissman for doing a collegiate mountain bike race on our singlespeeds, not coming in last, catching the ice cream man and drinking beer as recovery. Darryl Zengler for the introduction to Cherry Canyon and the Inside Edition of PTC, and of course taking Weissman to the hospital after

the jump gone wrong.

Most of all, my family — Mom, Dad, Matt, and Jon. I couldn't have followed any closer in Matt's footsteps, either clearing his name behind him or trying to live up to it. The answers to life's questions were no match for his straightforward logic and simple tactics. Mom, for providing all the love a son could hope for, and supporting me through good and bad decisions. Dad, for being a companion in my outdoor habits. Jon, for keeping me rooted firmly to the ground when I get high and mighty.

Thank you.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivations

The use of probability and statistics in guiding engineering design/maintenance decisions is a maturing concept with increasing importance. For example, national defense issues such as maintaining a naval nuclear fleet have significant cost and security effects if components are not reliable enough to maintain life at sea. Concurrently, the national laboratories in the USA are renowned for their superior computing resources and advances in computational science and engineering. This strong computational backbone supports, among other duties, the stockpile stewardship effort to certify the nuclear stockpile sans the luxuries of nuclear testing in the post-test era beginning in 1991 [28]. The goal is to certify the safety and condition of the stockpile, which changes in time due to complicated nuclear physics. In light of this goal and the test restrictions, certification must be done in confidence. With above-ground non-critical experiments, historical data, and computational recourse available, the mission survives, but with hardship and complexity. Superior computational prowess allows for higher fidelity modeling and large deployments across the lab machines to assess the predictability of numerical models to extrapolate information about the future of the stockpile with confidence. It is in the engineers' hands to maximize the use of compu-

tational resources to ensure that components and systems will function reliably during their lifetime. This analysis is crucial for both the design and maintenance processes. Naturally, performance cannot be exactly predicted by today's computational methods, and unless infinite tests are carried out, cannot be known precisely through experiment alone. Thus engineers require computational tools that can objectively quantify the likelihood of desired performance to guide design and maintenance decisions. Probabilistic and statistical tools objectify questions addressing likelihood, thus a basic understanding of these fields is assumed of the reader.

A typical regimen for applying probabilistic techniques to engineering systems is to specify probability distributions to input parameters, pass them through a computational model, and compare the output distribution with experiment. With good agreement between computational and experimental results, one can know the likelihood of desired performance and can thus make engineering decisions with confidence. The previous statement makes use of subjective terms like "good" and "confidence", where disparate interpretations of these terms between experts may result in divergent decision-making agendas. That statement also assumes one knows the input probability distributions, i.e., that this is a parameter distributed according to a uniform or Gaussian distribution. And what of the bounds or mean and variance? Thus, some level of objectivity of the description of uncertainty must be met to be to applied to a model.

Completely deterministic engineering systems accept known inputs, $X = (X_1, ..., X_n)$, and yield a deterministic response, Y (this notational conventional will be used throughout this work). Unfortunately, real-world systems contain noise that perturb the response. Such noise sources can arise from inhomogeneities of the inputs, measurement error, or experimental error. These sources of noise can be accounted for by applying distributions

to input parameters. Once again, how certain can one be of the specified distribution? Also, when computationally modeling an engineering system, noise can be generated from computational round-off due to machine precision.

The rudimentary technique applied to engineering models to yield the likelihood of outcomes are Monte Carlo (MC) methods. These simulations produce a probability distribution of outcomes based on the distributions of the inputs using the computational model. MC methods are meant to estimate the likelihood of responses over the entire input range. A more complex question is to ask how often an engineering system becomes noncompliant, i.e., responds outside of the desired range of a specified performance measure. Typical techniques for this problem are MC methods, first- and second-order reliability methods (FORM and SORM respectively), and importance sampling. These techniques are increasingly complex sampling algorithms with (roughly) increasing accuracy and efficiency.

Given a computational model of a system, $Y = F(X)$, the response Y depends on inputs $X = (X_1, ..., X_n)$ which can be a mix of random and deterministic parameters. The foundation of this thesis relies on the understanding of the concentration-of-measure phenomenon (CoM) that provides a subsequent recourse to upper-bound the likelihood of a system performing poorly. It is suggested that the reader have a comprehensive knowledge of set theory and probability theory, both of which are critical towards understanding the CoM theory and its applications.

1.2 Uncertainty quantification

Uncertainty quantification (UQ) is the quantitative assessment of the uncertainties present in a system, which typically consists of a model that attempts to predict the real-world outcome of an engineering system. Methods such as Monte Carlo sampling, evidence the-

ory, and fuzzy logic provide varying approaches to perform UQ. This work is concerned with the application of *concentration-of-measure* inequalities to perform UQ regarding the performance of engineering systems. Specifically, this framework rests on the context of *certification* — a mechanism for deciding whether a system is likely to perform safely and reliably within design specifications.

The certification process is sometimes described in terms of quantification of margins and uncertainties (QMU) (e.g., [58], [28], [54]). Suppose that a system is required to perform at a certain degree, or *threshold*, for its acceptable (safe) operation. The system is designed so that its performance under a worst-case scenario of potential operating conditions is somewhat higher than its threshold. This idea is commonly regarded as a safety factor, e.g., the threshold load of a ladder is usually set lower than load it could take before failing in order to ensure that its failure never occurs. A suitable measure M of the distance between the worst-case and desired performance constitutes the *performance margin*. However, because the systems of interest—and the conditions they operate in—are subject to variability, performance measures are stochastic in nature. Therefore, the precise values of the expected performance level and its threshold are often uncertain. Uncertainty sources include operating conditions such as loads and physical characteristics including geometry and material behavior. Let U denote a suitable measure of these uncertainties regarding the performance measure, then the *confidence factor*

$$\mathrm{CF} = \frac{M}{U} \tag{1.1}$$

may be taken as a rational basis for certification. Thus, if CF is sufficiently larger than 1, the system may be regarded as safe and reliable and a candidate for certification. Interpreted

differently, the system's expected performance is far from failure with little variability for the given design specifications.

A first obvious strategy of attempting certification is by means of experimental testing and statistical sampling. This is indeed the only avenue of certification possible when no information is available *a priori* of the system behavior. However, the amount of tests required for purely empirical certification can be prohibitively high, especially in systems for which testing is expensive (see §2.2.5) in time and/or money, essentially rendering the empirical approach impractical. The problem is exacerbated when systems operate under extreme conditions outside the range of direct laboratory testing, e.g., large space structures and high-energy systems such as fusion reactors [10].

Under these conditions, numerical *modeling* can provide a worthwhile alternative to testing. Complex engineering systems can often be modeled through the application of sound physical and engineering principles, and the resulting numerical models can be used to accurately predict the performance of the system. In these cases, the key concern is how the availability of a model can be exploited in order to achieve certification with the least number of tests. Model-based certification reduces the number of tests required for certification if it provides sufficient *predictiveness*. The assessment of model predictiveness is typically accomplished through *verification and validation* (V&V), i.e., through a careful assessment of numerical modeling and solution errors.

The attractiveness of this scenario is obfuscated by the challenge of rendering it in *rigorous* and precise mathematical terms, including developing a computational toolbox enabling its efficient implementation. Rigorous UQ is meant specifically as a set of mathematically provable inequalities that provide rigorous upper bounds for the probability-of-failure of a system. While V&V and UQ have been the focus of extensive research in recent years

(e.g., [51], [58], [53], [28], [54]), a rigorous mathematical and computational framework just described is nowhere outlined in the literature. For instance, in the context of QMU rigorous definitions of M and U — and tractable means of computing them — are often left unspecified. The problem is compounded when system certification requires the quantification of uncertainties in multiple performance measures simultaneously. In these cases, *ad hoc* methods for amalgamating uncertainties, e.g., by root-mean squares, are often used without clear justification. A distinct link between CF and the probability-of-failure is omitted. Additionally, strict quantitative measures of model predictiveness, the series of V&V tests required for computing such measures, and the exact manner in which models supplement rigorous certification are also often absent, effectively reducing the value of QMU and model-based certification to that of a compelling but imprecise and heuristic conceptual framework.

The aim of this thesis is to develop, apply, and extend a rigorous theory of uncertainty quantification and certification in which the desired probability-of-failure upper bounds are supplied through *concentration-of-measure* inequalities. Loosely, the concentration-of-measure phenomenon is a consequence of the phenomenon that functions of a large number of variables, i.e., functions in high-dimensional spaces, tend to exhibit small local oscillations with respect to each variable tend to be nearly constant. Moreover, fluctuations can be controlled through elementary yet powerful and nontrivial quantitative inequalities called concentration-of-measure inequalities (see [32] for a monograph, [7] for a survey). These tools have found broad applications in functional analysis, complexity theory, probability, and statistics. However, the application of concentration-of-measure inequalities to perform uncertainty quantification lacks to date.

Concentration-of-measure inequalities applied to uncertainty quantification intend to

supply: probability-of-failure upper bounds resulting in rigorous and conservative certification; precise definitions and quantitative measures of margins and uncertainties enabling such certification; a distinct relationship between the confidence factor and probability-of-failure; a rigorous framework for model-based certification, including precise quantitative measures of model predictiveness and the effect of *unknown unknowns*; and the series of verification and validation tests required for assessing model predictiveness. Prior to formally presenting these developments, UQ is presented below as it has evolved, with an emphasis of using the lexicon adopted by the national laboratories. Specifically, a review of quantification of margins and uncertainties (QMU) and verification and validation (V&V) follows with an introduction to concentration-of-measure in the next section. These descriptions are provided for the reader to become acclamated to the current lexicon of uncertainty quantification. It will be noted that this work's lexicon of verification and validation do not align exactly with those presented here — the quantification of verification and validation lie in rigorously defined values denoted as "diameters".

1.2.1 Quantification of margins and uncertainties

The quantification of margins and uncertainties (QMU) is a formalism for evaluating confidence in estimating reliability of complex engineering systems ([28])—a distinct predecessor to this work on concentration-of-measure inequalities for uncertainty quantification. QMU came into being after underground testing (UGT) of nuclear weapons ceased in 1991, and the effort to stockpile stewardship became science-based (SBSS). Its ebbing history to present day is outlined in [28], clearly stating that no universally accepted definition of QMU exists, but identifies advances towards a universally accepted formalism for UQ that is needed to instill confidence in decision-makers. They note that some forms of QMU are not statistically

based; some use or dismiss subjectivity of expert opinion; and that QMU may ultimately "drive [engineers] into an endless search for impossibly quantitative answers to ill-posed questions". QMU is "not itself a procedure for certifying..., but the formal framework it provides, along with the common language for all parties involved, can be very beneficial to that process". This work, besides providing unambiguous rigor to UQ, also benefits from bridging the formalism gap between this form of UQ and total system certification. The proceeding is a presentation of the formalism of QMU defined in [28].

QMU relies on three basic functions — metric definition, developing *gates*, and evaluating uncertainty [58]. Gates formalize the measure of certification through propagation of uncertainty and allowable margins at the component level. To begin, consider the first-order Taylor expansion of the function $F(x_1, \ldots, x_n)$ in each of its variables:

$$\delta F_i = \frac{\partial F}{\partial x_i} \delta x_i \tag{1.2}$$

where uncertainties are accumulated at the component level in some normed sense:

$$\delta F = \sqrt{\sum_{i=1}^{N} \left(\frac{\partial F}{\partial x_i} \delta x_i \right)^2}. \tag{1.3}$$

The method presented in this work for defining a rigorous UQ methodology borrows much of its lexicon and assimilates many of its definitions to QMU. Certification criteria, margins, and uncertainties are defined for both methods, but the computation and interpretation of these values vary significantly enough to distinguish the methods.

Fundamental assertions of accuracy must be made of the model and between a model and the real world — precisely the gaol of V&V. These tools are used to certify systems in both QMU and UQ via CoM frameworks.

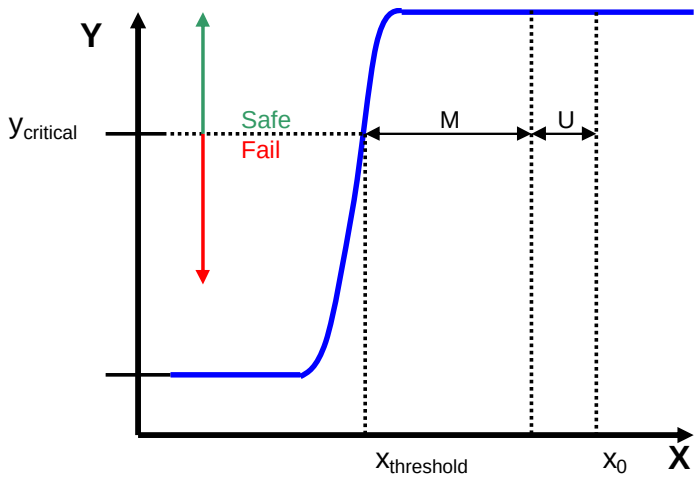


Figure 1.1: QMU definitions marked up on a typical "cliff" chart

1.2.2 Verification, validation, and certification

V&V has been defined in many ways qualitatively, but the lack of a firm quantitative description has propagated many such acceptable definitions in the literature. This obfuscation has not qualitatively affected the agreed meaning of certification, which is solely dependent upon passing both verification and validation measures of success, however they may be defined. For example, [52] briefly describes the evolution of the definitions for V&V through a multitude of engineering guideline development organizations such as IEEE, ANS, DoD, AIAA, and ASME. The latest accepted terminology for V&V has evolved to:

Verification: The process of determining that a model implementation accurately represents the developer's conceptual description of the model and the solution to the model.

Validation: The process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model.

Performing V&V succinctly strictly from these definitions could imaginably lead to strikingly different V&V campaigns between independent parties.

1.2.2.1 Verification

The premise of verification is to evaluate a numerical model's ability to accurately and repeatably produce a solution to the real world application it is intended to model — a task that bears much emphasis on correctness of the numerical implementation at every possible level. Firstly, verification can be divided into code and solution verification.

Code verification covers numerical algorithm verification and software quality engineering (SQE) [52]. A developer of a complex code must demonstrate that the numerical algorithms are implemented correctly and produce expected results, which is a natural request for computational codes that continually grow in complexity and must critically assessed. Secondly, SQE is the process of asserting that the architectures of machines being used have consistent systems operations and source compiling software.

Solution verification pertains to the numerical accuracy of the solution of the partial differential equations (PDEs), which allow for such complex codes to discretize a general system, thus providing numerical tractability. Verification benchmarks can serve this purpose with simple geometries and loading cases where perhaps analytical solutions can be verified against the numerical solution. The two ways this numerical error estimation can be approached is through *a priori* and *a posteriori* error approximation methods ([52]). *A priori* methods strictly employ knowledge of the numerical algorithm which approximates the PDE operators and the corresponding initial and boundary conditions. *A posteriori* error approximation uses all of the *a priori* information, as well as order-of-accuracy methods.

The collection of all these tasks comprises one accepted flavor of verification. Indeed they are each powerful and provide insight to the accuracy of the code and its solution. Although this work does not explicitly define verification as above, basic attempts have been made to have code, algorithmic, and solution correctness, as well as repeatability and

reproducibility. Lumped together, these provide a model that is acceptable to perform the uncertainty quantification attempted in this work.

The concentration-of-measure interpretation of verification adds another layer for the verification tract, relying on these numerical error assessments above to ensure that a code is formally prepared to be exercised for its intended purpose, but more importantly to define a rigorous way in making sense of the uncertainties that affect the model and how they affect the occurrence of interesting events. In the applications of this methodology, the impetus was on expedient attempts at tractable methodologies while the strict numerical accuracy checks were left to be done at a later time. Therefore, one will not see these strict numerical accuracy checks as defined by the national laboratories' lexicon of verification in this work. Current work on verification using benchmark examples in the UQ via CoM framework are being developed, but will not appear in this work.

1.2.2.2 Validation

The premise of validation is to compare the accuracy of a model against its real world counterpart. Institutions that attempt to define and use validation as a decision tool differ in the accepted activities that comprise the validation campaign. For example, the AIAA guide ([1]) terms validation as only the assessment of model accuracy in comparison to the associated experiment. In addition, [2] defines validation to include addressing the issue of using the computational model to extrapolate or interpolate conditions towards the intended use of the model, and assessing the estimated accuracy against pre-specified accuracy requirements in the intended range of use. These additions are earmarked as predictive capability tasks in [1]. Validation experiments, sometimes referred to as "proof tests" or "system performance tests" ([52]) are inherently different from scientific discovery or calibration ex-

periments, where phenomena are being explored or physical parameters are being probed, respectively. Thus validation experiments deliver predictive capability information to the computational customer through intimate interworkings between computationalists and experimentalists. This multidisciplinary, cross-communication culture identifies precedents needed for clearly defined validation terminology for all parties involved.

It should be noted that among the various lexicography for validation ([55]), one should be aware of the issue of using data for validation and calibration: "Validation tests against independent data that have not also been used by calibration are necessary in order to be able to document the predictive capability of a model." A separation of available data, whether historical or readily sampled, must be made for calibration and prediction, otherwise predictiveness can only be made for the calibration range, a self-defeating purpose for assessing predictiveness of a model.

The heart of validation is a comparison between model and experimental performance measures, but what should a comparison of the two entail? It is suggested that high-quality validation metrics ([52]) be defined to assert confidence in the comparison, which is an active issue in the literature ([69]).

In this work, a validation campaign is made by clearly defining its goals and expected results, accounting for experimental limitations, allowing the supposed predictive code to identify sensitive parameters that may affect experimental capabilities, and quantitatively computing a validation measure in terms of diameters through rigorous steps guided by optimization algorithms, all based upon the concentration-of-measure phenomenon. This is one particular way of doing validation which places emphasis on unambiguous performance quantities and rigorous methodologies to compute them, with enough flexibility to address specific computational and experimental scenarios. For example, in hypervelocity

impact experiments using light gas guns, velocity cannot be controlled precisely, and the concentration-of-measure inequalities that guide the validation campaign account for uncontrollable parameters in design space. Also, this framework allows for trivial validation metrics, i.e., the comparison between nominally similar computations and experiments is simply their difference.

1.2.2.3 Certification

Certification, in the scope of this work, is the unambiguous measure of the ratio of margins M to uncertainties U in a system, where uncertainties are computed through verification and validation campaigns in the form of "diameters". Certification according to QMU is defined through each portal of the system, or component-wise, while the concentration-of-measure-based UQ framework is system-wise with the applicable extension to component-wise certification (§4.2).

At the end of the day, certification through UQ techniques provides a number or numbers for which decisions are to be made from accordingly. Expert opinion, even in this complete objective UQ framework, must play a role in decision-making. Certification requires confidence of a system to perform within its desired operating specifications given the conditions it is likely (or unlikely) to experience, which can be gained from verification and validation assessments. Though certification has a rigorous definition, confidence is an intrinsic value with different weighting scales between decision-makers, and has a fuzziness of its own ([58]). Scientific judgment, according to [58], "is always essential in reaching a decision on the basis of incomplete or inconclusive evidence and therefore has always played a significant role in certifying".

Scientific judgment can also play a role early in the UQ process — usually as confidence

in the values of inputs or models as priors in Bayesian probability theory, subject to the opinion of the prescribing experts.

1.2.3 Subjective probability theory

The formulation of uncertainty quantification expressions in this work are fundamentally derived from the axioms of probability theory, whereby the notion of probability is the limit of a relative frequency of an event occurring and is thus an objective probability measure. The relative frequency interpretation is usually also fostered by those without sufficient understanding of the axioms of probability theory, as many would use the limit of the ratio of heads or tails landing to the number of total coin flips to express the probability of landing a head or tails. The second interpretation to probability — a degree of belief, or subjective probability — has gained popularity where expert opinion plays a definitive role in decision-making for engineering systems. Bayes theorem is the fundamental concept to subjectivistic theory, where updating the prior distribution $p(A)$ of an event A is based on new evidence (or expert opinion) E produces a posterior distribution $p(A|E)$:

$$p(A|E) = p(A)\frac{p(E|A)}{p(E)} \tag{1.4}$$

There is something to be said here for the topic of rare events and systems where few samples can be recorded and used for relative frequency calculations of interesting (but rare) events. Dealing with low probabilities-of-failure corresponding to very rare events may sometimes favor the subjective probability interpretation because the lack of statistical significance or experimental support rely more heavily on expert opinion to make decisions. For example, [4] points out that frequencies in geological sciences on the order of 10^{-11} per year correspond to events that are nearly impossible if the age of the earth is used

as a reference for occurrence time. This makes sense in their choosing of subjectivistic theory of probability as the "appropriate framework within which expert opinions, which are essential to the quantification process, can be combined with experimental results and statistical observations to produce quantitative measures of the risks from these systems".

Though the Bayesian framework presents advantages where expert opinion and rare events preclude the use of subjectivity, expert opinion may be considerably different between experts and the use of judgment may hinder the believability of the outcome. Thus, decision-making in subjectivity may have valid fundamentals, but the subjectivity propagated to posteriors may be considered invalid by other experts, whose subjectivity may adversely affect decision-making. This work seeks to provide a rigorous framework for uncertainty quantification, and the reasonable tract blooms from the objective definition of probability based on probability theory axioms.

1.2.4 Concentration-of-measure

This section presents a background of the concentration-of-measure phenomenon for the convenience of the reader. The interested reader is referred to [32] for a monograph and to [7] for a survey.

Recognition of the concentration-of-measure phenomenon as such may be traced back to an observation by Lévy [33] that functions on high dimensional spheres with small local oscillations, i.e., whose modulus of continuity can be controlled, are strongly concentrated around their mean value with respect to the uniform Lebesgue measure on the hypersphere. The study of this phenomenon was pioneered in the early seventies by V. Milman in his work on the asymptotic geometry of Banach spaces ([48], [47], [46], [49]). For an isoperimetric interpretation of the concentration-of-measure phenomenon and powerful applications to

geometry, refer to [20], [19], and [21]. For early probabilistic results in the context of sums of independent random variables, refer to [23] (Hoeffding's inequality), [12] (Chernoff bound), and to [59] and [15] for quantitative inequalities pertaining to the Glivenko-Cantelli ([11], [18]) convergence of empirical distributions. Far-reaching extensions, that in particular provide dimension-free concentration-of-measure inequalities in product spaces, have more recently been advanced by M. Talagrand (see [67], [65], [63], [64], [62], [61], [56]). For a selection of relevant articles in probability and statistics, refer to [9], [42], [41], [8], [31], [6], [30], and [29].

A brief compendium of representative concentration-of-measure inequalities is collected in the following section. Whereas only the simplest McDiarmid concentration-of-measure inequalities explained in §2 are employed in this work in §3, the more advanced inequalities described in §4 supply avenues for extension of the present QMU methodology to systems including correlated inputs, inputs with known probability distributions, unbounded inputs and other cases of interest.

Chapter 2

Concentration-of-Measure

2.1 Concentration-of-Measure Inequalities

The concentration-of-measure phenomenon, an observation of measure tightening in high dimensions, expresses probabilities of outlier events occurring by means of upper-bounded inequalities. These inequalities can be rearranged to resemble expressions dependent upon margins M and uncertainties U of a system. It is easiest to begin with the most fundamental concentration-of-measure inequality that will be used in most cases in this work:

2.1.1 McDiarmid's inequality

McDiarmid's inequality is perhaps the fundamental example of concentration-of-measure inequality. Let $X_1, \ldots, X_M$ be M random variables with values in spaces $E_1, \ldots, E_M$. Let F be a one-dimensional function of $X_1, \ldots, X_M$. Write

$$D_F^2 := \sum_{i=1}^{M} \sup_{(x_1,\ldots,x_{i-1},x_{i+1},\ldots x_M)\in E_1\times\cdots E_{i-1}\times E_{i+1}\times\cdots\times E_M} \sup_{(A_i,B_i)\in E_i^2} |F(x_1,\ldots,x_{i-1},A_i,x_{i+1},\ldots x_M) - F(x_1,\ldots,x_{i-1},B_i,x_{i+1},\ldots x_M)|^2. \tag{2.1}$$

Remark 2.1.1. *The computational cost (the speed and the storage requirements) of this optimization, as well as the quality of its solution, depend on the size of the search space.*

For relatively large dimensional search spaces, this optimization may become impractical due to several reasons:

1. *The computational cost associated with the optimizations prescribed in Eq. (2.1) may exceed the capabilities of contemporary computational resources for applications of interest. Furthermore, each evaluation of F may be overly computationally extensive and the uncertainty quantification for such systems may demand considerable computational resources deep into the peta- and exaflop regime.*

2. *A single optimization in Eq. (2.1) is not necessarily a convex optimization problem. Therefore, finding the global optimal is intractable in general. On the other hand, evolutionary optimization algorithms (such as genetic algorithms [50] and simulated annealing [27]) provide effective solution strategies for global optimization. Nevertheless, their success (convergence to the global optimum and the rate of convergence) degrades with increasing number of random variables — the curse of dimensionality.*

Typically D_F is referred to as the *diameter* of F. In Eq. (2.1) the suprema are taken with respect to variables in the spaces $E_1,\ldots, E_M$. (Here and subsequently throughout this work standard notation of probability theory applies—see e. g., Chapter 2 of [17] for an introduction.) Let $\mathbb{P}$ and $\mathbb{E}$ be the measure of probability and expectation associated with the random variables $X_1,\ldots,X_M$ and taking $X := (X_1,\ldots,X_M)$. The following theorem ([43]), also known as the *bounded-differences inequality*, bounds the fluctuations of $F(X)$ away from its mean without *a priori* knowledge of the probability distribution of the random variables $X_1,\ldots,X_M$.

Theorem 2.1.2 (McDiarmid, 1989)**.** *Let the random variables $X_1,\ldots,X_M$ be independent.*

Then

$$\mathbb{P}[F(X) - \mathbb{E}[F] \geq r] \leq \exp\left(-2\frac{r^2}{D_F^2}\right). \tag{2.2}$$

Observe that if the spaces $E_i = (a_i, b_i)$ and

$$F(X) = \frac{1}{M}\sum_{i=1}^{M} X_i\,, \tag{2.3}$$

then Hoeffding's inequality [23] is recovered as a special case

$$\mathbb{P}\left[\frac{1}{M}\sum_{i=1}^{M} X_i - \frac{1}{M}\sum_{i=1}^{M}\mathbb{E}[X_i] \geq r\right] \leq \exp\left(-2M\frac{r^2}{\left(\sum_{i=1}^{M}(b_i - a_i)^2/M\right)}\right). \tag{2.4}$$

2.1.2 Convex-distance inequality

The bounded-differences inequality is a special case of the more powerful *convex-distance inequality* [62]. In particular, the convex-distance inequality applies in cases where the bounded-differences inequality fails. Assume that $X_1, \ldots, X_M$ are independent variables, each taking values in a measurable set E. Let $X = (X_1, \ldots, X_M)$ and $\mathbb{P}[A] := \mathbb{P}[X \in A]$, and $A \subset E^M$ be an arbitrary measurable subset of E. For $\alpha \in [0,1]^M$, define the weighted Hamming distance from the point $x \in E^M$ to A as

$$d_\alpha(x, A) := \inf_{z \in A} \sum_{i:\, z_i \neq x_i} i^M |\alpha_i|. \tag{2.5}$$

Letting $\|\alpha\|^2 := \sum_{i=1}^{M} \alpha_i^2$, define the convex distance of X from the set A as

$$d_T(x, A) := \sup_{\alpha \in [0,1]^M, \|\alpha\|=1} d_\alpha(x, A) \tag{2.6}$$

then the following theorem holds.

Theorem 2.1.3. *For any subset $A \subset E^M$ with $\mathbb{P}[X \in A] \geq \frac{1}{2}$ and $t > 0$,*

$$\min\left\{\mathbb{P}[A], \mathbb{P}[d_T(X, A) \geq t]\right\} \leq \exp\left(-\frac{t^2}{4}\right). \tag{2.7}$$

The convex-distance inequality originates in a remarkable series of papers by Talagrand [65], [62], [66]. The preceding statement of the inequality is taken from [38].

2.1.3 Concentration inequalities with correlated random variables

The concentration-of-measure phenomenon is not limited to the case of independent random inputs and also arises when the inputs are correlated. Suppose, for definiteness, that $E_i = [0, 1]$ for all i. Let $f\big(X_1, \ldots, X_M\big)$ be the probability density of the inputs and denote by $f\big(X_j, \ldots, X_M | (X_1, \ldots, X_{i-1}, X_i) = (x_1, \ldots, x_{i-1}, x_i)\big)$ the law of $X_j, \ldots, X_M$ conditioned on $(X_1, \ldots, X_{i-1}, X_i) = (x_1, \ldots, x_{i-1}, x_i)$. A matrix Γ measuring the correlations between pairs of random variables X_i may then be defined as follows: $\Gamma_{ij} := 0$ if $i > j$; $\Gamma_{ii} := 1$; and, for $i < j$,

$$\begin{aligned}\Gamma_{ji} = \sup_{x_i, z_i \in [0,1]^2} \sup_{(x_1, \ldots, x_{i-1}) \in [0,1]^{i-1}} \\ \|f(X_j, \ldots, X_M | (X_1, \ldots, X_{i-1}, X_i) = (x_1, \ldots, x_{i-1}, x_i)) - \\ f(X_j, \ldots, X_M | (X_1, \ldots, X_{i-1}, X_i) = (x_1, \ldots, x_{i-1}, z_i))\|_{TV}\,,\end{aligned} \tag{2.8}$$

where $\|.\|_{TV}$ denotes the total variation norm over probability measures. For instance, if $(X_i)_{1 \leq i \leq M}$ is a Markov Chain on $[0, 1]$ with uniformly contracting transition kernels, i.e.,

$$\alpha := \sup_{i, x_{i-1}, z_{i-1}} \|f(X_i | X_{i-1} = x_i) - f(X_i | X_{i-1} = z_i)\|_{TV} < 1\,, \tag{2.9}$$

then

$$\|\Gamma\| \leq \frac{1}{1-\alpha^{\frac{1}{2}}}\,. \tag{2.10}$$

Write $\|\Gamma\|$ for the operator norm of the matrix Γ. Then the following theorem applies:

Theorem 2.1.4. *Let F be 1-Lipschitz over $[0,1]^M$. Then*

$$\mathbb{P}[|F - \mathbb{E}[F]| \geq r] \leq 2\exp\left(-\frac{r^2}{2\|\Gamma\|^2}\right). \tag{2.11}$$

Refer to [57] for a proof of the following theorem, and to [39], [40], [31], [24] for related results.

2.1.4 Concentration inequalities with functions with unbounded oscillations

Concentration-of-measure inequalities are not limited to functions with bounded differences or inputs taking their values on compact spaces. General concentration-of-measure inequalities can be obtained by controlling the Lipschitz regularity of the output function and the tail of the random variables X_i at infinity. This control can be achieved by means of analytical inequalities called logarithmic Sobolev inequalities [22]. Let f be a nonnegative measurable function f over a measure space $(E, \mathcal{B}, \mu)$ such that $\int f \ln(1+f) d\mu < \infty$. Define the entropy of f as

$$\mathrm{Ent}_\mu(f) := \int f \ln f d\mu - \int f d\mu \ln\left(\int f d\mu\right). \tag{2.12}$$

Refer to Section 5 of [32] for the following theorem and to references therein.

Theorem 2.1.5. *Let μ be a probability measure on the Borel sets of a metric space (E, d) such that for some $C > 0$ and all locally Lipschitz function f on E*

$$\mathrm{Ent}_\mu(f^2) \leq 2C \int |\nabla f|^2 d\mu \tag{2.13}$$

with

$$|\nabla f|(x) := \limsup_{y \to x} \frac{|f(x) - f(y)|}{|x - y|} \,. \tag{2.14}$$

Then for every 1-Lipschitz integrable function $F : E \to R$ and for every $r \geq 0$

$$\mu\left(\{Y \geq \int F d\mu + r\}\right) \leq \exp\left(-\frac{r^2}{2C}\right) \,. \tag{2.15}$$

Eq. (2.13) is deemed the logarithmic Sobolev inequality with constant C. The application of Theorem 2.1.5 to product of metric spaces (E_i, d_i) follows from the observation that if the measures μ_i satisfy the logarithmic Sobolev inequality

$$\mathrm{Ent}_{\mu_i}(f^2) \leq 2C_i \int |\nabla_i f|^2 d\mu_i \tag{2.16}$$

for every locally Lipschitz function f on E_i, where $|\nabla_i f|$ is the generalized modulus of gradient on E_i defined as in Eq. (2.14), then the product measure $\mu = \mu_1 \times \cdots \times \mu_M$ satisfies the logarithmic inequality (see Corollary 5.7 of [32])

$$\mathrm{Ent}_\mu(f^2) \leq 2 \max_{1 \leq i \leq M} C_i \int |\nabla f|^2 d\mu \tag{2.17}$$

with $|\nabla f|^2 = \sum_{i=1}^M |\nabla_i f|^2$, for every locally Lipschitz function f on $E = E_1 \times \cdots \times E_M$. The particular application to Gaussian distributions follows simply from the observation that

a normal centered Gaussian distribution on $\mathbb{R}^n$ satisfies the logarithmic Sobolev inequality with constant 2.

2.1.5 Concentration inequalities for empirical processes

Concentration inequalities can also be used to obtain powerful and very quantitative estimates for empirical processes defined by sampling, ([65], [62], [66], [41]). Let $Y^1, \ldots, Y^N$ be independent random variables, not necessarily identically distributed, in some measurable space $(E, \mathcal{B})$. Let $\mathcal{F}$ be some countable family of real-valued measurable functions on $(E, \mathcal{B})$ such that $\|f\|_\infty \leq b < \infty$ for every $f \in \mathcal{F}$. Let

$$Z := \sup_{f \in \mathcal{F}} \big| \sum_{i=1}^{N} f(Y_i) \big| \tag{2.18}$$

or

$$Z := \sup_{f \in \mathcal{F}} \big| \sum_{i=1}^{N} (f(Y_i) - \mathbb{E}[f(Y_i)]] \big|. \tag{2.19}$$

In addition, let

$$\sigma^2 := \sup_{f \in \mathcal{F}} \sum_{i=1}^{N} \mathrm{Var}[f(Y_i)]. \tag{2.20}$$

Then the following holds [41].

Theorem 2.1.6. *For any positive real number ϵ and x,*

$$\mathbb{P}\left[Z \geq (1+\epsilon)\mathbb{E}[Z] + \sigma\sqrt{2\kappa x} + \kappa(\epsilon) b x \right] \leq \exp(-x) \tag{2.21}$$

with $\kappa = 4$, $\kappa(\epsilon) = 2.5 + 32\epsilon^{-1}$. *Moreover,*

$$\mathbb{P}\left[Z \leq (1-\epsilon)\mathbb{E}[Z] - \sigma\sqrt{2\kappa' x} - \kappa'(\epsilon)x\right] \leq \exp(-x) \tag{2.22}$$

with $\kappa' = 5.4$ *and* $\kappa'(\epsilon) = 2.5 + 43.2\epsilon^{-1}$.

In the particular case in which the random variables Y_i are identically distributed, theorem 2.1.6 furnishes powerful estimates for the empirical process

$$\mu_N := \frac{1}{N}\sum_{i=1}^{N} \delta_{Y_i} \tag{2.23}$$

by observing that

$$Z = \sup_{f \in \mathcal{F}} N|\mu_N(f) - \mu(f)|. \tag{2.24}$$

2.2 Concentration-of-Measure Inequalities Cases

The setting of a typical UQ problem is as follows: consider a system with possibly multiple output performance measures $Y : \Omega \to E_1 \times \cdots \times E_N$ on a given probability space $(\Omega, \mathcal{U}, \mathbb{P})$. In particular, $E_1, \ldots, E_N$ denote Euclidean spaces endowed with the standard metric and $\mathbb{P}$ is a probability measure. Suppose that the safe operation of the system requires that $Y \in A$ for some open *admissible set* $A \subset E_1 \times \cdots \times E_N$. Ideally, the support of the probability measure associated to Y would be contained within A, i.e.,

$$\mathbb{P}[Y \in A] = 1. \tag{2.25}$$

Systems satisfying this condition can be certified with complete certainty. However, this absolute guarantee of safe performance may be unattainable, e.g., if $\mathbb{P}$ lacks compact

Table 2.1: Number of tests required for certification as a function of the probability-of-failure tolerance

failure tolerance (ϵ)	number of tests (m)
1	0
10^{-1}	115
10^{-2}	23,025
10^{-3}	3,453,877
10^{-4}	460,517,018
10^{-5}	57,564,627,324
10^{-6}	6,907,755,278,982

support, $Y \notin A$ for a measurable set of support of the probability measure associated to Y, or evaluations prohibitively expensive in practice to obtain. In these cases, the condition in Eq. (2.25) may be relaxed to

$$\mathbb{P}[Y \in A^c] \leq \epsilon \tag{2.26}$$

for some appropriate *certification tolerance* ϵ, where $A^c = E_1 \times \cdots \times E_N \backslash A$ is the *inadmissible set*. Inequality 2.26 expresses the requirement that the probability of system failure be acceptably small and gives rigorous mathematical expression to the QMU conceptual view of certification.

A *conservative* certification criterion is obtained when the probability-of-failure $\mathbb{P}[Y \in A^c]$ is bounded from above and the upper bound is verified to be below the certification tolerance ϵ. Evidently, for an upper bound to be useful it must be *tight*, i.e., it must be close to the actual probability-of-failure $\mathbb{P}[Y \in A^c]$. Therefore, the essential mathematical and computational challenge is to obtain tight upper bounds to the probability-of-failure of the system.

A first strategy that naturally suggests itself is to bound the probability-of-failure empirically, i.e., solely by means of experimental testing. Suppose that m tests are performed with outcomes $Y^1, \ldots, Y^m$. With this data set the empirical probability measure is extracted

as

$$\mu_m = \frac{1}{m}\sum_{i=1}^{m}\delta_{Y^i}. \tag{2.27}$$

Then, an application of Hoeffding's inequality [23] gives, with probability $1-\epsilon'$,

$$\mathbb{P}[Y \in A^c] \leq \mu_m[A^c] + \sqrt{\frac{1}{2m}\log\frac{1}{\epsilon'}}. \tag{2.28}$$

Hence, the inequality

$$\mu_m[A^c] + \sqrt{\frac{1}{2m}\log\frac{1}{\epsilon'}} \leq \epsilon \tag{2.29}$$

supplies a conservative certification condition. Note that Inequality 2.28 can be improved by using Chernoff's inequality instead of Hoeffding's inequality when the probability-of-failure $\mathbb{P}[Y \in A^c]$ is known. Since that probability-of-failure is unknown, it can't be used in practice to give a bound on the necessary number of tests one has to use Hoeffding's inequality.

The certification criterion of Eq. (2.29) reveals that the number of experiments required to certify a system based on statistical sampling alone is of the order of $\frac{1}{2}\epsilon^{-2}\log\frac{1}{\epsilon'}$. The number-of-tests requirement is shown in Table 2.1 as a function of the probability-of-failure tolerance ϵ with $\epsilon' = \epsilon$. It is clear from this table that the number of tests required for the purely empirical certification of a system can be prohibitively expensive if the tests are costly and the required probability-of-failure is low.

It is natural to require that the confidence level on the estimation of the probability-of-failure is at least of the order of the required maximum probability-of-failure. That is why $\epsilon' = \epsilon$ is a natural chose. For $\epsilon' > \epsilon$ the uncertainty on the estimation of the probability-of-failure would be larger than the probability-of-failure itself, which translates into the fact

that the probability-of-failure may be, with probability ϵ', above the certification threshold ϵ.

When empirical certification is not an option, the question that naturally arises is how models can be employed to reduce the number of tests required for certification. Thus, the goal of *model-based certification* is to achieve certification with minimal testing. To frame this question in mathematical terms, suppose that the behavior of the system is exactly described by an unknown function $Y = G(X, Z)$ of random input variables (X, Z). For definiteness, assume that X and Z are independent and range over known intervals, and that no additional statistical information about the input parameters is available. These assumptions can be readily relaxed, but such extensions will not be pursued here in the interest of simplicity. Suppose in addition that the behavior of the system is modeled by means of a function $Y = F(X)$. Thus X collects the input variables that are accounted for by the model, whereas Z collects those variables that are unaccounted for, or *unknown unknowns*. Evidently, if an exact model were available no testing would be required to achieve certification. In general, models cannot be expected to be exact and the degree of predictiveness of the models needs to be carefully assessed with V&V.

The manner in which concentration-of-measure supplies probability-of-failure upper bounds with the aid of validated models is summarized in this section through a sequence of representative scenarios of increasing complexity.

2.2.1 Scenario 1: Exact model, single performance measure whose mean is known

Begin by assuming that the mean performance $\mathbb{E}[Y]$ is known exactly and the model is perfect, i.e., there exists a random vector $X : \Omega \to \chi_1 \times \cdots \times \chi_M$ and a known function

$F : \chi_1 \times \cdots \times \chi_M \to E_1 \times \cdots \times E_N$ such that the relation $Y = F(X)$ describes the system exactly. These assumptions represent ideal conditions in which all uncertainty regarding the response of the system is *aleatoric uncertainty*, i.e., stems from the stochastic variability of the system, and there is no *epistemic uncertainty*, i.e., the behavior of the system is known exactly, including its mean response. Begin by considering the case in which certification depends on a single performance measure, i.e., $N = 1$. Under these assumptions, the admissible set is of the form $A = [a, \infty)$, where a is the minimum threshold for safe system operation. Then, if F is integrable (and that is the only assumption on F) and the input parameters are independent McDiarmid's inequality states that

$$\mathbb{P}\left[F(X) - \mathbb{E}[F(X)] \leq -r\right] \leq \exp\left(-2\frac{r^2}{D_F^2}\right) \tag{2.30}$$

where

$$\begin{aligned} D_F^2 := \sum_{k=1}^{M} & \sup_{(x_1,\ldots,x_{k-1},x_{k+1},\ldots x_M)\in\chi_1\times\ldots\chi_{k-1}\times\chi_{k+1}\times\cdots\times\chi_M} \\ & \sup_{(A_k,B_k)\in\chi_k^2} |F(x_1,\ldots,A_k,\ldots x_M) - F(x_1,\ldots,B_k,\ldots x_M)|^2 \end{aligned} \tag{2.31}$$

is the *verification diameter* of the system. Bound 2.30 can be re-written in the form

$$\mathbb{P}[Y \in A^c] \leq \exp\left(-2\frac{(\mathbb{E}[Y] - a)_+^2}{D_F^2}\right) \tag{2.32}$$

where $x_+ := \max(0, x)$, whence it follows that the inequality

$$\frac{(\mathbb{E}[Y] - a)_+}{D_F} \geq \sqrt{\log\sqrt{\frac{1}{\epsilon}}} \tag{2.33}$$

Table 2.2: Minimum confidence factor CF required to stay within a pre-specified probability-of-failure tolerance ϵ

failure tolerance (ϵ)	1	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
confidence factor (CF)	0	1.07298	1.51743	1.85846	2.14597	2.39926	2.62826

supplies a conservative certification criterion.

Comparison of Eq. (2.33) and Eq. (1.1) affords the identification:

$$M = (\mathbb{E}[Y] - a)_+ \tag{2.34a}$$

$$U = D_F. \tag{2.34b}$$

Thus, in the absence of epistemic uncertainty, i.e., for systems for which an exact model is available and whose mean performance is exactly known, *the margin M is the difference between the mean performance and the threshold*, or zero if this difference is negative, and *the uncertainty U equals the verification diameter of the system.* With these identifications, the certification criterion can be expressed in the form

$$\mathrm{CF} = \frac{M}{U} \geq \sqrt{\log \sqrt{\frac{1}{\epsilon}}}. \tag{2.35}$$

This inequality establishes a clear correspondence between the probability-of-failure tolerance ϵ and the confidence factor CF. This correspondence is shown in tabular form in Table 2.2. Thus, concentration-of-measure inequalities supply precise definitions of margin measures, uncertainty measures, and minimum confidence factors that guarantee the safe operation of the system to within a pre-specified probability-of-failure tolerance.

Several additional aspects of the certification method described above are noteworthy. Firstly, the only information about the model that is required for certification is the mean

performance and the diameter of the system. In particular, the response function $F(X)$ need not be interpolated or otherwise represented, and can effectively be treated as a black box. Secondly, only ranges of input parameters, and not their detailed probability distribution functions, need be known for certification. Thirdly, the present scenario provides an extreme example of how the availability of a high-fidelity model helps to reduce the number of tests required for certification: when an exact model is available, the need for experimental testing is eliminated altogether.

It is also interesting to note that the verification diameter of Eq. (2.31), which provides a rigorous measure of the aleatoric uncertainty in the response of the system, represents the largest deviation in system performance that is recorded when each input parameter is allowed to vary in turn between pairs of values spanning its entire range. Evidently, the computation of the verification diameter of a system entails an optimization over input-parameter space seeking to identify those large deviations in the input parameters that result in the largest deviations in the output parameters. It bears emphasis that consideration of large deviations is essential and that, in particular, linearized sensitivity analysis is not sufficient for rigorous certification in general. Specific optimization algorithms for the computation of verification diameters are discussed in §3.2.

It should be noted that this work uses the term *verification* in a somewhat expanded sense relative to other conventional uses of the term (see, e.g., [51], [53]), discussed in §1.2.2.1. Thus, a common use of the term verification is to signify the process of assessing how well a numerical model approximates the underlying physical laws governing the system, often expressed as a system of partial differential equations. However, in the present context the term verification naturally refers to the process of quantifying all aleatoric uncertainties, whether arising from numerical errors, from the statistical variability of the

input parameters, from the intrinsic stochasticity of the model, or from other sources. Thus, verification aims to quantify how precisely a model can predict the response of the system. A distinguishing characteristic of verification is that it is achieved solely by exercising the model and without reference to experimental data. The concentration-of-measure framework renders verification, often a conceptually appealing but ambiguous and imprecise term, in precise quantitative terms. Thus, as already noted, verification is rigorously quantified by the verification diameter, in the sense that once the verification diameter of the system is known, the system can be certified rigorously and conservatively.

For completeness and in order to illustrate the assumptions required for Eq. (2.30) to hold, a simple version of McDiarmid's inequality [43] (which is also known as a bounded differences inequality, and where the constant 2 in the exponential has been replaced by $\frac{1}{2}$) is included with a sketch of its proof. The sharpening of Inequality 2.37 to Eq. (2.30) is technically involved — refer to [43] for a complete account. Independence is not a necessary condition for McDiarmid's inequality — refer to [44] for an extension of that inequality to centering sequences (for instance to martingales).

Theorem 2.2.1. *Assume that the random variables $X_1, \ldots, X_M$ are independent. Then*

$$\mathbb{P}[F(X) - \mathbb{E}[F] \geq r] \leq \exp\left(-\frac{1}{2}\frac{r^2}{D_F^2}\right) \tag{2.36}$$

and

$$\mathbb{P}[F(X) - \mathbb{E}[F] \leq -r] \leq \exp\left(-\frac{1}{2}\frac{r^2}{D_F^2}\right). \tag{2.37}$$

Proof. This is an adaptation of the proof of Corollary 1.17 in [32]. By Chebyshev's inequality then,

$$\mathbb{P}[F(X) - \mathbb{E}[F] \geq r] \leq e^{-\lambda r}\mathbb{E}\Big[\exp\big(\lambda(F - \mathbb{E}[F])\big)\Big] \tag{2.38}$$

for all $\lambda \geq 0$. In addition,

$$\mathbb{E}\Big[\exp\big(\lambda(F-\mathbb{E}[F])\big)\Big] = \mathbb{E}\Big[\mathbb{E}\big[\exp(\lambda(F-\mathbb{E}[F|\mathcal{F}_{n-1}]))|\mathcal{F}_{n-1}\big]\exp(\lambda(\mathbb{E}[F|\mathcal{F}_{n-1}]-\mathbb{E}[F]))\Big] \tag{2.39}$$

where $\mathcal{F}_k$ denotes the σ-algebra generated by $X_1,\ldots,X_k$. Recall that, for every bounded function f of zero mean with respect to a measure ν and all $\lambda \geq 0$, Jensen's inequality gives

$$\int e^{\lambda f} d\nu \leq \int\int e^{\lambda(f(x)-f(y))} d\nu(x) d\nu(y) \leq \sum_{i=0}^{\infty} \frac{(D_f\lambda)^{2i}}{(2i)} \leq \exp(D_f^2\lambda^2/2) \tag{2.40}$$

with $D_f := \sup_{x,y}|f(x)-f(y)|$. Applying Inequality 2.40 to the integration with respect to the law of X_n, the following is obtained

$$E\big[\exp(\lambda(F-\mathbb{E}[F|\mathcal{F}_{n-1}]))|\mathcal{F}_{n-1}\big] \leq \exp(D_n^2\lambda^2/2) \tag{2.41}$$

where

$$D_n := \sup_{x_1,\ldots,x_{n-1},x,y} |F(x_1,\ldots,x_{n-1},x) - F(x_1,\ldots,x_{n-1},y)|. \tag{2.42}$$

It follows that

$$\mathbb{E}\Big[\exp\big(\lambda(F-\mathbb{E}[F])\big)\Big] = \exp(D_n^2\lambda^2/2)\mathbb{E}\Big[\exp(\lambda(\mathbb{E}[F|\mathcal{F}_{n-1}]-\mathbb{E}[F]))\Big] \tag{2.43}$$

and, by induction,

$$\mathbb{E}\Big[\exp\big(\lambda(F-\mathbb{E}[F])\big)\Big] = \exp(D_F^2\lambda^2/2). \tag{2.44}$$

Combining this inequality with Eq. (2.38) and taking $\lambda = r/D_F^2$ finally gives

$$\mathbb{P}[F(X) - \mathbb{E}[F] \geq r] \leq \exp\left(-\frac{1}{2}\frac{r^2}{D_F^2}\right). \tag{2.45}$$

Inequality 2.37 is obtained by replacing F by $-F$. □

Additionally observe that Hoeffding's inequality [23]

$$\mathbb{P}\left[\frac{1}{M}\sum_{i=1}^{M} X_i - \frac{1}{M}\sum_{i=1}^{M}\mathbb{E}[X_i] \geq r\right] \leq \exp\left(-2M\frac{r^2}{\left(\sum_{i=1}^{M}(b_i - a_i)^2/M\right)}\right) \tag{2.46}$$

follows from McDiarmid's inequality as a special case when the spaces E_i are equal to intervals (a_i, b_i) and

$$F(X) = \frac{1}{M}\sum_{i=1}^{M} X_i. \tag{2.47}$$

2.2.2 Scenario 2: Exact model, multiple performance measures whose mean is known

The preceding framework can be extended to the case in which certification depends on multiple performance measures. To this end, suppose that A is an arbitrary subset of $\mathbb{R}^N$, i.e., $N \geq 1$ and $E_i = \mathbb{R}$, $i = 1, \ldots, N$, and that the mean performance $\mathbb{E}[Y]$ is known and belongs to the interior of A. Then, the following concentration-of-measure inequality is deduced from McDiarmid's inequality

$$\mathbb{P}[Y \in A^c] \leq \inf_{s,r \in \mathbb{R}^N \,:\, \mathbb{E}[Y] + \prod_{i=1}^{N}(-s_i, r_i) \subset A} \sum_{i=1}^{N}\left[\exp\left(-2\frac{r_i^2}{D_{F_i}^2}\right) + \exp\left(-2\frac{s_i^2}{D_{F_i}^2}\right)\right] \tag{2.48}$$

where D_{F_i} is the verification diameter of the response function F_i and the infimum is taken over all hyperrectangles contained in A. This inequality follows simply from Mc-

Diarmid's inequality by observing that, for all $s, r \in \mathbb{R}^N$ ($s := (s_1, \ldots, s_N)$) such that $\mathbb{E}[Y] + \prod_{i=1}^N (-s_i, r_i) \subset A$, one has

$$A^c \subset \cup_{i=1}^N \{Y_i - \mathbb{E}[Y_i] \geq r_i\} \cup \{Y_i - \mathbb{E}[Y_i] \leq -s_i\}, \tag{2.49}$$

whence it follows that

$$\mathbb{P}[Y \in A^c] \leq \sum_{i=1}^N \big(\mathbb{P}[Y_i - \mathbb{E}[Y_i] \geq r_i] + \mathbb{P}[Y_i - \mathbb{E}[Y_i] \leq -s_i]\big). \tag{2.50}$$

As in the preceding scenario, the inequality

$$\inf_{s,r \in \mathbb{R}^N \,:\, \mathbb{E}[Y] + \prod_{i=1}^N (-s_i, r_i) \subset A} \sum_{i=1}^N \left[\exp\left(-2 \frac{r_i^2}{D_{F_i}^2} \right) + \exp\left(-2 \frac{s_i^2}{D_{F_i}^2} \right) \right] \leq \epsilon \tag{2.51}$$

now supplies a conservative certification criterion. It follows that in the case of multiple performance measures, certification can be achieved by the computation of the verification diameters of each of the components of the response function.

Suppose, for example, that $A = \prod_{i=1}^N (a_i, +\infty)$, where a_i is the threshold of the ith performance measure, and suppose that $a_i \leq \mathbb{E}[Y_i]$, $i = 1, \ldots, N$. Then, the certification inequality 2.51 reduces to

$$\sum_{i=1}^N \exp\left(-2 \frac{(\mathbb{E}[Y_i] - a_i)_+^2}{D_{F_i}^2} \right) \leq \epsilon. \tag{2.52}$$

It should also be noted that Eq. (2.52) is based on the worst-case scenario that these events do not intersect and therefore the sum of their respective probabilities is an upper bound for the union. In practical applications, there may be an overlap between the indi-

Table 2.3: Aggregate confidence factor for a system characterized by two performance measures with confidence factors CF_1 (rows) and CF_2 (columns)

CF_i	1.0	1.2	1.4	1.6	1.8	2.0
1.0	0.808348	0.909127	0.965192	0.989139	0.997179	0.999381
1.2	0.909127	1.045670	1.135196	1.178736	1.194371	1.198758
1.4	0.965192	1.135196	1.270207	1.352168	1.386639	1.397003
1.6	0.989139	1.178736	1.352168	1.487759	1.563896	1.591443
1.8	0.997179	1.194371	1.386639	1.563896	1.701007	1.772316
2.0	0.999381	1.198758	1.397003	1.591443	1.772316	1.911394

vidual failure events that constitutes this bound. This point is not addressed here, but will be considered in a sequel work.

By analogy to the case of a single performance measure, margins and uncertainty measures can be introduced as

$$M_i = (\mathbb{E}[Y_i] - a_i)_+ \tag{2.53a}$$

$$U_i = D_{F_i} \tag{2.53b}$$

for each performance measure in turn. Then, Eq. (2.52) can be rewritten in the form

$$\mathrm{CF} = \sqrt{\log\left(\frac{1}{\sqrt{\sum_{i=1}^{N}\exp\left(-2(\mathrm{CF}_i)^2\right)}}\right)} \geq \sqrt{\log\sqrt{\frac{1}{\epsilon}}} \tag{2.54}$$

where

$$\mathrm{CF}_i = \frac{M_i}{U_i}. \tag{2.55}$$

Evidently, the certification criterion of Eq. (2.54) reduces to Eq. (2.35) in the case of a single performance measure. It is interesting to note that, in the case of multiple performance measures, the confidence factor CF follows as an aggregate of the confidence factors CF_i of each of the performance measures according to the composition rule of Eq. (2.54).

However, it should be carefully noted that neither margins nor uncertainties can be aggregated independently of each other to define an overall margin and an overall uncertainty of the system. The confidence factor aggregation relation is shown in Table 2.3 for a system characterized by two performance measures. As expected, the aggregate confidence factor is smaller than each of the confidence factors corresponding to the individual performance measures. Thus, lack of confidence in individual performance measures compounds and the overall confidence in the system decreases with the addition of every new performance measure. However, because the individual confidence factors enter the aggregation relation of Eq. (2.54) through exponentials, it follows that the aggregate confidence factor—and the certification process itself—is dominated by those performance measures having the smallest individual confidence factors. Conversely, performance measures having large confidence factors have negligible effect on the overall confidence factor of the system and can be safely removed from consideration in the certification process.

2.2.3 Scenario 3: Exact model, single performance measure whose mean is unknown

In the foregoing, the mean performance $\mathbb{E}[Y]$ of the system is assumed known *a priori*. However, in most situations of practical interest such information is not available and the mean performance must be estimated instead. Suppose that, to this end, one performs m evaluations of the model $F(X)$ based on an unbiased sampling of the input parameters, resulting in predicted performances $Y^1, Y^2, \ldots, Y^m$. Define the *estimated mean performance* corresponding to these calculations as

$$\langle Y \rangle = \frac{1}{m} \sum_{i=1}^{m} Y^i. \tag{2.56}$$

Start by additionally assuming that there is one single performance measure, $N = 1$, and that the safe operation of the system requires that $Y \geq a$ for some threshold a, i.e., $A = [a, \infty)$. The probability $\mathbb{P}[Y \in A^c]$ can now only be determined to within confidence intervals reflecting the randomness of the estimated mean $\langle Y \rangle$. Under these conditions, the following inequality is obtained

$$\mathbb{P}\left[\mathbb{P}[Y \in A^c] \geq \exp\left(-2\frac{[\langle Y \rangle - a - \alpha]_+^2}{D_F^2}\right)\right] \leq \epsilon' \tag{2.57}$$

where ϵ' is a pre-specified *estimation tolerance* and

$$\alpha = \sqrt{\frac{-\log \epsilon'}{D_F m}}. \tag{2.58}$$

Inequality 2.57 follows simply from an application of McDiarmid's inequality to $\langle Y \rangle$, with the result

$$\mathbb{P}[\mathbb{E}[Y] - \langle Y \rangle \leq -\alpha] \leq \epsilon' \tag{2.59}$$

whence it follows that, with $\mathbb{P}$ probability $1 - \epsilon'$,

$$A^c \subset \{Y - \mathbb{E}[Y] \leq a + \alpha - \langle Y \rangle\}. \tag{2.60}$$

Next observe that Eq. (2.57) simply states that, with probability $1 - \epsilon'$,

$$\mathbb{P}[Y \in A^c] \leq \exp\left(-2\frac{[\langle Y \rangle - a - \alpha]_+^2}{D_F^2}\right) \tag{2.61}$$

and the certification criterion of Eq. (2.26) now becomes

$$\frac{[\langle Y\rangle - a - \alpha]_+}{D_F} \geq \sqrt{\log\sqrt{\frac{1}{\epsilon}}}\,. \tag{2.62}$$

This certification criterion is again of the form of Eq. (2.35) with margin and uncertainty measure

$$M = [\langle Y\rangle - a - \alpha]_+ \tag{2.63a}$$

$$U = D_F. \tag{2.63b}$$

Comparison of Eq. (2.34) and Eq. (2.63) shows that the estimation of the mean performance of the system effectively reduces the margin by the amount α. Evidently, this margin decrease can be reduced to an arbitrarily small value by carrying out a sufficiently large number of model evaluations. The certification criterion of Eq. (2.63) again shows that, as in the case of known mean performance and in the absence of epistemic uncertainty, certification can be rigorously achieved from the sole knowledge of the verification diameter of the system and an estimate of its mean performance.

2.2.4 Scenario 4: Exact model, multiple performance measures whose means are unknown

For completeness, proceed to record the extension of the preceding case to multiple performance measures, $N \geq 1$, and arbitrary A in $\mathbb{R}^N$. In this case, with probability $1-\epsilon'$,

$$\mathbb{P}[Y \in A^c] \leq \inf_{s,r\in\mathbb{R}^N\,:\,\mathbb{E}[Y]+\prod_{i=1}^N(-s_i-\alpha_i,r_i+\alpha_i)\subset A} \sum_{i=1}^N \left(\exp\left(-2\frac{r_i^2}{D_{F_i}^2}\right) + \exp\left(-2\frac{s_i^2}{D_{F_i}^2}\right)\right) \tag{2.64}$$

where $\alpha \in \mathbb{R}^N$ and

$$\alpha_i = D_{F_i}\sqrt{\frac{\log\frac{2N}{\epsilon'}}{2m}} \tag{2.65}$$

Inequality 2.64 again follows from McDiarmid's inequality by observing that

$$\mathbb{P}\left[|\mathbb{E}[Y^i] - \langle Y\rangle^i| \geq \alpha_i\right] \leq \frac{\epsilon'}{N}. \tag{2.66}$$

Hence, $|\mathbb{E}[Y^i] - \langle Y\rangle^i| \leq \alpha_i$ for all i with probability $1-\epsilon'$, with the result that for all $s, r \in \mathbb{R}^N$ such that $\langle Y\rangle + \prod_{i=1}^N(-s_i - \alpha_i, r_i + \alpha_i) \subset A$,

$$\{Y \in A^c\} \subset \cup_{i=1}^N \{Y_i - \mathbb{E}[Y_i] \geq s_i\} \cup \{\mathbb{E}[Y_i] - Y_i \geq r_i\} \tag{2.67}$$

with probability $1-\epsilon'$, and the certification criterion finally becomes

$$\inf_{s,r\in\mathbb{R}^N\,:\,\langle Y\rangle+\prod_{i=1}^N(-s_i-\alpha_i,r_i+\alpha_i)\subset A} \sum_{i=1}^N \left(\exp\left(-2\frac{r_i^2}{D_{F_i}^2}\right) + \exp\left(-2\frac{s_i^2}{D_{F_i}^2}\right)\right) \leq \epsilon. \tag{2.68}$$

For example, suppose, as in Scenario 2, that $A = \prod_{i=1}^N (a_i, +\infty)$, where a_i is the threshold of the ith performance measure, and suppose that $a_i + \alpha_i \leq \langle Y_i\rangle$, $i = 1, \ldots, N$. Then, the certification inequality of Eq. (2.68) reduces to

$$\sum_{i=1}^N \exp\left(-2\frac{(\langle Y_i\rangle - a_i - \alpha_i)_+^2}{D_{F_i}^2}\right) \leq \epsilon \tag{2.69}$$

where now

$$\alpha_i = D_{F_i}\sqrt{\frac{\log\frac{N}{\epsilon'}}{2m}} \tag{2.70}$$

in view of the one-side form of the admissible intervals. By analogy to the case of a single performance measure, introduce the margins and uncertainty measures

$$M_i = (\langle Y_i \rangle - a_i - \alpha_i)_+ \tag{2.71a}$$

$$U_i = D_{F_i} \tag{2.71b}$$

for each performance measure in turn. Then, Eq. (2.69) can be rewritten in the form of Eq. (2.54) with the individual performance measure confidence factor defined as in Eq. (2.55). As in the case of a single performance measure, observe that the need to estimate the mean performance has the effect of reducing the individual performance measure margins in the amounts α_i. These are controllable margin decreases that can be reduced to any desired extent by carrying out a sufficiently large number of model evaluations. Again, in the absence of epistemic uncertainty, certification can be rigorously achieved from the sole knowledge of the verification diameters of the individual performance measures and estimates of their mean performance.

2.2.5 Scenario 5: Inexact model

In practice F denotes a numerical or analytical model of a physical system whose output is the random vector Y. The model accounts for some of the input parameters X that determine the performance of the system. If the model were perfect then, for all X, $F(X)$ would exactly equal the outcome Y of an experiment performed with an identical set of input parameters. In general, Y and $F(X)$ are not equal owing to:

i) imperfections in the model,

ii) the existence of additional unknown random input parameters, or *unknown unknowns*,

not accounted for in the model.

These two sources of error and epistemic uncertainty can be analyzed by supposing that the exact response of the physical system is governed by a function $G(X, Z)$, generally unknown, of the random variables X that are accounted for by the model and additional unknown unknowns Z. Even if no unknown unknowns exist and the model accounts for all input parameters of the system, in general $G(X, Z) \neq F(X)$ owing to the limited fidelity of the model $F(X)$. Evidently, these sources of error and epistemic uncertainty, namely, the existence of unknown unknowns and the limited fidelity of the model, must be carefully assessed as part of the certification of the system.

To this end, begin by considering the case of a single performance measure, $N = 1$, and by noting that

$$\{Y \leq a\} \subset \{F(X) \leq a + h\} \cup \{G(X, Z) - F(X) \leq -h\}, \tag{2.72}$$

where $G - F$ may be regarded as a *modeling-error function* and h is an arbitrary number, leading to the estimate

$$\mathbb{P}[Y \leq a] \leq \mathbb{P}[F(X) \leq a + h] + \mathbb{P}[G(X, Z) - F(X) \leq -h] \tag{2.73}$$

and to the conservative certification criterion

$$\mathbb{P}[F(X) \leq a + h] + \mathbb{P}[G(X, Z) - F(X) \leq -h] \leq \epsilon. \tag{2.74}$$

Ideally h should be chosen to minimize the sum of the probabilities on the left-hand side of Eq. (2.74).

It has been observed in previous sections how to obtain a bound for $\mathbb{P}[F(X) \leq a + h]$, i.e., for the probability-of-failure predicted by the model (with an additional margin h). Certification now additionally requires a bound on $\mathbb{P}[G(X, Z) - F(X) \leq -h]$, i.e., the probability that the predicted and measured performance differ significantly. This probability measures the deleterious effect on predictiveness of all model imperfections, whether resulting from the limited fidelity of the model or from unknown unknowns, and may therefore be regarded as a measure of *epistemic uncertainty*. As is evident from Eq. (2.74), the epistemic uncertainty has the effect of decreasing the effective probability-of-failure tolerance. In particular, certification is not possible if the model is not sufficiently faithful, i.e., if $\mathbb{P}[G(X, Z) - F(X) \leq -h] \geq \epsilon$.

The epistemic uncertainty can again be rigorously bounded by means of concentration-of-measure inequalities. Thus, a direct application of McDiarmid's inequality to $G - F$ gives the concentration-of-measure bound

$$\mathbb{P}[G(X, Z) - F(X) \leq -h] \leq \exp\left(-2\frac{(\mathbb{E}[G - F] + h)_+^2}{D_{G-F}^2}\right) \tag{2.75}$$

where the *validation diameter*

$$\begin{aligned} D_{G-F}^2 := \sum_{k=1}^{M} \sup_{(x_1,\ldots,x_{k-1},x_{k+1},\ldots x_M)\in E_1\times\ldots E_{k-1}\times E_{k+1}\times\cdots\times E_M} \sup_{z\in E_{M+1}} \sup_{(A_k,B_k)\in E_k^2} \\ |F(x_1,\ldots,x_{k-1},A_k,x_{k+1},\ldots x_M) - G(x_1,\ldots,x_{k-1},A_k,x_{k+1},\ldots x_M,z) \\ - F(x_1,\ldots,x_{k-1},B_k,x_{k+1},\ldots x_M) + G(x_1,\ldots,x_{k-1},B_k,x_{k+1},\ldots x_M,z')|^2 \end{aligned} \tag{2.76}$$

supplies a measure of the *epistemic uncertainty* of the system. In particular, the validation diameter measures the extent to which the predictions of the model deviate from observation.

In practice $\mathbb{E}[G-F]$ and $\mathbb{E}[F]$ are not known and must instead be estimated. Let $\langle G-F\rangle$ and $\langle F\rangle$ be the empirical means of $G-F$ and F, respectively, estimated from m nominally identical model evaluations and experiments. Take $\alpha_F := D_F m^{-\frac{1}{2}}(-\log\epsilon')^{\frac{1}{2}}$ and $\alpha_{G-F} := D_{G-F} m^{-\frac{1}{2}}(-\log\epsilon')^{\frac{1}{2}}$. Then with probability $1-2\epsilon'$ for all $h \in [0,a]$,

$$\mathbb{P}[F(X) \leq a+h] \leq \exp\left(-2\frac{(\langle F\rangle - a - h - \alpha_F)_+^2}{D_F^2}\right) \tag{2.77}$$

and

$$\mathbb{P}[G(X,Z) - F(X) \leq -h] \leq \exp\left(-2\frac{(\langle G-F\rangle + h - \alpha_{G-F})_+^2}{D_{G-F}^2}\right). \tag{2.78}$$

It therefore follows that the inequality

$$\inf_h \exp\left(-2\frac{(\langle F\rangle - a - h - \alpha_F])_+^2}{D_F^2}\right) + \exp\left(-2\frac{(\langle G-F\rangle + h - \alpha_{G-F})_+^2}{D_{G-F}^2}\right) \leq \epsilon \tag{2.79}$$

supplies a conservative certification criterion. A near optimal choice for h is given by matching the expressions in the exponentials, with the result

$$h = \frac{(\langle F\rangle - a - \alpha_F)D_{G-F} + (\alpha_{G-F} - \langle G-F\rangle)D_F}{D_F + D_{G-F}} \tag{2.80}$$

whence the certification criterion becomes

$$CF = \frac{M}{U} \geq \sqrt{\log\sqrt{\frac{2}{\epsilon}}} \tag{2.81}$$

with

$$M = \big(\langle F\rangle - a - \alpha_F - \alpha_{G-F} + \langle G - F\rangle\big)_+ \tag{2.82a}$$

$$U = D_F + D_{G-F} \equiv U_A + U_E. \tag{2.82b}$$

It is interesting to note that the total uncertainty U is indeed the sum of an aleatoric uncertainty U_A, measured by the verification diameter D_F, and an epistemic uncertainty U_E, measured by the validation diameter D_{G-F}. Generalizations to the cases of multiple performance measures, $N > 1$, estimated mean performance, and arbitrary admissible set follow along similar lines as those presented in the preceding scenarios.

Observe that the effect of epistemic uncertainty, whether due to lack of fidelity of the model or the existence of unknown unknowns, is to decrease the margin M by $\alpha_{G-F}+\langle G-F\rangle$ and increase the total uncertainty U by D_{G-F}. It should also be carefully noted that the determination of the validation diameter D_{F-G} requires the simultaneous and coordinated execution of the model F and experiments G for equal known parameters X, in order to assess the fidelity of the model, and the repetition of experiments for equal known parameters X, in order to assess the effect of the unknown unknowns. Since experiments are often costly and time consuming, the value and practicality of model-based certification thus depends critically on the ability to quantify epistemic uncertainties by means of a sufficiently small number of experiments. However, for a sufficiently predictive model the function $G-F$ involved in the evaluation of validation diameter D_{G-F} can be expected to exhibit much less variation than either F or G, thus enabling the computation of D_{G-F} by means of a rapidly converging iterative scheme requiring a small number of model evaluations and experiments.

Thus, the precise manner in which a predictive model cuts down on the number of experiments required for certification is by restricting the need of testing to the specific purpose of *validation*, presumably a much less testing-intensive task than purely empirical certification of the system regarded as a black box. Evidently, the requisite number of experiments depends critically on the quality of the model, as well as on the method of optimization used to determine the epistemic uncertainties. An extreme case is furnished by a perfect model that accounts for all input parameters, in which case the quantification of the total uncertainty requires no experiments. In general, the purpose and benefit of model-based certification is to reduce the number of experiments required for certification through the formulation of a sufficiently predictive model. Herein lies the promise, as well as the challenge, of model-based certification.

Fig. 2.1 unambiguously depicts the series of computations required for the various scenarios for certification outlined above. Note that if certification is not achieved, improvements must be made on the model or experimental procedures before repeating the iteration towards certification.

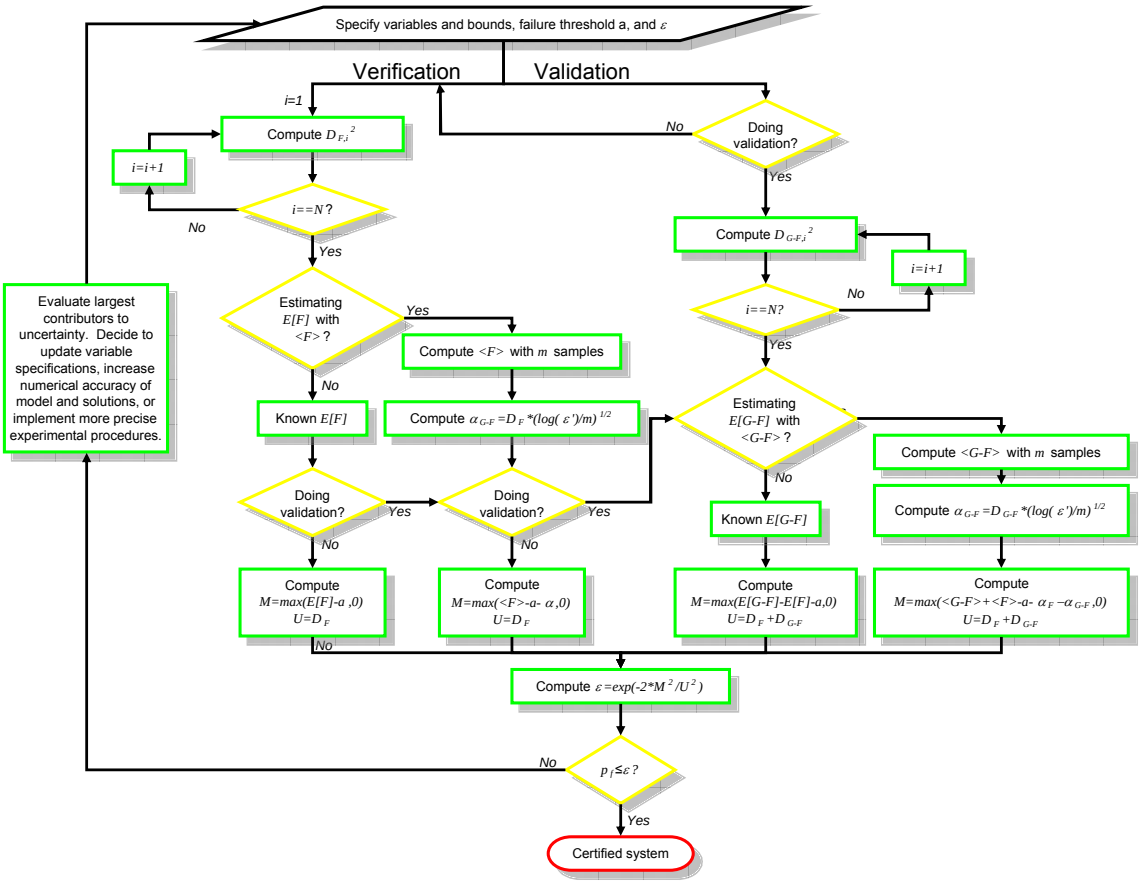


Figure 2.1: Flowchart of the rigorous steps of V&V suggested by UQ via CoM with switches for the scenarios discussed in §2.2

Chapter 3

Applications

The applications in this chapter serve to illustrate the practicality and implementation of the theory developed for the concentration-of-measure UQ framework in the previous chapter.

3.1 Linear System

This example is meant to showcase a verification computation according to §2.2.3 that mocks a stochastic simulation typically encountered in computational engineering applications where random inputs are propagated through a linear algebra computation. This example is expandable to take arbitrarily many random inputs that are entries of a stochastic linear system

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}. \tag{3.1}$$

Take the system $\mathbf{A}x = b$, where $\mathbf{A}$ is size $n \times n$ and symmetric, and x and b are size $n \times 1$. The unique entries in $\mathbf{A}$ and b are random, resulting in $N = \frac{n(n+1)}{2} + n$ random

variables, which are scaled according to n. Solving for x yields a column vector from which the performance measure, $y = \max(x)$ is taken.

Consider $\mathbf{A}$ likened to a stiffness matrix constructed in a finite element scheme, where one wishes to compute the nodal displacements x, given the forces b at each node. A measure of performance for the system may be its maximal nodal displacement, consistent with $y = \max(x)$. For example, consider the nodal stiffness matrix of a nonperfect material (i.e., metal with imperfections or grain structure), represented by $\mathbf{A}$ that may have some uncertainty reflecting these inhomogeneities in its entries. Also consider that the forces b on the nodes may have uncertainty due to imperfections (e.g., machine precision or bulk material loading uncertainties) in the numerical model. In this regard, this example represents the core computation of a typical solid mechanics solution algorithm.

Some numerical considerations should be highlighted. To obtain a nontrivial solution ($x = 0$), $\mathbf{A}$ and b are controlled to ensure positive definiteness by adding 1 to each entry of b and to the diagonal entries of $\mathbf{A}$. The bounds given to the random entries in b and $\mathbf{A}$ are normalized by n to maintain consistent order-of-magnitude results when n changes.

3.1.1 Results

Two cases were run using $n = 50$ ($N = 1325$), where the bounds on each random entry is a uniform distribution between $[-h, h]$, where h =1.0e-2 in case 1 and $h = 5.0$ in case 2. These cases will serve to illustrate the effect of concentration-of-measure for tightly bounded (case 1) and loosely bounded (case 2) random variables affecting a system. Intuitively, case 1 will exhibit lower uncertainty interpreted through diameters and a tighter concentration of Monte Carlo samples about its mean.

If not otherwise specified, all computations in this work were done on the following

machine:

- **Architecture**: Intel x86_64 Linux server
- **Memory**: 32 GB DDR2-667 main memory, 16 x 2 GB
- **Processors**: 4 Intel®Xeon ™
- **Disk**: 500 GB Serial ATA, 7200 RPM HD w/16 MB Databurst Cache
- **File System**: NTFS, factory
- **Operating System**: Fedora Core Linux 2.6.13-1.1526_FC4smp

Table 3.1: Concentration-of-measure UQ results for the linear system application

Case	Diameter (D^2)	Mean ($\mathbb{E}[y]$)
1	1.256e-4	1.00183
2	2.277e2	2.87929

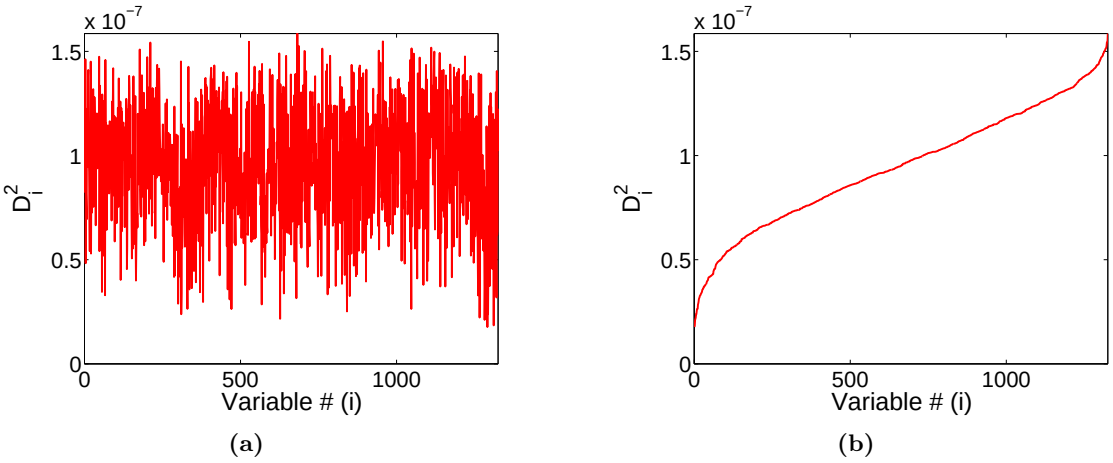
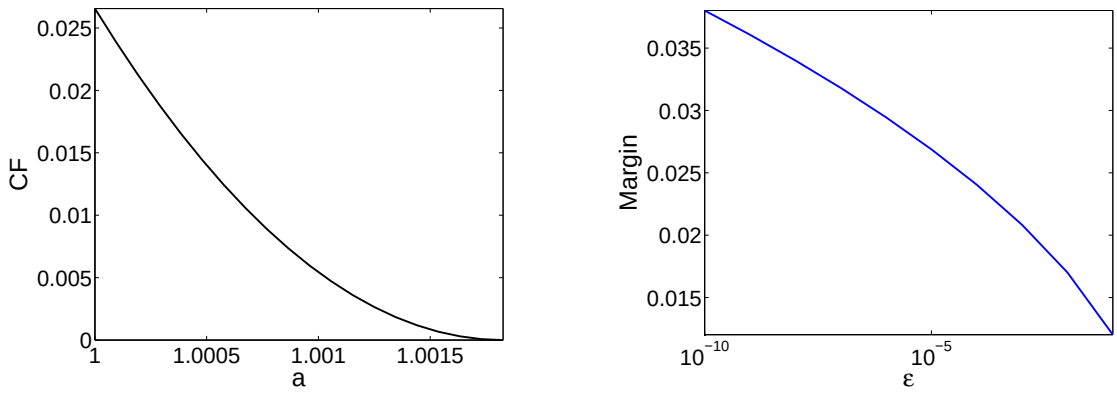
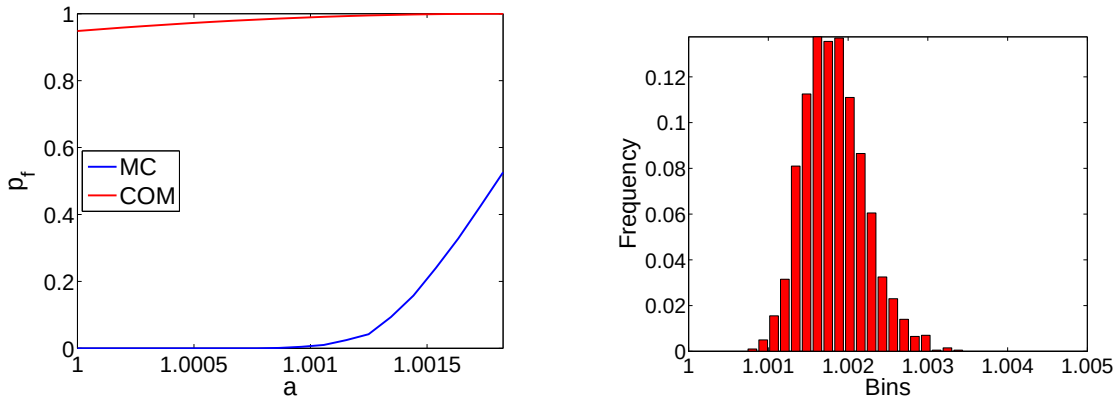


Figure 3.1: Case h =1e-2: (a) Individual diameters according to random variable number and (b) diameters in ascending order



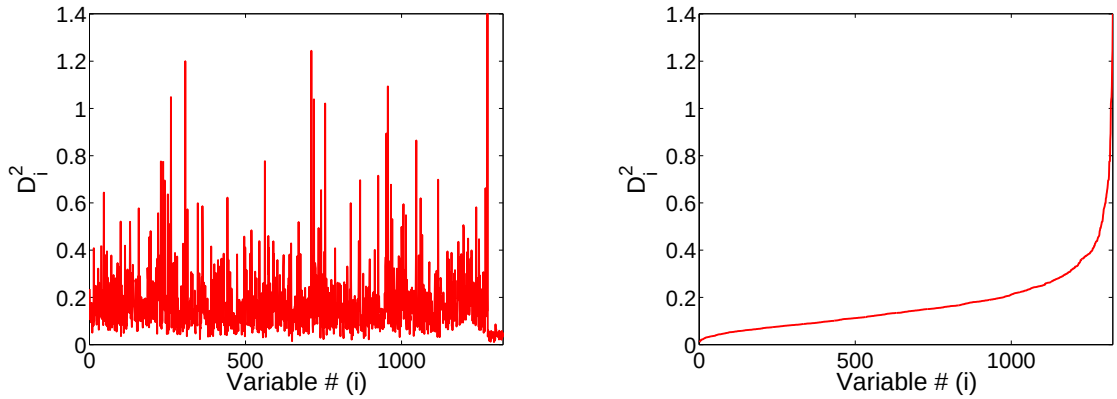
(a) (b)

Figure 3.2: Case h =1e-2: (a) Certification factor versus failure tolerance a and (b) margins required for certification at a level ϵ



(a) (b)

Figure 3.3: Case h =1e-2: (a) Probability-of-failure versus failure threshold a and (b) a histogram of performance measures using 2000 samples



(a) (b)

Figure 3.4: Case $h = 5$: (a) Individual diameters according to random variable number and (b) individual diameters in ascending order

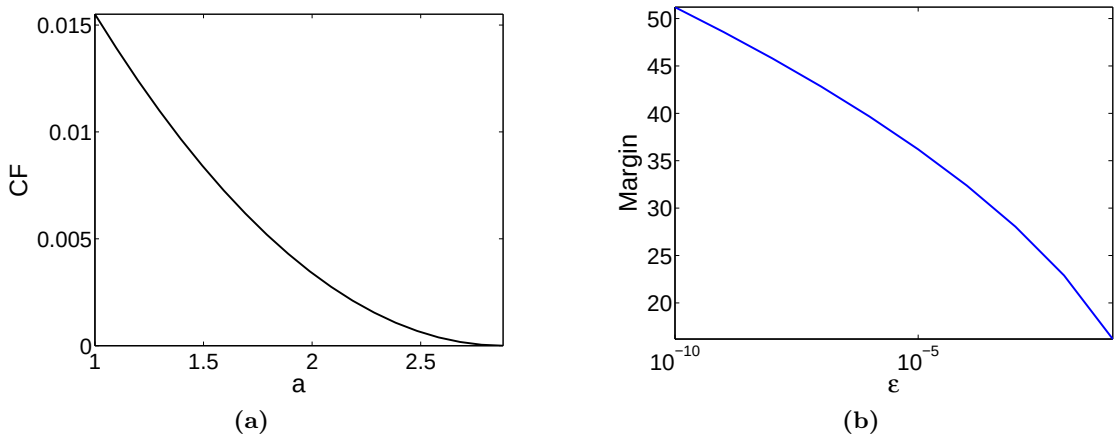


Figure 3.5: Case $h = 5$: (a) Certification factor versus failure threshold a and (b) margins needed for certification at a level ϵ

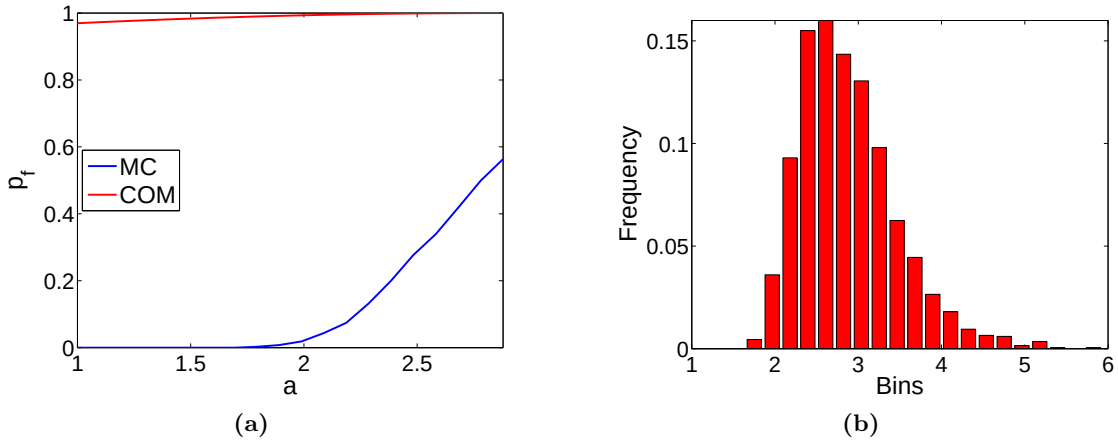


Figure 3.6: Case $h = 5$: (a) Probability-of-failure versus failure threshold a and (b) a histogram of performance measures using 2000 samples

3.1.2 Discussion

The figures above are typical presentations of results for a concentration-of-measure UQ analysis, illustrating individual diameters (Fig. 3.1(a) and Fig. 3.1(b)), confidence factors for various failure thresholds (Fig. 3.2(a)), margins required for various certification criteria (Fig. 3.2(b)), comparison of p_f estimates from concentration-of-measure to Monte Carlo sampling for various failure thresholds (Fig. 3.3(a)), and a histogram of Monte Carlo samples (Fig. 3.3(b)). Note that Scenario 1 (§2.2.1) inequalities are used here, assuming that the mean calculated is the exact mean.

Fig. 3.1(a) and Fig. 3.4(a) show the magnitude of diameters associated to individual random variables — the largest individual "subdiameters" can be interpreted as the random variables that contribute most significantly to the total uncertainty. Note the intuitive lower order of magnitude for subdiameters in case 1 compared to case 2 — larger uncertainty in inputs generally results in more uncertainty in the output as collected by subdiameters into a total diameter. An ordering of subdiameter values in Fig. 3.1(b) and Fig. 3.4(b) exemplifies some regularity of the magnitude of subdiameters associated to different random variables.

Fig. 3.2(a) and Fig. 3.5(a) show the confidence factor computed for increasing values of failure threshold. As a tends closer towards $\mathbb{E}[y]$, the margin M decreases while the total diameter D^2 is constant, resulting in a lower confidence factor. Fig. 3.2(a) exhibits higher confidence factors for the range of a compared to Fig. 3.5(a) mostly due to the diameter value discrepancies, and partly due to the different magnitudes of margins dictated by the magnitude of the mean.

Fig. 3.2(b) and Fig. 3.5(b) show the margins needed to recover a failure probability upper bounded by ϵ — this allows a system designer to interpret how much further the system mean needs to be from the failure threshold ($M = (\mathbb{E}[y] - a)_+$) to recover a desired p_f. Note that

noncertifiable systems may require a margin greater than the maximum difference between a mean value and legitimate failure threshold, demonstrating the noncertifiable nature of the system.

Fig. 3.3(a) and Fig. 3.6(a) show the estimates of p_f for various failure thresholds a using Monte Carlo sampling and CoM. Note that the CoM estimate is usually higher than Monte Carlo sampling estimates when enough samples are available to do Monte Carlo analyses (see the end of §5 for a discussion of cases which one should and shouldn't use CoM to do UQ) and exhibits the conservative nature of CoM p_f estimates.

Fig. 3.3(b) and Fig. 3.6(b) show the probability densities of the outcomes for the given ranges of random variable inputs through histograms using 2000 Monte Carlo samples. Note the abscissa values show that case 1 is more tightly centered about its mean that case 2. These figures should logically correspond to the p_f values computed by Monte Carlo sampling in Fig. 3.3(a) and Fig. 3.6(a).

The meticulous analysis above, stemming from computations of the diameter and mean, shows that the CoM framework provides important insight into the interpretation of system uncertainties and performance. Such analyses are critical in any CoM UQ computation to completely understanding the effect of system performance by uncertainties, and will be provided for each application to follow.

3.2 Ring Explosion

This section applies, by way of demonstration, the concentration-of-measure approach to the problem of predicting the state of maximum compression to an imploding/exploding ring. The behavior of the ring in the high-energy regime is strongly nonlinear and dynamical and the calculation of performance measures requires the solution of an initial value problem in

time. Because the system undergoes multiple bifurcations and generally traverses a rough energy landscape, the performance measures are expected to depend sensitively on initial conditions and on the parameters of the system. In addition, in the high-energy regime the shape of the ring at maximum compression is expected to be highly crumpled and irregular. Thus, while straightforward in its definition, the example of an imploding ring does pose a nontrivial and illuminating test of the theory.

That the certification criterion of Eq. (2.62) is indeed conservative, i.e., that if Eq. (2.62) is satisfied then the probability-of-failure is indeed less than ϵ with probability $1 - \epsilon'$, has been rigorously proven mathematically. However, several practical questions remain to be ascertained. A first question concerns whether the bounds furnished by concentration-of-measure are tight enough to supply a practical means of certification. A second question concerns the means of calculation of the aleatoric and epistemic uncertainties, as measured by their corresponding verification and validation diameters, including the relative efficiencies of optimization algorithms and the number of system evaluations and experimental tests required for the computation of the uncertainty measures. The imploding ring example presented in this section sheds useful light on these and other related questions.

3.2.1 Test case description

Consider a ring described by means of a bead model consisting of n point masses interacting through two- and three-body potentials. The ring starts from an equilibrium circular configuration and is imparted an inward radial velocity. The objective of the analysis is to characterize the state of the ring at the point of maximum compression. Refer to [37] for details of the numerical modeling of the ring.

A typical trajectory of the ring is shown in Fig. 3.8. As may be observed in the figure, the

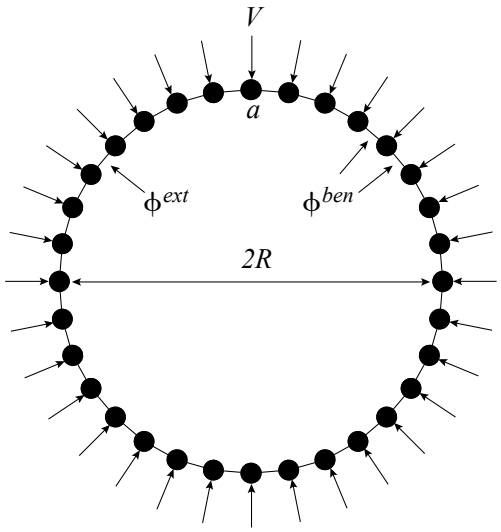


Figure 3.7: Schematic of ring implosion and explosion test

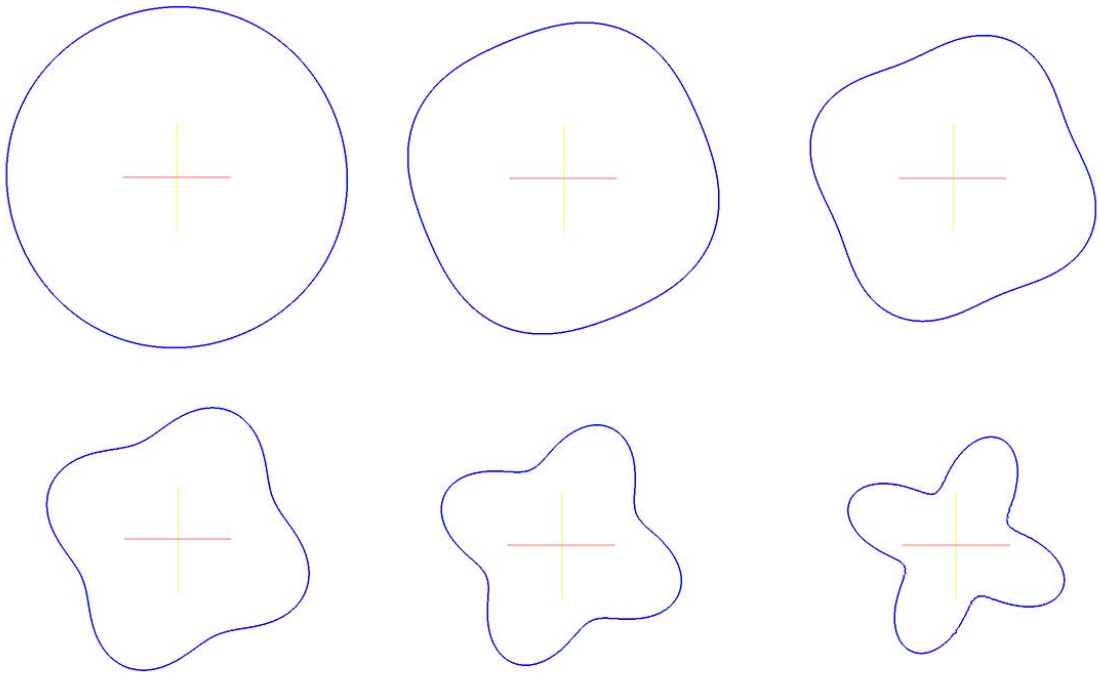


Figure 3.8: Ring implosion test. Crumpling of circular ring resulting from implosion (left to right, top to bottom)

ring buckles early on and subsequently undergoes extensive crumpling, with the amplitude of the resulting crumpling increasing monotonically up to the point of maximum compression.

Suppose that the objective of the analysis is to predict the bending energy E^{ben} at the point of maximum compression of the trajectory. A natural measure of the extent of compression of the ring is supplied by the extensional energy E^{ext}. Likewise, the bending energy E^{ben} provides a natural measure of the degree of crumpling of the ring in compression. Therefore, an appropriate performance measure is given by the value of the bending energy E^{ben} of the ring at the time the extensional energy E^{ext} attains its first maximum. Thus, simulations stop when E^{ext} attains its first maximum in time, at which point the value of E^{ben} is recorded as the single performance measure Y of interest. Additionally assume that the proper operation of the system requires Y to be above a certain threshold a. Thus, the system fails when the bending energy achieved at the point of maximum compression falls short of the threshold value.

The calculations described earlier implicitly define a response function $Y = F(X)$ that returns the performance measure Y as a function of the parameters X of the system. These parameters include material constants, geometry, loading, initial conditions, and numerical parameters such as the time step used for numerical integration. For definiteness, regard all parameters to be certain, and therefore treat them as fixed constants, with the exception of the bending stiffnesses D of the bond-pairs in the ring, which are assumed to be uncertain. Specifically, divide the circumference of the ring into M segments of equal length. Each of the beads in a particular section $i = 1, \ldots, M$ is assigned a variable value X_i for the bending stiffness within a certain range. In order to investigate the effect of the range of the inputs on the probability-of-failure estimates, the bending stiffness ranges considered are $[0.95, 1.05]$, $[0.99, 1.01]$, $[0.995, 1.005]$, and $[0.999, 1.001]$. In addition, in order to examine the

Table 3.2: Values of fixed parameters used in imploding ring calculations

Parameter	Value
Number of beads (n)	256
Extensional stiffness (C)	1.0
Individual bead mass density (m)	1.0
Ring radius (R)	0.25
Initial time step (Δt)	6.14×10^{-4}
Initial velocity fluctuation (ΔV)	$V_0/3.0$
Initial velocity wave number (k)	4
Minimum time step	1.0×10^{-6}
Richardson extrapolation energy tolerance	1.0×10^{-3}

corresponding effect of the nonlinearity of the model, the mean initial velocities considered are $V_0 = 3$, 3×10^{-1}, 3×10^{-2}, and 3×10^{-3}. The remaining values of the parameters used in calculations are recorded in Table 3.2.

3.2.2 Uncertainty quantification analysis

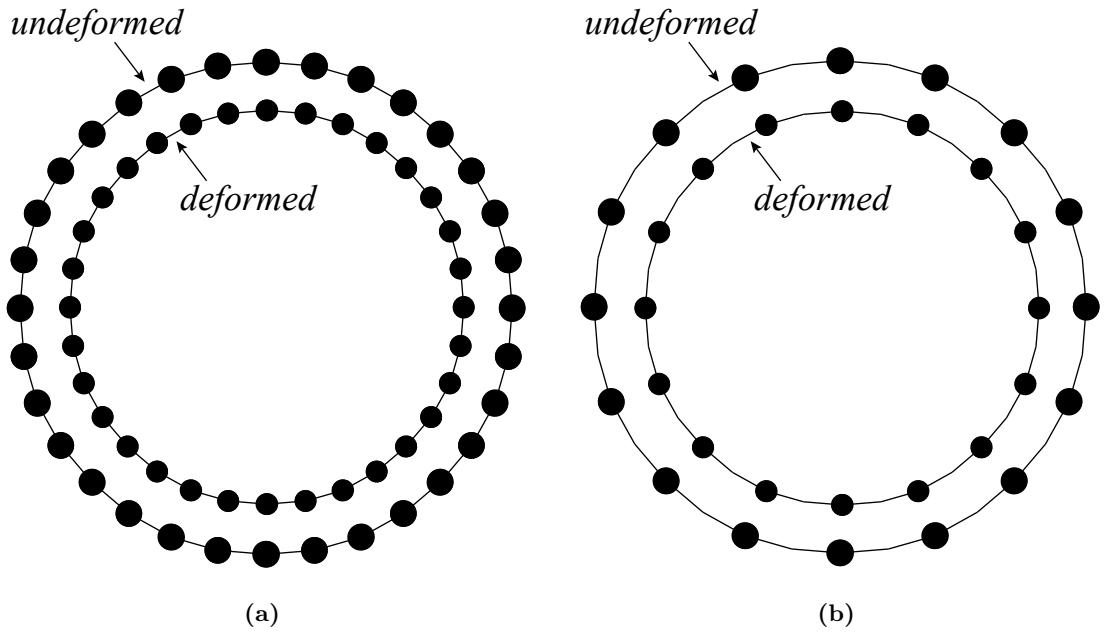


Figure 3.9: Ring configurations considered in the definition of the coarse-grained model. (a) Exact model; (b) Coarse-grained model

The fundamental calculations to be carried out for purposes of uncertainty quantification

are: the calculation of the estimated mean performance $\langle Y \rangle$ as in Eq. (2.56); the verification diameter D_F of the response function as in Eq. (2.31); and the validation diameter D_{G-F} as in Eq. (2.76). In lieu of experimental data, G is taken as the response function of a higher resolution ring model discretized with 768 beads, with F then representing a coarse ring discretized with 256 beads obtained by coarse-graining G. As in the case of actual experimental data, the evaluations of G are expensive, which places a premium on optimization methods that require the least number of evaluations of G. Considerations are made towards appropriate scaling of masses, and bending and axial stiffnesses between F and G as detailed in [37].

For large deviations from a perfect circular configuration, the coarse-grained model F deviates significantly from the exact model G, especially as a result of the strong nonlinear dependence of the bending energy on the bond-pair angle and of the inability of the coarse-grained geometry to resolve fine wrinkles in the deformed configuration of the exact model. It bears emphasis that the upscaling model described in [37] is not proposed as an accurate method of coarse-graining but, contrariwise, it is intended as a rough approximation for purposes of ascertaining the uncertainties introduced by modeling assumptions.

In addition, in order to introduce unknown unknowns of the type that are likely to be encountered in practice, i.e., arising from fine unresolved length scales, introduce nodal mass perturbations in fine model G that are subgrid relative to the coarse model F, i.e., preserve the aggregate masses of all beads in F. This constraint is satisfied by adding a mass Z to the beads of G explicitly accounted for in F, and subtracting masses $Z/2$ to the remaining beads in G. This distribution of mass over the fine model is indeed undetectable on the level of resolution of the coarse model. In this manner, the added mass Z becomes the sole unknown unknown of the system. In calculations, the added mass Z is either zero

or is varied bead-wise so that the variation of the total mass of the ring is in the range $[-0.1, 0.1]$.

The estimation of the mean performance $\langle Y \rangle$ can be carried out simply by means of Monte Carlo sampling. The computation of the diameters D_F and D_{G-F} requires an optimization over parameter space. Owing to the roughness of the energy function landscape of the ring and the lack of explicit derivatives of the response function, global optimization methods such as genetic algorithms and simulated annealing naturally suggest themselves in the computation of D_F, but gradient methods are also worth exploring if the function exhibits regularity in the design space. All calculations employ the quasi-Newton method and genetic algorithms implemented in Sandia National Laboratories' DAKOTA Version 4.0 software package. By contrast, the computation of D_{G-F} requires costly experimental tests and global optimization algorithms are often not viable due to their slow convergence. The expectation, however, is that for a sufficiently high-fidelity model the function $G-F$ exhibits much less variation than either F or G, and that, therefore, the computation of D_{G-F} can be carried out by means of rapidly converging iterative schemes such as a quasi-Newton iteration.

3.2.3 Results

3.2.3.1 Perfect model

Begin by supposing that the model F is perfect as in §2.2.1. Fig. 3.10 depicts the dependence of the distribution in performance on the number of random inputs. The distributions are calculated directly by 10^3 Monte Carlo samples, with the sampling carried out in two ways: assuming a uniform distribution of the inputs over their intervals of variation, referred to as "uniform distribution" in the figure; and assuming that the inputs are at the end

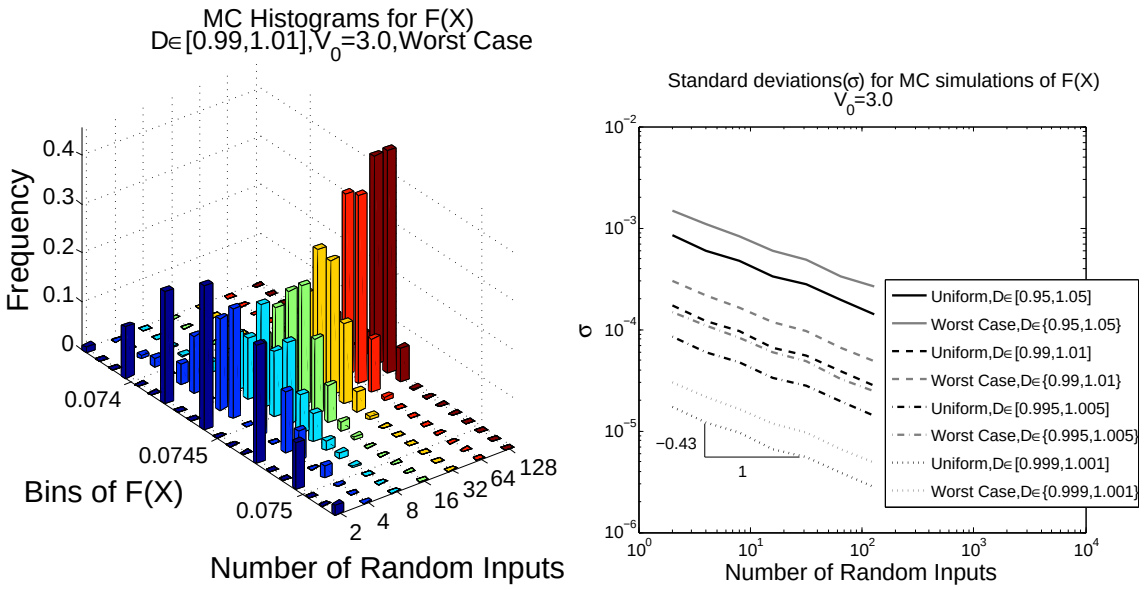


Figure 3.10: Monte Carlo calculations of the distribution of the performance measure as a function of the number of random variable inputs of the system. (a) Histograms of the performance measure. (b) Standard deviation (σ) on the distribution

points of their intervals of variation, referred to as "worst case" in the figure. As expected, uniform sampling results in a lower probability-of-failure than worst-case sampling. All Monte Carlo calculations presented subsequently are carried out using worst-case sampling. The standard deviation of the distribution is found to scale as $M^{-0.43}$ with the number of random variable inputs M, Fig. 3.10b. A marked concentration of the performance histogram as the number of inputs is increased is clearly evident in the figure, which vividly demonstrates the concentration-of-measure phenomenon on which the present approach is predicated.

The dependence of the verification diameter D_F of the response function on the number M of input parameters is shown in Fig. 3.11. For linear F one has the scaling $D_F \sim M^{-\frac{1}{2}}$, and hence the power-law behavior ($D_F \sim M^{-0.4}$ in the quasi-Newton calculations, $D_F \sim M^{-0.43}$ in the genetic algorithm calculations) evident in Fig. 3.11 owes

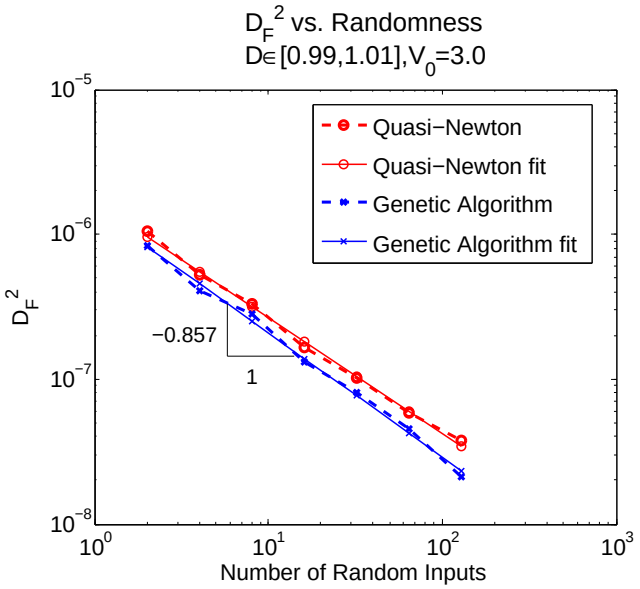


Figure 3.11: Ring implosion test. Dependence of the verification diameter (as computed by either quasi-Newton or a genetic algorithm) on the number of random inputs of the system, showing clear power-law behavior

to the lack of regularity of the response function $F(X)$. This scaling behavior serves to illustrate one of the principal strengths of the concentration-of-measure approach, namely, that the uncertainty bounds become sharper as the number of random variables increases. Therefore, concentration-of-measure bounds are particularly attractive in the context of certification of large complex systems with many random inputs. It is also interesting to note from Fig. 3.11 that both the genetic algorithm and quasi-Newton results exhibit ostensibly identical behavior.

The margin plots in Fig. 3.12 below for Scenario 1 (and Scenario 3 in Fig. 3.13 and Scenario 5 in Figs. 3.17 and 3.18) apply an initial implosion velocity $V_0 = 3$. They show the dependence of the required margin on the number of random inputs for four input ranges: $D = [0.95, 1.05]$, $[0.99, 1.01]$, $[0.995, 1.005]$, and $[0.999, 1.001]$ for the coarse model F, with corresponding intervals for the fine model G for Scenario 5 figures in Figs. 3.17 and 3.18.

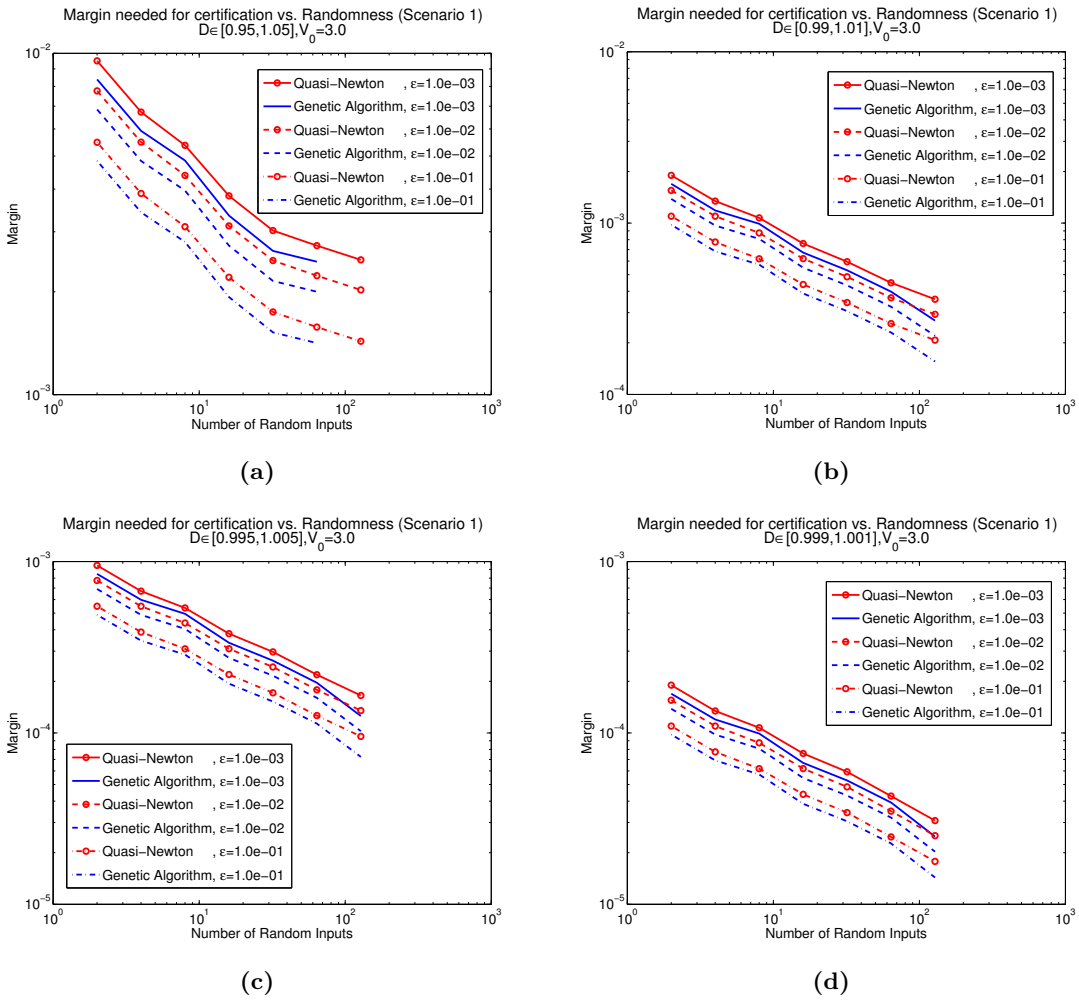


Figure 3.12: Minimum design margins required in order to guarantee various probabilities-of-failure for a perfect model and exact mean performance case

A number of features of the results presented in Fig. 3.12 and Fig. 3.13 immediately come to prominence. Interestingly, despite the extremely nonlinear behavior of the system the margins required to guarantee its safe performance take modest values which, presumably, should be attainable by practical designs. The modest range and good behavior of those design margins illustrates the feasibility and practicality of concentration-of-measure inequalities as a basis for the certification of complex systems. A second feature of interest is the steady decrease of the design margins with the number of random inputs, a decrease that is in keeping with the expected behavior of concentration-of-measure inequal-

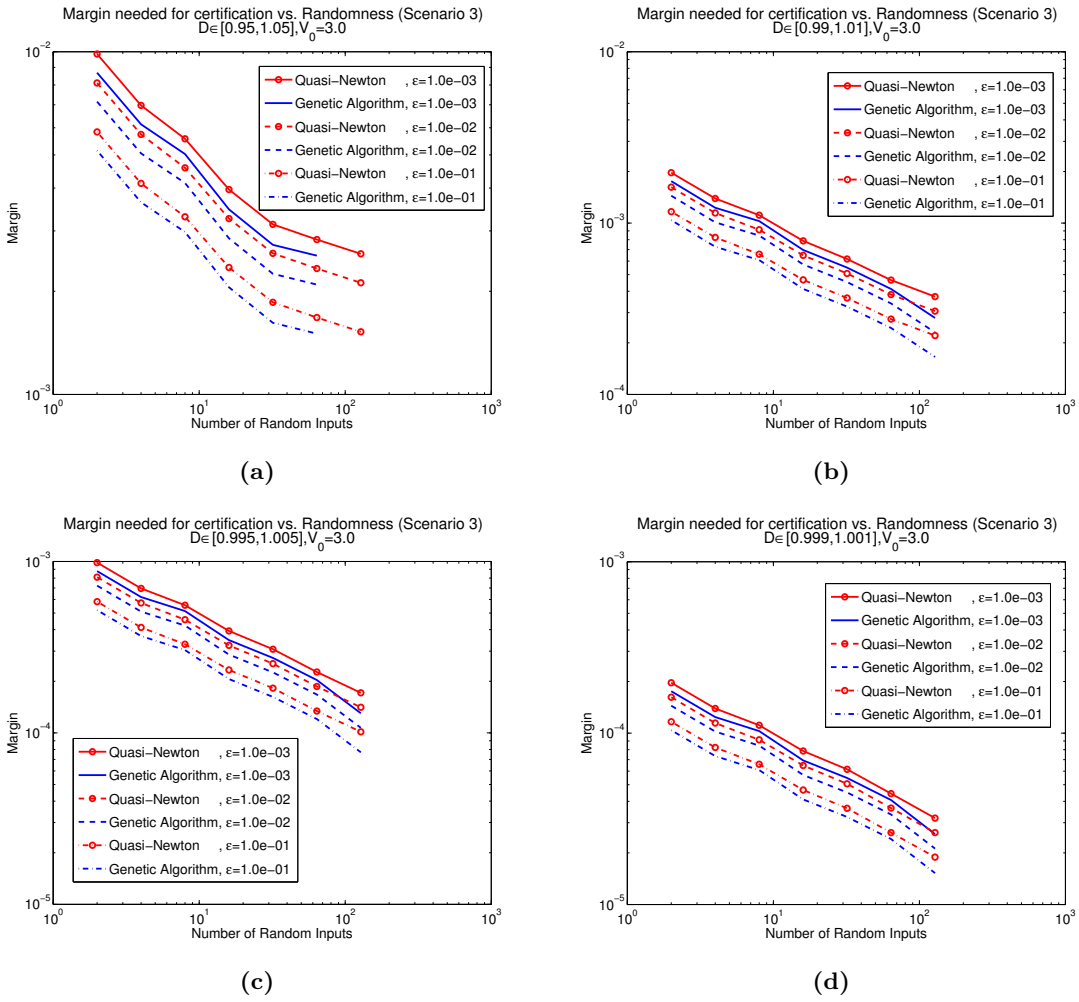


Figure 3.13: Minimum design margins required in order to guarantee various probabilities-of-failure for a perfect model and estimated mean performance case

ities. Thus, the concentration-of-measure approach to certification pursued in the present work, which is based on the simple Hoeffding inequality, is particularly effective for systems with a large number of uncorrelated or weakly correlated random inputs. Extensions of concentration-of-measure inequalities to correlated inputs are summarized in Appendix A but will not be pursued here in the interest of simplicity. A final feature of interest, which is evident from a comparison of Fig. 3.12 and Fig. 3.13, concerns the small increase of the required design margins that results from estimating the mean performance of the system empirically. This small effect suggests that in practice mean performances can be safely

estimated empirically and that the mean performance of the system need not be known exactly.

3.2.3.2 Inexact model

Next is the investigation of the effect of modeling uncertainties, including unknown unknowns. The precise manner in which the exact model is coarse-grained and unknown unknowns are introduced has been described in §3.2.1. The determination of the residual probability-of-failure tolerance requires the estimation of the mean deviation $\langle G_i - F_i \rangle$ between predicted and measured performance measures and the computation of the validation diameters $D_{G_i - F_i}$. Again, note that this determination requires the simultaneous execution of nominally identical calculations and experiments, a process that gives precise form to the notion of validation.

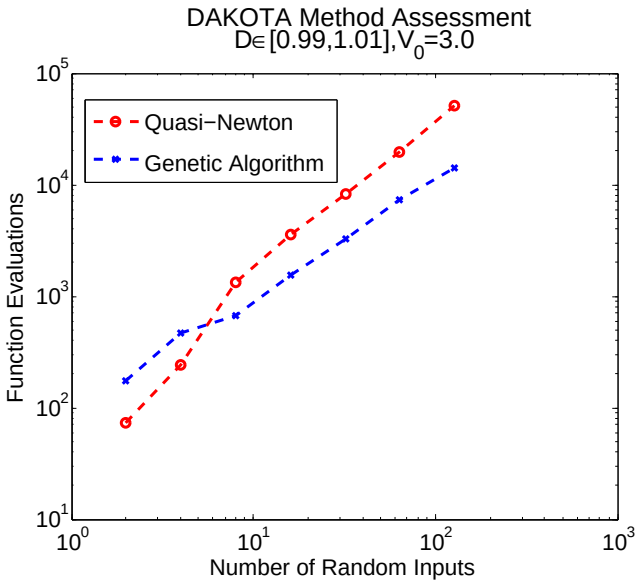


Figure 3.14: Comparison of the number of objective function evaluations to convergence of the validation diameter D_{G-F} using a quasi-Newton and a genetic algorithm as function of the number of random input parameters for the imploding ring case

As stated in §2.2.5, in the context of certification, models are useful precisely because,

for sufficiently predictive models, the objective function required to compute the validation diameter, Eq. (2.76), may be expected to exhibit much less variation than the response function itself, with the result that the evaluation of the validation diameter can be based on rapidly-convergent iterative procedures, which in turn can greatly reduce the number of experimental tests required for certification. This working assumption may be tested in the case of the imploding ring example. Fig. 3.14 shows plots of the number of iterations to convergence using a quasi-Newton iteration and a genetic algorithm as function of the number of random input parameters. As may be seen in the figure, for a small number of input parameters the quasi-Newton iteration converges more rapidly than the genetic algorithm and the number of iterations to convergence is manageably small. As the complexity of the system increases, the performance gap between the genetic algorithm and the quasi-Newton iteration narrow and, for a large number of input parameters the genetic algorithm requires fewer iterations to convergence.

In a certification context, the effect of the limited model fidelity is an effective reduction in the probability-of-failure tolerance (see §2.2.5). Thus, the computed probability-of-failure must now be compared against a reduced tolerance that accounts for modeling errors. Fig. 3.15 shows that, for the particular choice of upscaling model used in the calculations, the validation diameter D_{F-G} is greatly in excess of the verification diameter D_F. Thus, the epistemic or modeling uncertainty U_E completely overshadows aleatoric uncertainty U_A resulting from the intrinsic variability of the inputs and any stochastic nature of the system. The reason for the large epistemic uncertainty is clearly illustrated in Fig. 3.16, which shows that fine crumpling of the ring occurs in the late stages of compression, as in Fig. 3.16(a). This localized crumpling greatly contributes to the bending energy of the ring, i. e., to the chosen measure of performance. However, the localized crumpling cannot be resolved by

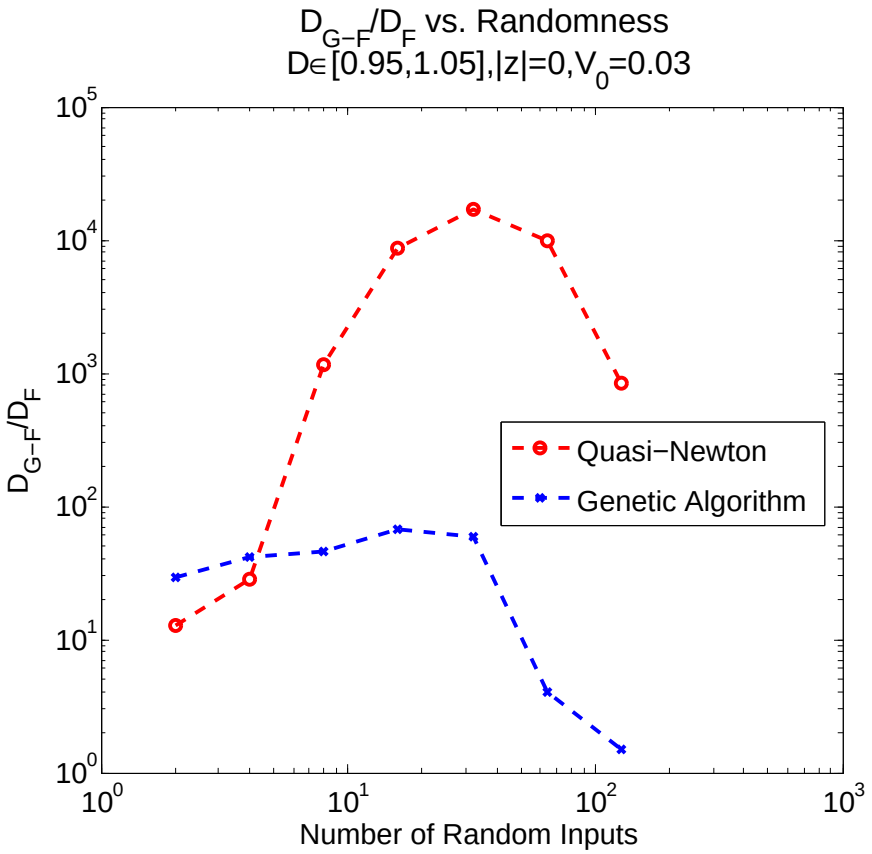
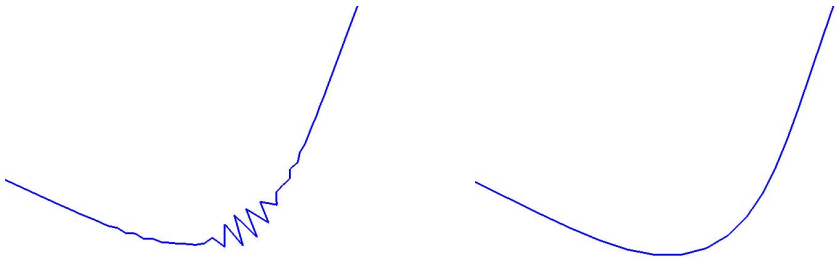


Figure 3.15: Ring implosion test. Ratio of validation diameter, D_{G-F}, to verification diameter, D_F

the coarse model and is completely suppressed, as in Fig. 3.16(b), with the result that the bending energy of the ring is greatly underestimated.



(a) **(b)**

Figure 3.16: Closeup of a snapshot of the ring trajectory showing: (a) localized crumpling in the exact system represented by $G(X, Z)$ and (b) the failure of the coarse-grained model $F(X)$ to resolve the localized crumpling

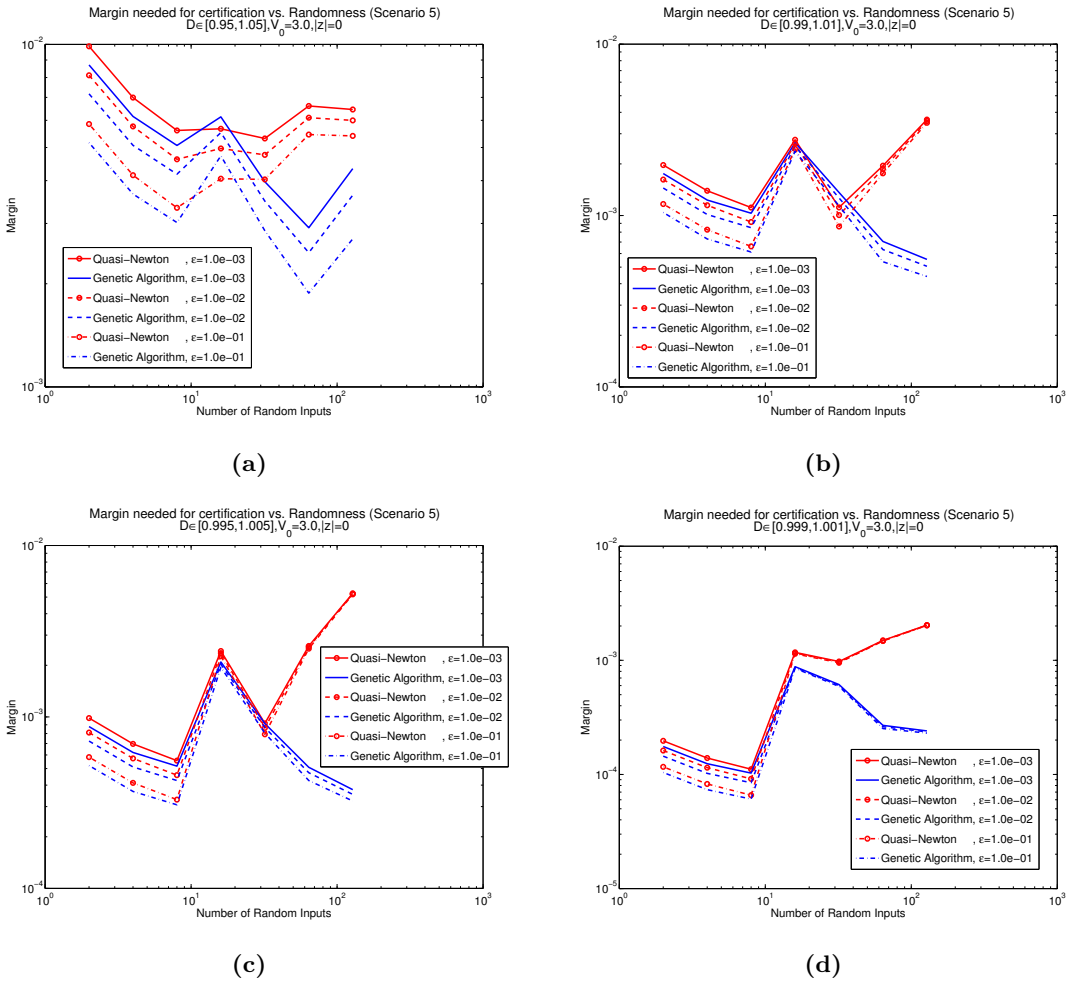


Figure 3.17: Minimum design margins required in order to guarantee various probabilities-of-failure for an inexact model, estimated mean performance, and no unknown unknowns case

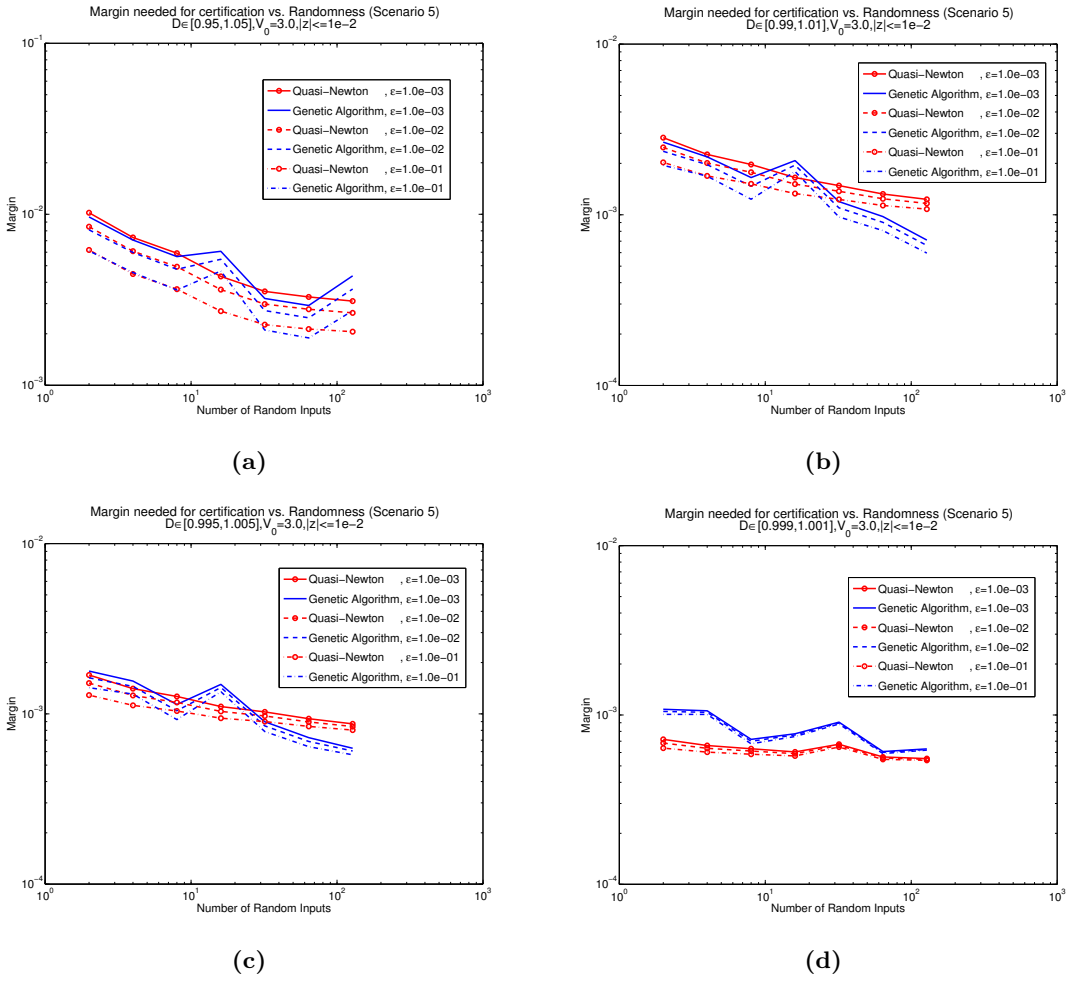


Figure 3.18: Minimum design margins required in order to guarantee various probabilities-of-failure for an inexact model, estimated mean performance, and unknown unknowns added to the system in the form of added bead masses resulting total ring mass variations in the range $[-0.1, 0.1]$

In order to compensate for the large epistemic uncertainty, the certification of the system requires larger margins than otherwise required in the case of the perfect model (see Fig. 3.17). This is particularly so for large numbers of random input parameters owing to the lack of concentration exhibited by the inexact model, a behavior that is in sharp contrast to the perfect model which, as noted earlier, exhibits strong concentration with increasing system size. The behavior of the system is not significantly altered by the introduction of unknown unknowns (see Fig. 3.18). Remarkably, despite the large level of epistemic uncertainty and the introduction of unknown unknowns the minimum margins required for certification remain within relatively modest values attainable by practical designs. Thus, while the concentration advantage may be diminished, concentration-of-measure inequalities remain a viable avenue for certification in the context of inexact models and unknown unknowns.

3.3 Optimal Control Strategies

This example seeks optimal controls that minimize a probability-of-failure upper bound supplied by concentration-of-measure inequalities. The resulting optimal controls then maximize the design margin M, or reduce the uncertainty U in the operation of the system, as measured by the system diameter, or both. Optimal control within this UQ framework can be thought of as concentration-of-measure optimal control, or COMOC.

This section assesses COMOC in a specific area of application: positioning accuracy of robotic arm maneuvers, where the robotic arms are modeled as interconnected three-dimensional rigid bodies [36, 35]. Uncertainty is introduced by first assuming that the lengths of the arms are random, and secondly the system also experiences random forcing due to side wind. A version of McDiarmid's concentration inequality is used here where the

mean of the system performance is estimated through an empirical average as in §2.2.3.

Investigated here is a particular robot arm maneuver whose successful operation requires a minimum arm tip positioning accuracy, where the deterministic analysis of the nominal geometry of the system without wind forces can be found in [34], as well as details of the discrete mechanics and optimal control for constrained systems (DMOCC) methodology. Results of numerical experiments are collected in §3.3.1. In the particular example under consideration, COMOC reduces the concentration-of-measure probability-of-failure upper bound by about one order of magnitude with respect to the deterministic optimal control.

3.3.1 Test case: Minimizing the probability-of-failure for a robot arm maneuver

The deterministic robot arm maneuver from [34] is considered first in the presence of geometrical uncertain arm lengths then in uncertain operating conditions represented by the presence of uncertain wind forces in addition to the uncertain lengths. All calculations use the reduced variational time-stepping scheme in [34] obtained via the discrete null space method with nodal reparameterization. Further DMOCC considerations for this maneuver are detailed in [34].

The performance measure Y is the placement accuracy of the arm tip, i.e., the distance from the arm tip and its prescribed location $\boldsymbol{x}_H$ at the end of the maneuver's duration of $t_N = 1.5$ units of time. Thus, in this case $Y = \|\boldsymbol{x}_N - \boldsymbol{x}_H\|$ is obtained for a candidate control sequence $\boldsymbol{\tau}_d$ by stepping forward in time and the initial conditions both described in [34]. From an optimal control point of view, this is similar to a shooting method. Let a be the acceptable radius of the arm tip from the desired point $\boldsymbol{x}_H$ — whereby acceptable operation is $Y \leq a$ and failure is $Y > a$. The goal is to find a control sequence $\boldsymbol{\tau}_d$ for which

the probability-of-failure $\mathbb{P}[Y > a]$ is minimal. Often, however, the probability $\mathbb{P}[Y > a]$ is not known explicitly. In these cases, one seeks instead to minimize a concentration-of-measure upper bound of the probability-of-failure. The resulting objective function to be minimized is the probability-of-failure for an assumed perfect model with estimated mean, as in §2.2.3

$$p_f = \exp\left(-2\frac{M^2}{U^2}\right) = \exp\left(-2\frac{(a - \langle Y\rangle - \alpha)_+^2}{D_F^2}\right) \tag{3.2}$$

This choice of objective function in for the optimal control $\boldsymbol{\tau}_d$ seeks to maximize *confidence* in the safe operation of the system either by increasing the placement margin M, i.e., by decreasing $\langle Y\rangle$, or by reducing the uncertainty U of the maneuver, i.e., by reducing the diameter D_F, or the collective ratio.

Each objective function evaluation of Eq. (3.2) requires the computation of the empirical average response $\langle Y\rangle$ and the diameter D_F. For all calculations in this example, the mean response is computed by Monte Carlo sampling and the system diameter and optimal controls $\boldsymbol{\tau}_d$ are computed by separate simulated annealing schemes described in §A.1. The starting controls for the algorithm are set to the deterministic controls computed in [34].

3.3.1.1 Uncertain Geometry

Begin with $M = 2$ uncertain lengths where l^1 can vary randomly in a range of 5% and l^2 varies randomly in the range of 0.1% around the given value, respectively. These values ensure that their influence on the system's uncertainty is of the same order of magnitude.

Fig. 3.19 shows the mean performance evolution, system diameter, and concentration-of-measure probability-of-failure upper bound along the simulated-annealing algorithm evolution for the determination of the optimal controls. Expectedly, both the positioning accuracy of the maneuver, measured by the mean response $\langle Y\rangle$ with $m = 100$, and the

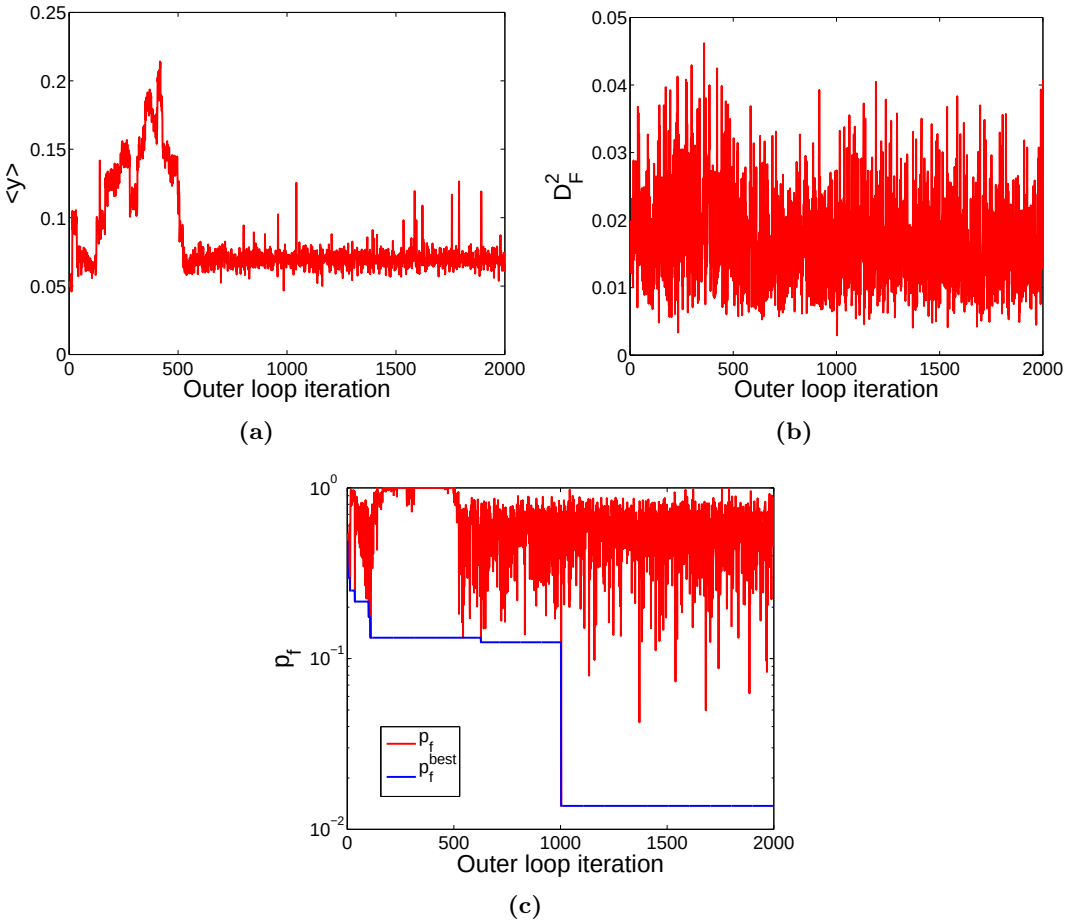


Figure 3.19: Uncertain geometry: Simulated annealing algorithm evolution for the determination of the optimal controls. Evolution of: (a) mean performance (b) system diameter and (c) concentration-of-measure probability-of-failure upper bound

uncertainty in the operation of the maneuver, measured by the diameter D_F, show a decreasing tendency. The concentration-of-measure probability-of-failure upper bound correspondingly decreases from $p_f = 0.49$ to $p_f^{\text{best}} = 0.013$ with $\langle Y \rangle = 0.0693$ and $D_F^2 = 0.00292$ corresponding to the best control sequence. This reduction in probability-of-failure can be interpreted as an increase in the *confidence* that may be placed in the acceptable operation of the maneuver, as measured by the confidence factor. For the optimal control sequence with $p_f^{\text{best}} = 0.013$ computed via Eq. (2.61), the mean has been recomputed with $m = 10000$ such that one can assume it to be exactly the mean performance of the system. Then using

Eq. (2.32) (with $M = (a - \mathbb{E}[Y])_+$), $\mathbb{E}[Y]$ is computed to be $\mathbb{E}[Y] = 0.0542 < \langle Y \rangle$, and $p_f^{\text{best}} = 0.000472$ for the same $D_F^2 = 0.00292$.

Confidence in the system requires that the design margin be large in relation to the uncertainty, which underscores the importance of quantifying—and mitigating by means of optimal control—system uncertainties for purposes of certification. The application of COMOC to increase design confidence in the particular example of the robot-arm maneuver becomes obvious.

3.3.1.2 Uncertain Wind Forces and Uncertain Geometry

For this case, in addition to the uncertain lengths (retaining the same bounds), each body is affected by a random wind force in every time step, hitting the body's surface around the center of mass in a prescribed location. Each component of each three-dimensional force vector varies randomly between in $[-0.001, 0.001]$. With 15 time steps, two bodies, and three wind force components on each body, a total of $M = 92$ uncertain variables are introduced. Fig. 3.20 shows the mean performance evolution with $m = 100$, system diameter, and concentration-of-measure probability-of-failure upper bound along the simulated-annealing algorithm evolution. Even though only 50 objective function values have been computed, the probability-of-failure upper bound has been improved from $p_f = 1.0$ to $p_f^{\text{best}} = 0.1158$ and $D_F^2 = 0.01098$. Observe that the p_f^{best} has been found for a control sequence that leads to (local) minima in the mean and diameter, respectively. Again, assuming that the mean resulting from $m = 10000$ samples yields exactly the mean performance $\mathbb{E}[Y]$, then p_f^{best} is computed using Eq. (2.32) (with $M = (a - \mathbb{E}[Y])_+$), where $\mathbb{E}[Y] = 0.0545 > \langle Y \rangle$, with $p_f^{\text{best}} = 0.0567$ for the same $D_F^2 = 0.01098$. Here, even though $E[Y] > \langle Y \rangle$, p_f^{best} decreases because $\mathbb{E}[Y] \leq \langle Y \rangle + \alpha$ is satisfied (where $\alpha = 0.0225$) — thus M decreases when using

Eq. (2.32) instead of Eq. (2.61).

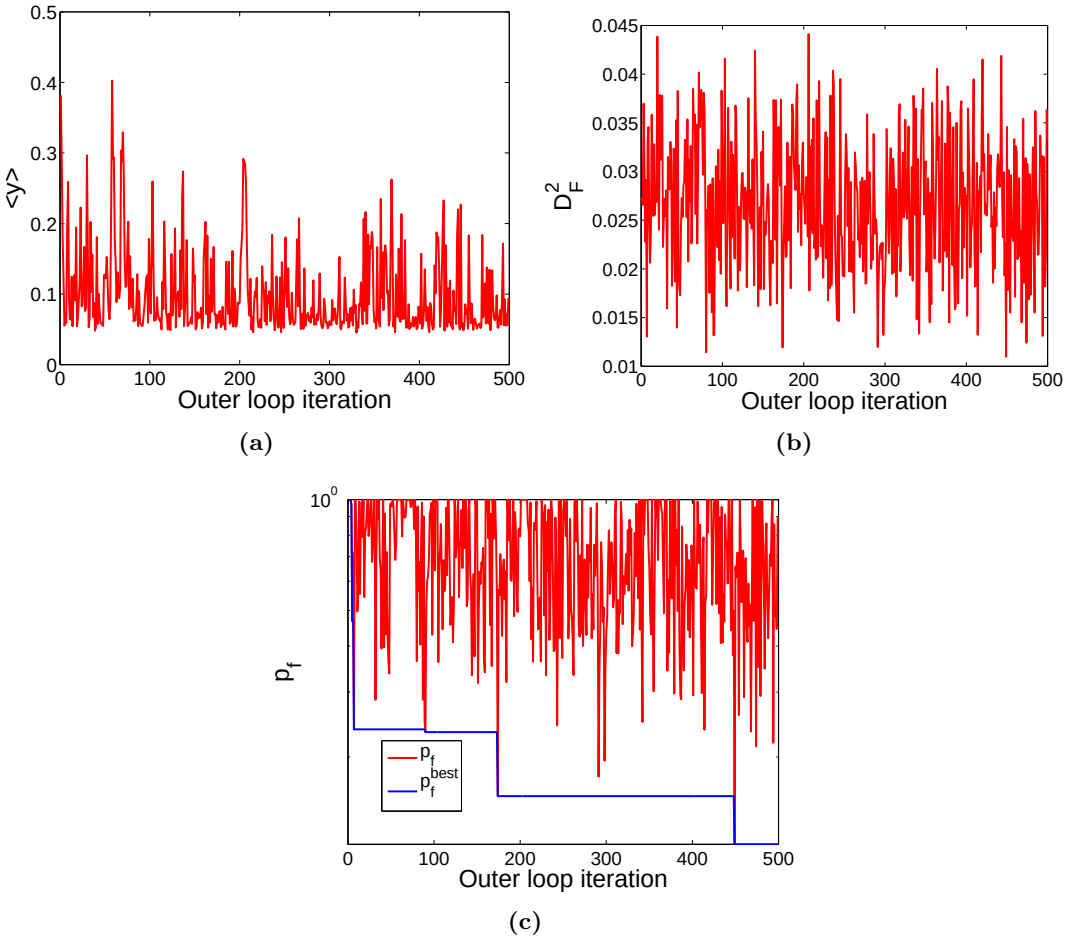


Figure 3.20: Uncertain wind forces and uncertain geometry: Simulated annealing algorithm for the determination of the optimal controls. Evolution of: (a) mean performance (b) system diameter and (c) concentration-of-measure probability-of-failure upper bound

3.3.2 Remarks

The greatest limitation of the COMOC implementation developed here is its computational cost. Each evaluation of the objective function (probability-of-failure) requires the calculation of the system diameter for a particular control, requiring multiple evaluations of the system's equations of motion. To reduce the computational expense to a tractable level, the controls have been constrained to remain close to the initial deterministic solution. The

MATLAB implementation prohibited use of parallelism in the optimizations, which would greatly decrease wallclock time of the entire COMOC simulation. It is likely that further gains in design confidence could be achieved from an unrestricted control optimization, but the computational resources and infrastructure required for such an optimization are beyond the scope of this first attempt. In view of these present limitations, the formulation of efficient COMOC implementations that alleviate its computational expense clearly suggests itself as a subject of further research.

3.4 PSAAP Hypervelocity Impact

3.4.1 Problem setting

The overarching PSAAP (Predictive Science Academic Alliance Program) application is the simulation of hypervelocity impact to be used as part of the uncertainty quantification campaign at the California Institute of Technology. Ballistic impact can be simulated using fully-Lagrangian finite-element methodology such as presently available in the Center's Virtual Testing Facility (VTF). An extensive Lagrangian methodology has been developed by Caltech's ASC/ASAP Center, including: composite finite elements specially designed for finite-deformation plasticity and ballistic impact dynamic conditions; cohesive elements for simulating brittle fracture and fragmentation; shear-band and spall elements; thermomechanical coupling; graph-based parallel fracture and fragmentation; variational integrators; variational r and h-adaption; variational contact and friction; and a suite of multiscale material models.

$F(X)$ represents the computational model that attempts to accurately represent the physical system, $G(X, Z)$. The system of interest for the PSAAP application is a hyperve-

locity impact between a 0.070 in-diameter 440c stainless steel sphere projectile and a 304 stainless steel 6 in-square plate 0.105 in thick. The finite-element package Abaqus is employed, which takes X as a random variable input and is completely deterministic otherwise. The computational size of the model is characterized by having $29,414$ linear tetrahedral elements and being symmetric about the $y = 0$ plane. The model supports a contact algorithm, a plasticity model, and element erosion as elements become overly distorted from the severe impact (in the absence of a fracture capability).

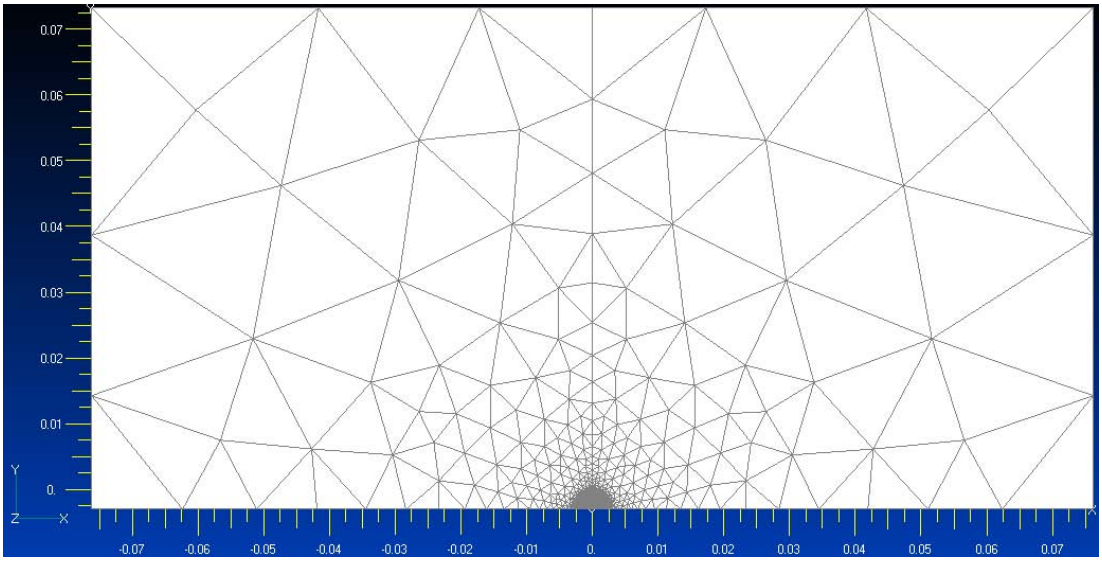
The performance measure for the system is the resulting perforation area in m^2 measured at a right angle from the plane of the plate. Experimentally, this measurement is made by shining a light down through the perforation and using software to determine the perimeter of the perforation and calculate the area inside of that perimeter. Refer to §B for details of the area measurement.

Computationally, the active elements of the final time step that are above a cutoff speed of 100 m/s are retained to eliminate disconnected material and ejecta from affecting perforation area and plotted in the plane of the plate in 2D. MATLAB image processing is used to calculate the ratio of black (solid) to white (perforated) area and then multiplied by the known area of the image to obtain total area of perforation. See Fig. 3.22 for a comparison for a single Abaqus model execution.

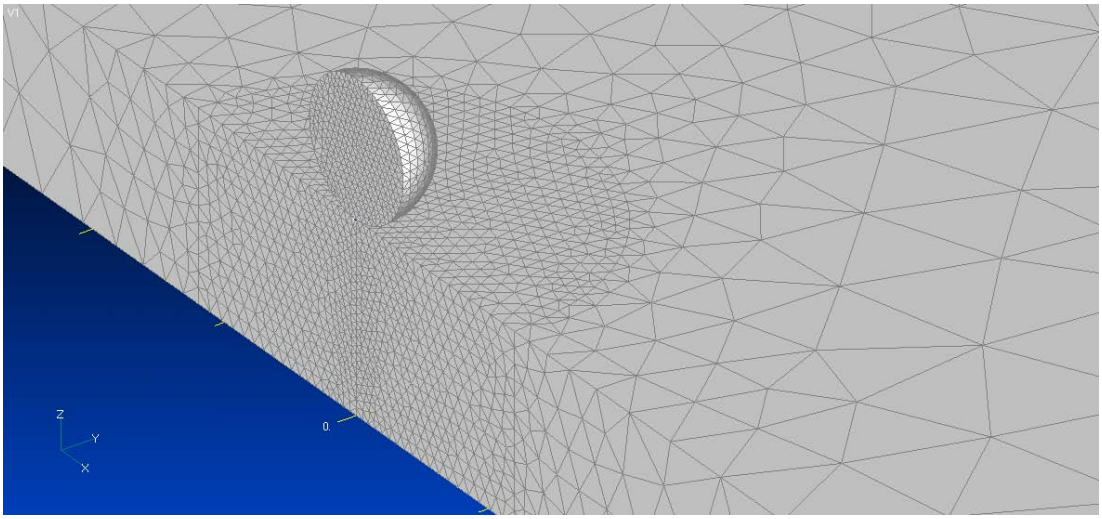
3.4.2 Verification

This application of the UQ framework uses six variables for the assessment of verification:

- Projectile Speed (V): [2147,2888] m/s
- Plate Obliquity (α_{plate}): [0,30] degrees in x-direction (symmetry about the 0° taken into account) — see Fig. 3.23 for the definition of the angle.



(a)



(b)

Figure 3.21: (a) Mesh used in PSAAP application. (b) Closeup near impact region

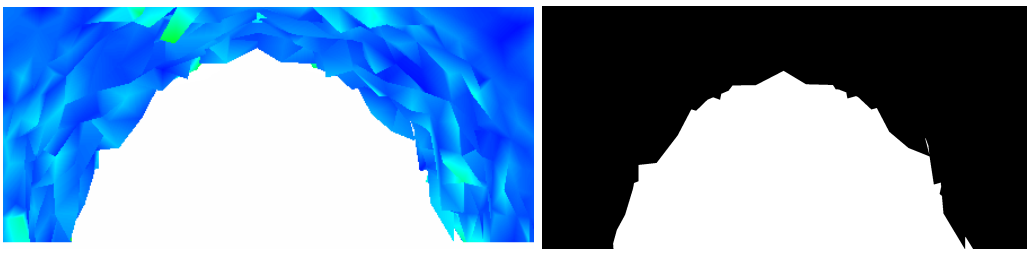


Figure 3.22: (a) Abaqus and (b) MATLAB visualizations of PSAAP performance measure

- Plate yield stress ($\sigma_{y_{\text{plate}}}$): [3.0e8,5.0e8] Pa
- Mesh size (Δx): Discretized by three mesh sizes of $\{14087, 29414, 60678\}$ elements
- Plate thickness (h_{plate}): Discretized by four plate thicknesses of $\{60, 78, 90, 105\}$ thousandths of an inch
- Time step (Δt): [5e-10,1e-9] s

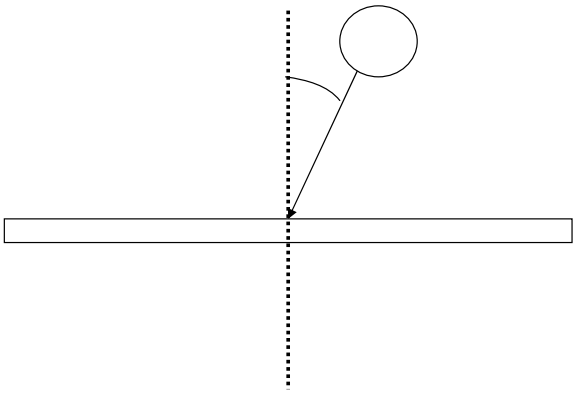


Figure 3.23: Side-view schematic of how obliquity is defined in PSAAP applications. Zero degrees corresponds to a perfectly perpendicular impact of the projectile to the plate

The study is characterized by doing diameter calculations for each variable with 20 function evaluations using DAKOTA's single objective function genetic algorithm from the JEGA package. The diameter is given in terms of m^4. The mean was estimated using 40 Monte Carlo samples from DAKOTA's nondeterministic sampling package.

Both the diameter and mean calculations were computed using two processors for each execution of Abaqus (v6.7), and six concurrent executions on Caltech CACR's *mind-meld* machine. The time for each execution varied, based heavily on the time step and mesh size variables. The median execution time was about 15 minutes.

Machine specifications for Caltech CACR's *mind-meld*:

- **Architecture**: Opteron Linux server

- **Memory**: 32 GB DDR2-667 main memory, 16 x 2 GB
- **Processors**: 8 dual core AMD Opteron 8220, 2.8 GHz, 1 MB Level 2 cache per core
- **Network Interconnect**: PCI-x, PCI-Express, Gigabit Ethernet
- **Disk**: nfs project work area, SAS home directories
- **Operating System**: Red Hat Linux 2.6.18-53.1.19

3.4.2.1 Results

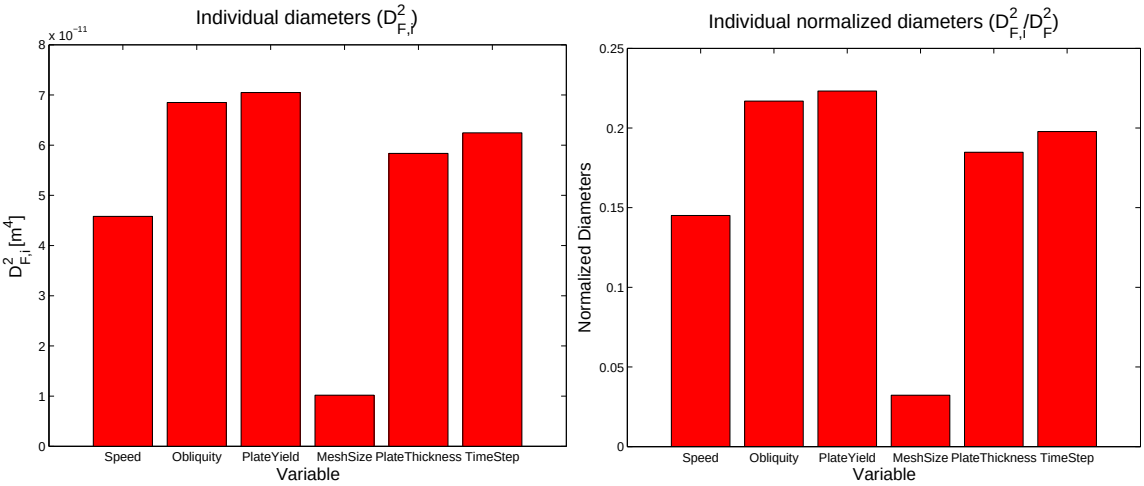


Figure 3.24: PSAAP verification exercise showing (a) individual subdiameters and (b) normalized subdiameters (D_i^2/D^2)

3.4.2.2 Discussion

Fig. 3.24(a) illustrates that the dominant contributors to uncertainty from the given domain of random variables reside in the plate yield and time step parameters within the given ranges, while Fig. 3.24(b) shows their relative contributions to the total diameter. These two figures open a door to many conclusions. First, the physical parameter oscillations with

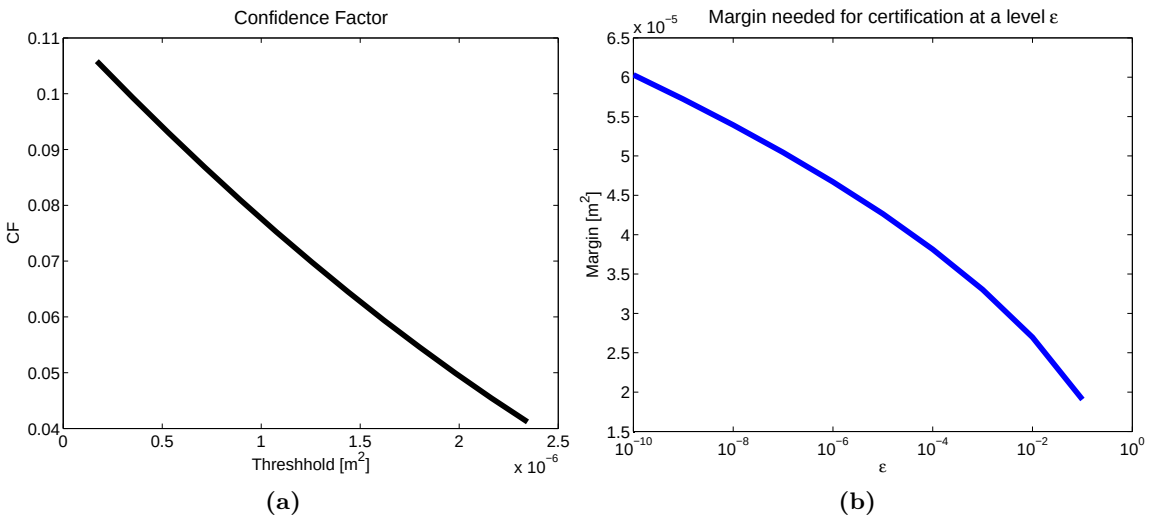


Figure 3.25: PSAAP verification exercise showing (a) confidence factors and (b) margins required for certification

regards to the given input domain are ostensibly larger collectively than those associated with numerical parameters. This suggests that most of the uncertainty in this model is attributed to the real-world parameters that are the only ones of interest to those performing experiments. If the majority of uncertainty were associated to the numerical parameters, then the model's performance is overly sensitive to numerical inadequacies. This suggests that the model cannot be *verified* in this framework because it is nonconvergent with respect to numerical parameters. This interpretation of verification is not dissimilar to that given in §1.2.2, since this framework identifies that the numerical parameters/algorithms do not accurately represent the modeler's conceptual description of a model by large uncertainties of performance. The latter years of this project aim to include more complex physics (fracture mechanics, temperature effects, rate sensitivity effects, first-principles parameters, void nucleation, etc.) that are not included in this model. This would suggest a more accurate model, but again, this framework will identify if that model can be verified numerically.

Fig. 3.25(b) shows the margin required for certification ($CF = M/U > 1$) at different

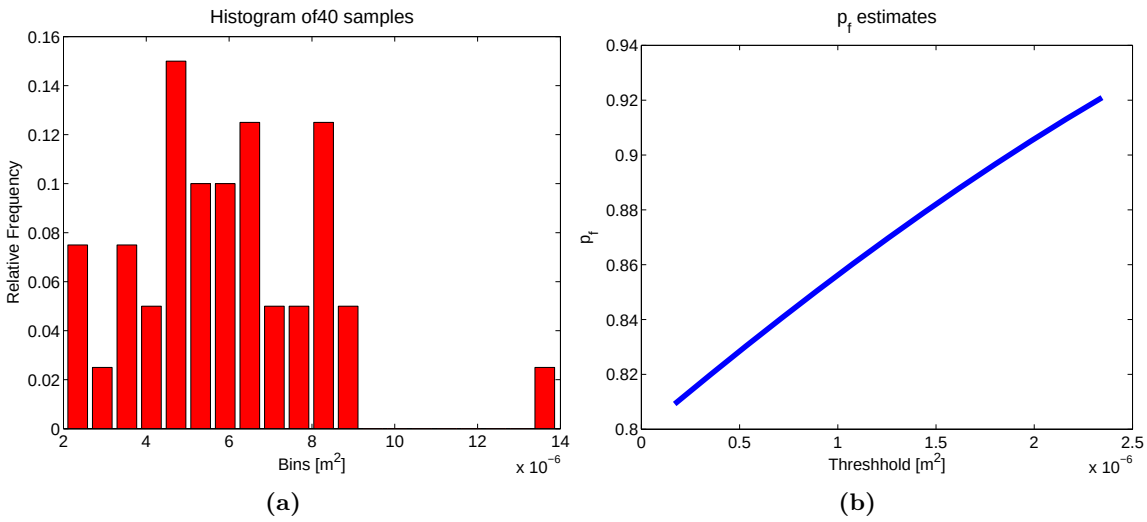


Figure 3.26: PSAAP verification exercise showing (a) a histogram of Monte Carlo samples used to estimate the mean and (b) the probability-of-failure computed by CoM UQ

levels of confidence ϵ. Fig. 3.25(a) shows a very low certification factor for even negative thresholds — that levels off at 0 for positive threshold values. These figures support the same interpretation — uncertainties dominate over the margin: $M/U < 1$ which is not certifiable for any respectable ϵ. It is the hope that the more intelligent physics codes will help reduce uncertainty in the model, as mentioned in the above paragraph. Note that the physical parameter specifications may be too broad to certify the system even with a perfect model, and only a shrinking of these bounds could provide the means to verification.

Fig. 3.26(a) shows that, of the 100 Monte Carlo samples taken to estimate the mean, the most observed perforation areas reside nearer to $0\,\mathrm{m}^2$. Fig. 3.26(b) shows the probability-of-failure computed from McDiarmid's concentration-of-measure inequality. In this case, trivial p_fs are recovered from meaningful margins, while nontrivial upper bounds of p_f are only obtained for the unobtainable negative perforation area threshold values.

3.4.3 Validation

Certain limitations of the associated experiments prohibit using velocity as a strictly controllable bounded random variable. Refer to §4.1.2 for a detailed discussion of how uncontrollable random variables are treated in this framework, a slight deviation from the validation campaign described in §2.2.5. The two-stage light gas gun used by Rosakis et al. to recover $G(X, Z)$ does not have repeatability or exact control over X. Some reasons for this are:

- Imprecise amount of hydrogen present in pump tube
- Imprecise amount of atmosphere present in the flight tube and target tank
- Natural variations and defects in material and material processes
- Magnitude of friction between bore and launch package
- Weight and fit of piston in the flight tube
- Head space of the AR section
- Fit of the mylar burst disc
- Potential blowby.

Since this application is characterized by including one controllable variable, $X_N = V$ among the controllable physical variables $(X_1, X_2) = (h_{plate}, \alpha_{plate})$, the deviation from the validation campaign in §2.2.5 uses some theory developed in §4 for exactly this case:

1. Compute $\bar{D}_{G-F,N}$ using Eq. (4.24).
2. For $i < N$, bound $\bar{D}_{G-F,i}$ using Eq. (4.27).

3. Bound $\bar{D}_G$ using $\bar{D}_G \leq \bar{D}_F + \bar{D}_{F-G}$.

4. Conclude via concentration inequalities Eq. (4.18) and Eq. (4.19).

3.4.3.1 Results

To date, only the subdiameter associated to the plate thickness has been recovered, since the experiments take nontrivial amounts of man-hours to setup, complete, and catalogue. After this thesis has been submitted, the last subdiameters should have been attempted and computed to complete the uncertainty quantification of this system — *the first UQ campaign of its kind in the world.* With the verification diameter suggesting noncertification, the validation diameter seems unlikely to be small for the given design domain, especially with the noisy $G(X, Z)$ function that is expected due to the range of competing complex phenomena of hypervelocity metallic impact — excavation, spall, vaporization, flange petalling, radial and circumferential crack nucleation and propagation, and shear plugging. Tractability of this method is shown regardless of the outcome.

The optimizer described in §A.2 was used here because of the noisy objective function. Tables 3.3–3.6 show the data by steps in the optimization. The columns of "Thickness", "Obliquity", and "Velocity" are given for each shot/simulation done. The experimental perforation area measurement falls under the "G result" column, and the simulation perforation area falls under the "F result" column. Recall that a triplet of values suggested by the optimizer translates into two pairs of plate thickness and obliquity settings — this distinction is made in the "Pairs" column where a triplet shares the same obliquity value between the two shots/simulations with the same pair number. The "Xprime?" column denotes whether the shot/simulation was associated to the value of the copied, or "primed", random variable set (1) or not (0). The "Pair" and "Xprime?" distinctions are necessary

to build the trajectory of triplets (h, α, h') suggested by the optimizer in objective function space in Fig. 3.27.

Table 3.3: PSAAP Validation Initialization Phase Data

Thickness	Obliquity	Velocity	F Result	G Result	Pair	Xprime?
105	0	2480	6.91591275e-06	5.35e-06	4	0
105	10	2719.28	8.32887225e-06	5.75e-06	1	0
105	21	2266.79	2.7426915e-06	0	2	0
105	30	2500	3.42846e-06	0	3	0
90	0	2402.9	6.4514025e-06	5.75e-06	4	1
90	10	1811.36	3.612645e-07	5.61e-06	1	1
90	21	2800	7.95743775e-06	6.77e-06	2	1
90	30	2825.75	1.413011475e-05	5.91e-06	3	1

Table 3.4: PSAAP Validation Iteration 1 Data

Thickness	Obliquity	Velocity	F Result	G Result	Pair	Xprime?
90	15	2509	6.45943725e-06	5.735e-06	6	1
60	15	2567	7.38419175e-06	8.859e-06	5	1
73	11	2635	8.1427455e-06	8.943e-06	3	1
73	15	2507	6.7249935e-06	8.735e-06	1	0
73	15	2631	7.79283225e-06	9.187e-06	1	1
73	15	2701	8.09181675e-06	8.233e-06	2	1
73	18	2260	4.872069e-06	6.752e-06	4	1
90	11	2320	4.93202925e-06	9.094e-06	3	0
90	15	2315	5.277996e-06	5.352e-06	5	0
90	15	2316	5.34094425e-06	6.813e-06	6	0
90	15	2741	8.1811485e-06	6.887e-06	2	0
90	18	2815	8.2037295e-06	8.357e-06	4	0

Table 3.5: PSAAP Validation Iteration 2 Data

Thickness	Obliquity	Velocity	F Result	G Result	Pair	Xprime?
105	2	2468	6.60886425e-06	3.76e-06	2	0
105	5	2255	4.4346285e-06	0	4	0
105	5	2648	7.73988075e-06	3.04e-06	5	0
105	5	2714	8.64536175e-06	3.24e-06	6	0
105	9	2201	4.103055e-06	0	1	0
73	5	2278	5.55292125e-06	7.4e-06	4	1
73	5	2438	6.5976165e-06	7.68e-06	5	1
90	2	2839	9.49428e-06	9.56e-06	2	1
90	5	2437	6.10914825e-06	5.42e-06	3	0
90	5	2556	7.76476125e-06	7.26e-06	3	1
90	5	2676	8.21572425e-06	8.54e-06	6	1
90	9	2157	3.64620825e-06	4.49e-06	1	1

Table 3.6: PSAAP Validation Iteration 3 Data

Thickness	Obliquity	Velocity	F Result	G Result	Pair	Xprime?
105	0	2146	3.42001575e-06	0	1	0
105	0	2390	5.57693775e-06	4.45e-06	4	0
105	0	2504	6.35877e-06	5.35e-06	5	0
105	0	2888	1.00148895e-05	7.39e-06	5	1
105	4	2344	4.88962575e-06	4.47e-06	3	0
73	0	2445	6.97044375e-06	7.78e-06	2	1
73	0	2782	8.96442975e-06	1.014e-05	4	1
90	0	2445	6.93045225e-06	5.68e-06	1	1
90	0	2767	9.17441775e-06	8.84e-06	2	0
90	4	2408	5.8591935e-06	6.6e-06	3	1

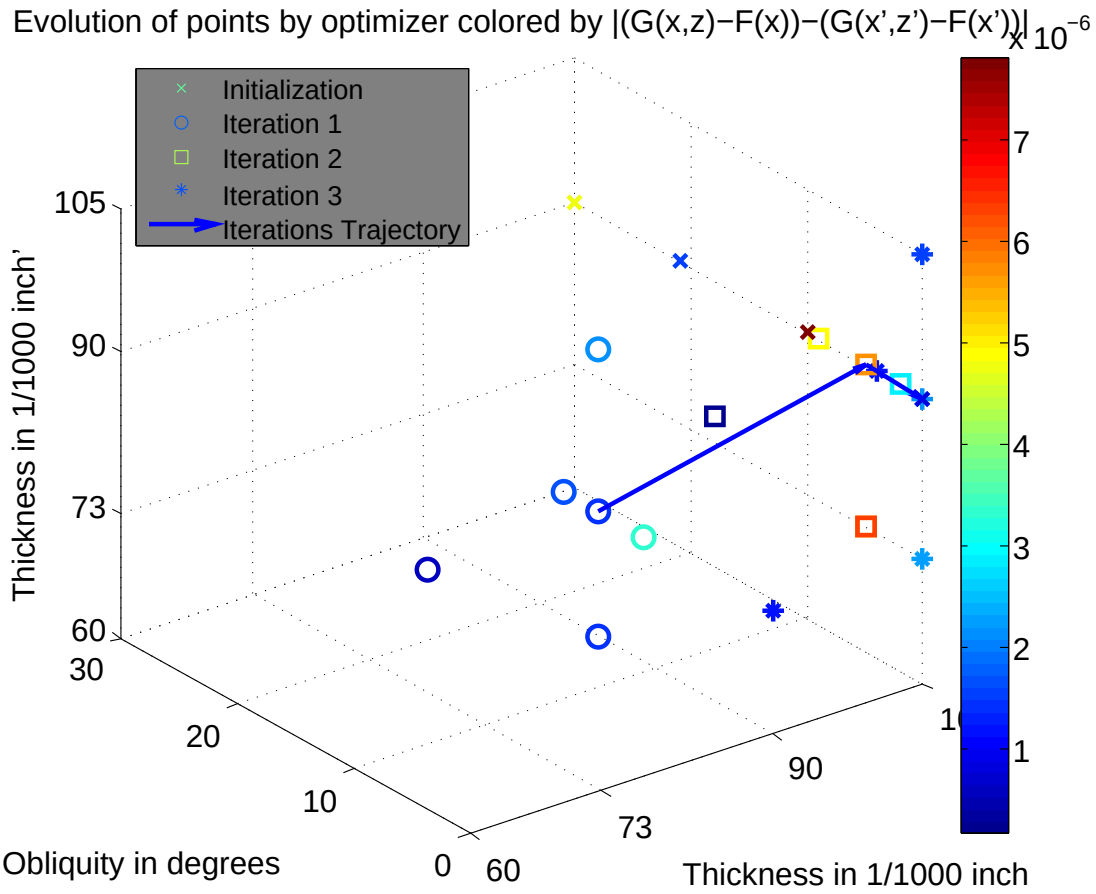


Figure 3.27: Visualization of the iterations for the PSAAP application using color to denote the objective function value as in Eq. (4.27) for the plate thickness subdiameter

3.4.3.2 Discussion

The validation diameter is the centerpiece of the UQ framework that has been developed in this work. It rigorously defines a tractable methodology to assess the uncertainty between a system and the model meant to predict its performance. It is not expected that the first attempt at computing a validation diameter suggests a certifiable system, as the system may inherently not be as uncertain as the validation diameter would suggest, since it is conservative. Between the noisy experimental results and a model that suggests a relatively large verification diameter, the expectation is that F and G will differ for every realization

suggest by the optimizer. The results above show that although the numerical and experimental results are within "ballpark" range of each other, they differ enough to estimate large subdiameters. The design point that optimizes the objective function is $(h, \alpha, h') = (105\ 1/1000\,\text{in}, 0°, 90\ 1/1000\,\text{in})$. With more objective function evaluations and unlimited averaging over (V, Z) at each triplet, it is assumed that some other point minimizes the objective function. The convergence criterion for the optimization algorithm was met at this point, and now only requires averaging over (V, Z) to confirm the objective function value at that point. This step will need to be done for each subdiameter computation because of the noisy objective function.

Fig. 3.27 illustrates the trajectory that the optimizer selected in the (h, α, h') space. The color of the points denote the value of the objective function of the optimization as in Eq. (4.27) for $X = (h, \alpha)$ and $x_1' = h'$. Note that the first iteration shows a cloud of points suggested by the optimizer that are relatively small in objective function value compared to the clouds of points associated to later iterations, suggesting that the optimizer is converging to a global supremum of the noisy objective function.

3.4.4 General discussion

The PSAAP application is the first of its kind to employ the suite of systematic verification and validation techniques to provide a measure of certification. Although this system shows that it cannot be certified for the given ranges, the verification subdiameters suggest where numerical inadequacies could reduce uncertainty, and the validation subdiameter computed thus far for plate thickness highlights the discrepancies between model and experiment, and also suggests improvements to be made for the next stages of the program for the physics modelers and the control of experiments and data measurement.

Chapter 4

Extensions to CoM UQ

4.1 UQ in a general setting

Often times, the restrictions of independence between random variables may be too strict for McDiarmid's inequality, Eq. (2.1), if there is correlation among them or independence cannot be verified. There must be a recourse to a concentration-of-measure inequality that relaxes this restriction. For this, the concentration-of-measure inequalities below exemplify the flexibility of the theory to adapt to these more general cases.

Let G be an arbitrary function of two random variables X, Z where X may be vector valued or infinite dimensional random variables and Z is an unknown unknown set of random variables. The goal is to obtain the sharpest possible bound for

$$\mathbb{P}[G(X, Z) \geq r] \tag{4.1}$$

with very limited information on the probability distribution of X and Z. The following theorems account for limited information in various inequalities.

4.1.1 Limited-information inequalities

4.1.1.1 Only $\mathbb{E}[G]$ and $\inf G$ are known

Consider the case that only $\mathbb{E}[G]$ and $\inf G$ are known and no further assumptions are made on X and Z; then the sharpest possible estimate is given by the following theorem

Theorem 4.1.1. *For all* $r > \inf_{x,z} G(x,z)$

$$\mathbb{P}[G(X,Z) \geq r] \leq \min\left(1, \frac{\mathbb{E}[G] - \inf_{x,z} G(x,z)}{r - \inf_{x,z} G(x,z)}\right). \tag{4.2}$$

The upper bound is optimal in the following sense: for all $r > \inf_{x,z} G(x,z)$ *and all* $\gamma > \inf_{x,z} G(x,z)$ *if there exists* (x_r, z_r) *such that* $G(x_r, z_r) = r$*, then with* μ *as an arbitrary probability distribution on* (X,Z)

$$\sup_{\mu\,:\,\mathbb{E}_\mu[G]=\gamma} \mathbb{P}_\mu[G(X,Z) \geq r] = \min\left(1, \frac{\gamma - \inf_{x,z} G(x,z)}{r - \inf_{x,z} G(x,z)}\right). \tag{4.3}$$

4.1.1.2 Inequality on distribution of noise deviation

Consider a function that is inherently noisy from epistemic uncertainties, Z. It may be convenient to frame an inequality in terms of the function's average over the epistemic uncertainty. The following theorem is an adaptation of a theorem found in [44]. Assume that $X := (X_1, \ldots, X_N)$ where the X_i are independent random variables (not necessarily identically distributed). Write $G_Z(x)$ the average of the variable $G(x,Z)$ with respect to Z. Take $\rho := G(X,Z) - G_Z(X)$ then $\mathbb{E}_\rho$ is the expectation with respect to that variable, where ρ can be considered a distribution of the epistemic noise about its mean.

Theorem 4.1.2.

$$\mathbb{P}[G(X,Z) \geq r] \leq \mathbb{E}_\rho \left[\exp\left(-2 \frac{(r - \mathbb{E}[G_Z(X)|\rho] - \rho)_+^2}{D_{G_Z}^2} \right) \right] \tag{4.4}$$

and

$$\mathbb{P}[G(X,Z) \leq r] \leq \mathbb{E}_\rho \left[\exp\left(-2 \frac{(\mathbb{E}[G_Z(X)|\rho] + \rho - r)_+^2}{D_{G_Z}^2} \right) \right] \tag{4.5}$$

where

$$D_{G_Z}^2 := \sum_{i=1}^{N} \sup_{x_1,\ldots,x_n,x_i'} |G_Z(\ldots,x_i,\ldots) - G_Z(\ldots,x_i',\ldots)|^2. \tag{4.6}$$

4.1.1.3 Inequalities on distribution of noise deviation extrema

Now, instead of working with the distribution of noise about its mean, consider the noise extrema. With

$$\rho_{\min} := \inf G(X,Z) - G_Z(X) \tag{4.7}$$

and

$$\rho_{\max} := \sup G(X,Z) - G_Z(X) \tag{4.8}$$

Theorems 4.1.3 and 4.1.4 will be useful.

Theorem 4.1.3.

$$\mathbb{P}[G(X,Z) \geq r] \leq \exp\left(-2 \frac{(r - \mathbb{E}[G] - \rho_{\max})_+^2}{D_{G_Z}^2} \right) \tag{4.9}$$

and

$$\mathbb{P}[G(X,Z) \leq r] \leq \exp\left(-2 \frac{(\mathbb{E}[G] + \rho_{\min} - r)_+^2}{D_{G_Z}^2} \right) \tag{4.10}$$

where D_{G_Z} is defined by Eq. (4.6).

Theorem 4.1.4. *If the entries of X are independent and X and Z are independent, then*

$$\mathbb{P}[G(X,Z) \geq r] \leq \exp\left(-2\frac{(r-\mathbb{E}[G])_+^2}{D_Z^2 + D_{G_Z}^2}\right) \tag{4.11}$$

and

$$\mathbb{P}[G(X,Z) \leq r] \leq \exp\left(-2\frac{(\mathbb{E}[G]-r)_+^2}{D_Z^2 + D_{G_Z}^2}\right) \tag{4.12}$$

where $D_Z := \rho_{\max} - \rho_{min}$ and D_{G_Z} is defined by Eq. (4.6).

4.1.1.4 Numerical model

It follows from Theorem 4.1.3 and 4.1.4 that one only needs to estimate $\mathbb{E}[G]$, D_Z, and D_{G_Z}. The estimation of D_{G_Z} can be simplified using a numerical model $F(X)$ and the following proposition

Proposition 4.1.5.

$$D_{G_Z} \leq D_F + D_{F-G_Z} \tag{4.13}$$

Let $Y^1, \ldots, Y^q$ be q independent copies of $G(X,Z)$. Then

$$< G >_q := \frac{\sum_{i=1}^q Y_i}{q}. \tag{4.14}$$

The following theorem is then a straightforward consequence of the results of the previous sections

Theorem 4.1.6. *Write*

$$\alpha_{\epsilon'} := \frac{D_Z^2 + (D_F + D_{G_Z - F})^2}{\sqrt{2q}} \sqrt{\ln \frac{1}{\epsilon'}} \tag{4.15}$$

where $D_Z := \sup(G(X,Z) - G_Z(X)) - \inf(G(X,Z) - G_Z(X))$

- *With probability* $1-\epsilon'$ *on the* q *samples*

$$\mathbb{P}[G(X,Z) \geq r] \leq \exp\left(-2\frac{(r- < G >_q - \alpha_{\epsilon'})_+^2}{D_Z^2 + (D_F + D_{G_Z - F})^2}\right) \tag{4.16}$$

or similarly,

$$\mathbb{P}[G(X,Z) \leq r] \leq \exp\left(-2\frac{(< G >_q - r - \alpha_{\epsilon'})_+^2}{D_Z^2 + (D_F + D_{G_Z - F})^2}\right). \tag{4.17}$$

4.1.2 Martingale inequality

Using Theorem 4.1.9, with no assumptions whatsoever made on $X_1, \ldots, X_N, Z$,

$$\mathbb{P}[G(X,Z) \geq r] \leq \exp\left(-2\frac{(r - \mathbb{E}[G])_+^2}{\bar{D}_G^2}\right) \tag{4.18}$$

and

$$\mathbb{P}[G(X,Z) \leq r] \leq \exp\left(-2\frac{(\mathbb{E}[G] - r)_+^2}{\bar{D}_G^2}\right) \tag{4.19}$$

where

$$\bar{D}_G^2 = \sum_{i=1}^{N} \bar{D}_{G,i}^2 \tag{4.20}$$

$$\begin{aligned} \bar{D}_{G,N} = \sup_{z, x_N, \ldots, x_1} \Big| & \mathbb{E}\left[G(X,Z) \middle| (Z, X_N, \ldots, X_1) = (z, x_N, \ldots, x_1)\right] \\ & - \mathbb{E}\left[G(X,Z) \middle| (X_{N-1}, \ldots, X_1) = (x_{N-1}, \ldots, x_1)\right] \Big| \end{aligned} \tag{4.21}$$

and for $i < N$

$$\begin{aligned}\bar{D}_{G,i} = \sup_{x_i,\ldots,x_1} \Big| & \mathbb{E}\left[G(X,Z)\big|(X_i,\ldots,X_1) = (x_i,\ldots,x_1)\right] \\ & - \mathbb{E}\left[G(X,Z)\big|(X_{i-1},\ldots,X_1) = (x_{i-1},\ldots,x_1)\right] \Big|\end{aligned} \tag{4.22}$$

where $\mathbb{E}[X|Y]$ denotes the conditional expectation of X given Y. Recall that if $X_1,\ldots,X_{N-1},(X_N,Z)$ are independent then

$$\begin{aligned}\bar{D}_{G,i} = \sup_{x_i,\ldots,x_1} \Big| & G_{Z,X_N,\ldots,X_{i+1}}(x_i,\ldots,x_1) \\ & - G_{Z,X_N,\ldots,X_i}(x_{i-1},\ldots,x_1)\Big|\end{aligned} \tag{4.23}$$

where $G_{Z,X_N,\ldots,X_i}$ stands for the function G averaged with respect to $Z,X_N,\ldots,X_i$.

$$\begin{aligned}\bar{D}_{G,N} = \sup_{z,x_N,\ldots,x_1} \Big| & G(z,x_N,\ldots,x_1) \\ & - \mathbb{E}\left[G_{X_N,Z}(x_{N-1},\ldots,x_1)\right]\Big|.\end{aligned} \tag{4.24}$$

Observe also that if the probability distributions of the random variables $X_1,\ldots,X_{N-1}$ are unknown then one can use (for $i < N$) the more conservative bound

$$\begin{aligned}\bar{D}_{G,i} \leq \sup_{x_{N-1},\ldots,x_i,x_i',\ldots,x_1} \Big| & G_{Z,X_N}(x_{N-1},\ldots,x_i,\ldots,x_1) \\ & - G_{Z,X_N}(x_{N-1},\ldots,x_i',\ldots,x_1)\Big|.\end{aligned} \tag{4.25}$$

To reduce computational cost, instead of the averaged expression in Eq. (4.25), use the noisy version by replacing $G_{Z,X_N}(x_{N-1},\ldots,x_i,\ldots,x_1)$ with $G(z,x_N,x_{N-1},\ldots,x_i,\ldots,x_1)$ and $G_{Z,X_N}(x_{N-1},\ldots,x_i',\ldots,x_1)$ with $G(z',x_N',x_{N-1},\ldots,x_i,\ldots,x_1)$ where z,z' are epistemic uncertainties. Generally, for a smoothed $G_Z(X)$, $G(X,Z) = G_Z(X) + \omega(Z)$ where

$\omega(Z)$ is considered noise. Then Eq. (4.25) becomes

$$\bar{D}_{G,i} \leq \sup_{x_1,\ldots,x_i,x_i',\ldots,x_N,z,z'} \left| G(x_1,\ldots,x_i,\ldots,x_N,z) - G(x_1,\ldots,x_i',\ldots,x_N,z') \right| \tag{4.26}$$

Now consider the appropriate expression for validation

$$\begin{aligned} \bar{D}_{G-F,i} \leq \sup_{x_1,\ldots,x_i,x_i',\ldots,x_N,z,z'} \Big| & G(x_1,\ldots,x_i,\ldots,x_N,z) - F(x_1,\ldots,x_i,\ldots,x_N) \\ & - G(x_1,\ldots,x_i',\ldots,x_N,z') + F(x_1,\ldots,x_i',\ldots,x_N) \Big| \end{aligned} \tag{4.27}$$

which that inequality used for the validation campaign, as in §3.4.3.

4.1.3 Proofs of general UQ

4.1.3.1 Proof of Theorem 4.1.1

The proof of Eq. (4.2) follows from the fact that on the subset $\Big\{(X,Z) \,:\, G(X,Z) \geq r\Big\}$ one has

$$\frac{G(X,Z) - \inf_{x,z} G(x,z)}{r - \inf_{x,z} G(x,z)} \geq 1. \tag{4.28}$$

The proof of Eq. (4.3) follows by choosing a measure μ such that (assuming $r > \gamma$)

$$\mu[\{(x_r,z_r)\}] = \frac{\gamma - \inf_{x,z} G(x,z)}{r - \inf_{x,z} G(x,z)} \tag{4.29}$$

with a remaining mass near a minimizer of G.

4.1.3.2 Centering sequences with bounded differences

The following lemma is taken from [23], which is also Lemma 3.2 of [44].

Lemma 4.1.7. *Let X be a real-valued random variable with $\mathbb{E}[X] = \mu$ with $a \leq X \leq b$ where a and b are constants. Then for any h*

$$\begin{aligned}\mathbb{E}[e^{hX}] &\leq \frac{b-\mu}{b-a}e^{ha} + \frac{\mu-b}{b-a}e^{hb} \\ &\leq e^{h\mu}\exp\left(\frac{1}{8}h^2(b-a)^2\right).\end{aligned} \tag{4.30}$$

Given a sequence $X = (X_1, X_2, \ldots)$ of integrable random variables, the corresponding difference sequence is $Y = (Y_1, Y_2, \ldots)$, where $Y_k = X_k - X_{k-1}$ using the convention $X_0 := 0$ and $Y_1 := X_1$. Let $\mu_k(x) := \mathbb{E}[Y_k | X_{k-1} = x]$. Define the distribution of the sequence X to be centering if for each $k = 2, 3, \ldots$, $\mu_k(x)$ is a nonincreasing function of x.

The following theorem is Theorem 2.3 of [44].

Theorem 4.1.8. *Let $X_1, X_2, \ldots, X_n$ be a centering sequence with corresponding differences $Y_k = X_k - X_{k-1}$, and suppose that there are constants a_k, b_k such that $a_k \leq Y_k \leq b_k$ for each k. Then for $t > 0$*

$$\mathbb{P}[X_n - \mathbb{E}[X_n] \geq t] \leq \exp\left(-2\frac{t^2}{\sum_{k=1}^n (b_k - a_k)^2}\right) \tag{4.31}$$

and

$$\mathbb{P}[X_n - \mathbb{E}[X_n] \leq -t] \leq \exp\left(-2\frac{t^2}{\sum_{k=1}^n (b_k - a_k)^2}\right). \tag{4.32}$$

The proof of the following theorem is similar to the proof of Theorem 6.7 and Corollary 6.9 of [43].

Theorem 4.1.9. *Let $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_n$ be a filter. Let the integrable random variable X be $\mathcal{F}_n$ measurable, and let $X_0, X_1, \ldots, X_n$ be the martingale obtained by setting $X_k := \mathbb{E}[X|\mathcal{F}_k]$. Suppose that for each $k = 1, \ldots, n$ there are constants a_k, b_k, such that $a_k \leq$*

$X_k - X_{k-1} \leq b_k$. *Then for any* $t > 0$, *a.s.*

$$\mathbb{P}\Big[X_n - \mathbb{E}[X_n|\mathcal{F}_0] \geq t\Big|\mathcal{F}_0\Big] \leq \exp\left(-2\frac{t^2}{\sum_{k=1}^n (b_k - a_k)^2}\right) \tag{4.33}$$

and

$$\mathbb{P}\Big[X_n - \mathbb{E}[X_n|\mathcal{F}_0] \leq -t\Big|\mathcal{F}_0\Big] \leq \exp\left(-2\frac{t^2}{\sum_{k=1}^n (b_k - a_k)^2}\right). \tag{4.34}$$

Proof. Observe that for $h > 0$

$$\mathbb{P}\Big[X_n - \mathbb{E}[X_n|\mathcal{F}_0] \geq t\Big|\mathcal{F}_0\Big] \leq e^{-ht}\mathbb{E}\Big[\exp(h(X_n - \mathbb{E}[X_n|\mathcal{F}_0]))\Big|\mathcal{F}_0\Big]. \tag{4.35}$$

Since $(\mathcal{F}_k)$ is a filter,

$$\mathbb{E}\Big[\exp(h(X_n - \mathbb{E}[X_n|\mathcal{F}_0]))\Big|\mathcal{F}_0\Big] = \mathbb{E}\Big[\exp(h(X_{n-1} - \mathbb{E}[X_n|\mathcal{F}_0]))\mathbb{E}[e^{h(X_n - X_{n-1})}|\mathcal{F}_{n-1}]\Big|\mathcal{F}_0\Big]. \tag{4.36}$$

Using Lemma 4.1.7, a.s.

$$\mathbb{E}[e^{h(X_n - X_{n-1})}|\mathcal{F}_{n-1}] \leq \exp\left(\frac{1}{8}h^2(b_n - a_n)^2\right) \tag{4.37}$$

It follows that

$$\mathbb{E}\Big[\exp(h(X_n - \mathbb{E}[X_n|\mathcal{F}_0]))\Big|\mathcal{F}_0\Big] \leq \exp\left(\frac{1}{8}h^2(b_n - a_n)^2\right)\mathbb{E}\Big[\exp(h(X_{n-1} - \mathbb{E}[X_n|\mathcal{F}_0]))\Big|\mathcal{F}_0\Big]. \tag{4.38}$$

By induction, it is deduced that

$$\mathbb{P}\Big[X_n - \mathbb{E}[X_n|\mathcal{F}_0] \geq t\Big|\mathcal{F}_0\Big] \leq e^{-ht}\exp\left(\frac{1}{8}h^2\sum_{k=1}^n (b_k - a_k)^2\right) \tag{4.39}$$

and conclude by taking $h = 4t/\sum_{k=1}^{n}(b_k - a_k)^2$. □

The following theorem is an adaptation of Lemma 3.16 of [45]

Lemma 4.1.10. *Let $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_n$ be a filter. the integrable random variable X be $\mathcal{F}_n$ measurable, and let $X_0, X_1, \ldots, X_n$ be the martingale obtained by setting $X_k := \mathbb{E}[X|\mathcal{F}_k]$. Write $Y_k := X_k - X_{k-1}$ the corresponding martingale difference, assume that each Y_k is bounded from above. Let Z be an arbitrary positive random variable. Then for any h*

$$\mathbb{E}\big[Ze^{h\sum_k Y_k}\big|\mathcal{F}_0\big] \leq \operatorname{esssup}\mathbb{E}\Big[Z\prod_{k=1}^{n}\mathbb{E}\big[e^{hY_k}\big|\mathcal{F}_{k-1}\big]\Big]. \tag{4.40}$$

4.1.3.3 Proof of Theorem 4.1.2

Observe that

$$\mathbb{P}[G(X,Z) \geq r] = \mathbb{P}\Big[G_Z(X) \geq r - \rho\Big]. \tag{4.41}$$

Hence

$$\mathbb{P}[G(X,Z) \geq r] = \mathbb{E}_\rho\Big[\mathbb{P}\big[G_Z(X) - \mathbb{E}[G_Z(X)|\rho] \geq r - \rho - \mathbb{E}[G_Z(X)|\rho]\big|\rho\big]\Big]. \tag{4.42}$$

Let $\mathcal{F}_n$ be the σ-algebra generated by $\rho, X_1, \ldots, X_n$. Write

$$Y_k := \mathbb{E}\Big[G_Z(X)\Big|\mathcal{F}_k\Big]. \tag{4.43}$$

Observe that $(\mathcal{F}_n)$ is a filtration (i.e., $\mathcal{F}_k \subset \mathcal{F}_{k+1}$). Using Theorem 4.1.9, a.s.

$$\mathbb{P}\left[G_Z(X) - \mathbb{E}[G_Z(X)|\rho] \geq r - \rho - \mathbb{E}[G_Z(X)|\rho]\big|\rho\right] \leq \exp\left(-2\frac{(r - \mathbb{E}[G_Z(X)|\rho] - \rho)_+^2}{D_{G_Z}^2}\right) \tag{4.44}$$

which concludes the proof of the theorem.

4.1.3.4 Proof of Theorem 4.1.3

The proof of Eq. (4.10) follows from McDiarmid's inequality by observing that

$$\begin{aligned}\mathbb{P}[G(X,Z) \leq r] &= \mathbb{P}\left[G_Z(X) \leq r - \rho\right] \\ &\leq \mathbb{P}\left[G_Z(X) \leq r - \rho_{\min}\right].\end{aligned} \tag{4.45}$$

4.1.3.5 Proof of Theorem 4.1.4

Observe that

$$\mathbb{P}[G(X,Z) \geq r] = \mathbb{P}\left[G(X,Z) - \mathbb{E}[G] \geq r - \mathbb{E}[G]\right]. \tag{4.46}$$

Define

$$Y_1 = \mathbb{E}\left[G(X,Z)\middle|X_1\right] - \mathbb{E}[G] \tag{4.47}$$

$$Y_k = \mathbb{E}\left[G(X,Z)\middle|X_1,\ldots,X_k\right] - \mathbb{E}\left[G(X,Z)\middle|X_1,\ldots,X_{k-1}\right] \tag{4.48}$$

for $2 \leq k \leq N$ and

$$Y_{N+1} = G(X,Z) - \mathbb{E}\Big[G(X,Z)\Big|X\Big]. \tag{4.49}$$

Using the independence between X and Z, $\mathbb{E}\left[G(X,Z)\middle|X\right] = G_Z(X)$ and $Y_{N+1} = \rho$. Using Theorem 4.1.9 conclude that

$$\mathbb{P}\left[G(X,Z) - \mathbb{E}[G] \geq r - \mathbb{E}[G]\right] \leq \exp\left(-2\frac{(r-\mathbb{E}[G])_+^2}{(\rho_{\max}-\rho_{\min})^2 + D_{G_Z}^2}\right) \tag{4.50}$$

which concludes the proof of the theorem.

4.1.4 UQ in one dimension

This section proposes theorems for the special case of uncertainty quantification in one aleatoric dimension. The use of concentration-of-measure phenomenon in one dimension decreases the sharpness of the inequalities expected in high dimensions but allows for unique expressions of the theory for such special cases.

Let G be a function of a single known random parameter X and an unknown random parameter Z; X and Z may be correlated. The goal is to obtain the sharpest possible bound for

$$\mathbb{P}[G(X,Z) \geq r]. \tag{4.51}$$

4.1.4.1 McDiarmid's inequality.

Using McDiarmid's inequality,

$$\mathbb{P}[G(X,Z) \geq r] \leq \exp\left(-2\frac{(r-\mathbb{E}[G])_+^2}{D_G^2}\right) \tag{4.52}$$

where

$$D_G = \sup_{x,z} G(x,z) - \inf_{x',z'} G(x',z'). \tag{4.53}$$

Case: X can't be precisely controlled. This was the situation for the PSAAP application in §3.4 prior to controllable variables such as plate obliquity and thickness being introduced, and just as (uncontrollable) velocity was being considered.

If the mean $\mathbb{E}[G]$ is known, the only thing that needs to be estimated/bounded is D_G. D_G given by Eq. (4.53) is the solution of an optimization problem and the value of z can't be controlled. The fact that the value of x can't be precisely controlled does not significantly affect the strategy (with only one input) since the interest is in estimating $\sup G$ and

$\inf G$. Then one can apply the optimization algorithm to the approximate control of x to achieve/estimate the sup and inf of G. More precisely X is controlled through parameters λ, in the sense that $X = H(\lambda, Z_2)$ where Z_2 is an unknown unknown. (Consider for the PSAAP example that attempts at controlling the projectile velocity could be by turning a knob for chamber hydrogen pressure — λ_1 — and the mass of the launch package — λ_2.) It follows that

$$D_G = \sup_{\lambda, z_1, z_2} G\left(H(\lambda, z_2), z_1\right) - \inf_{\lambda, z_1, z_2} G\left(H(\lambda, z_2), z_1\right). \tag{4.54}$$

Use of the numerical model $F(X)$. The numerical model F can be used to estimate D_G through the inequality

$$D_G \leq D_F + D_{G-F} \tag{4.55}$$

where

$$D_F = \sup_x F(x) - \inf_x F(x) \tag{4.56}$$

and

$$D_{G-F} = \sup_{x,z} \left(G(x,z) - F(x)\right) - \inf_{x,z} \left(G(x,z) - F(x)\right). \tag{4.57}$$

Since the value of x can't be precisely set (experimentally) one has to measure its precise value and use it in the numerical simulation in order to evaluate $G(x,z) - F(x)$.

Unknown mean. Let $Y^1, \ldots, Y^q$ be q independent copies of $G(X,Z)$. Take the average of these samples to be a mean estimator, $< G >_q$

$$< G >_q := \frac{\sum_{i=1}^q Y_i}{q}. \tag{4.58}$$

Using McDiarmid's inequality,

$$\mathbb{P}\left[< G >_q \leq \mathbb{E}[G] - r\right] \leq \exp\left(-2q\frac{r^2}{D_G^2}\right). \tag{4.59}$$

It follows that for

$$\alpha_{\epsilon'} := \frac{D_G}{\sqrt{2q}}\sqrt{\ln\frac{1}{\epsilon'}} \tag{4.60}$$

so then

$$\mathbb{P}\left[< G >_q \leq \mathbb{E}[G] - \alpha_{\epsilon'}\right] \leq \epsilon'. \tag{4.61}$$

Hence the following theorem

Theorem 4.1.11. *With probability $1-\epsilon'$ on the q samples, for all $r > \inf_{x,z} G(x,z)$*

$$\mathbb{P}\left[G(X,Z) \geq r\right] \leq \exp\left(-2\frac{(r - < G >_q - \alpha_{\epsilon'})_+^2}{D_G^2}\right) \tag{4.62}$$

with

$$\alpha_{\epsilon'} := \frac{D_G}{\sqrt{2q}}\sqrt{\ln\frac{1}{\epsilon'}}. \tag{4.63}$$

Optimal estimate with an empirical mean. Let $Y^1, \ldots, Y^q$ be q independent copies of $G(X,Z)$. Write

$$< G >_q := \frac{\sum_{i=1}^q Y_i}{q}. \tag{4.64}$$

Theorem 4.1.12. *With probability $1-\epsilon'$ on the q samples, for all $r > \inf_{x,z} G(x,z)$*

$$\mathbb{P}[G(X,Z) \geq r] \leq \min\left(1, \frac{< G >_q + \alpha'_\epsilon - \inf_{x,z} G(x,z)}{r - \inf_{x,z} G(x,z)}\right) \tag{4.65}$$

with

$$\alpha_{\epsilon'} := \frac{D_G}{\sqrt{2q}}\sqrt{\ln\frac{1}{\epsilon'}}. \tag{4.66}$$

Theorem 4.1.13. *With probability* $1-\epsilon'$ *on the* q *samples, for all* $r < \sup_{x,z} G(x,z)$

$$\mathbb{P}[G(X,Z) \leq r] \leq \min\left(1, \frac{\sup_{x,z} G(x,z) - < G >_q - \alpha'_\epsilon}{\sup_{x,z} G(x,z) - r}\right) \tag{4.67}$$

with

$$\alpha_{\epsilon'} := \frac{D_G}{\sqrt{2q}}\sqrt{\ln\frac{1}{\epsilon'}}. \tag{4.68}$$

Optimal bound from the empirical distribution. Write $Y = G(X,Z)$. Let $Y^1, \ldots, Y^q$ be q independent copies (samples) of Y. What is the best possible estimate/the sharpest bound on

$$\mathbb{P}[Y \geq r] \tag{4.69}$$

when $Y^1, \ldots, Y^q$ is the only available information on the probability distribution function? D_G or G itself may be precisely known in this case.

From those q samples one can compute an empirical mean and use McDiarmid's inequality, Eq. (4.62) or Eq. (4.65) to bound Eq. (4.69), but it will be shown that those bounds won't be optimal.

$Y^1, \ldots, Y^q$ being i.i.d. the only usable information lies in the empirical distribution

$$\nu_q := \frac{\delta_{Y^1} + \cdots + \delta_{Y^q}}{q}. \tag{4.70}$$

Let $\epsilon' \in (0,1)$ be given where $1-\epsilon'$ will be the confidence on the estimation. Write

$$\mathbb{B}(r,\epsilon',q) := \Big\{\Phi \,:\, \inf_{\mu} \mathbb{P}_\mu \left[\mathbb{P}_\mu[Y \geq r] \leq \Phi(r,\epsilon',\nu_q,q)\right] \geq 1-\epsilon'\Big\}. \tag{4.71}$$

In Eq. (4.71), μ stands for an arbitrary distribution on the range of G. An optimal estimator would minimize the following variational problem

$$\inf_{\Phi\in\mathbb{B}(r,\epsilon',q)} \sup_{\mu} \mathbb{E}\left[\big(\Phi(r,\epsilon',\nu_q,q) - \mathbb{P}_\mu[Y\geq r]\big)_+\right]. \tag{4.72}$$

Optimal estimator Begin with

$$C_n^k := \frac{n!}{k!(n-k)!}. \tag{4.73}$$

Write, for $p \in [0,1]$

$$h(p,\epsilon') := \max\Big\{h \,:\, \sum_{k=h}^{q} C_q^k p^k (1-p)^{q-k} \geq 1-\epsilon'\Big\}. \tag{4.74}$$

Using Bernoulli distributions in Eq. (4.72) one obtains that the optimal estimator only depends on $\nu_q[r,\infty]$ and q. To simplify notations, write for that optimal estimator

$$\Psi(s) := \Phi(r,\epsilon',\nu_q,q)\Big|_{q\nu_q[r,\infty]=s} \tag{4.75}$$

s is the number of samples above r among $Y^1,\ldots,Y^q$. $\Psi(s)$ is increasing in s and the "smallest" function such that

$$\text{For} \quad s \geq h(p,\epsilon'), \quad p \leq \Psi(s). \tag{4.76}$$

It follows that Ψ is the "smallest" function such that

$$0 \leq \inf_{p \in [0,1]} \Psi\left(h(p, \epsilon')\right) - p \tag{4.77}$$

which leads to

$$\Psi(s) := \sup_{p\,:\,h(p,\epsilon')=s} p. \tag{4.78}$$

Hence the following theorem

Theorem 4.1.14. *Writing s_q the number of samples in $Y^1, \ldots, Y^q$ above the threshold r, with probability $1 - \epsilon'$*

$$\mathbb{P}[G(X, Z) \geq r] \leq \sup_{p\,:\,h(p,\epsilon')=s_q} p \tag{4.79}$$

where the function h is defined by Eq. (4.74). Furthermore the upper bound/estimate is optimal if the only available information on the probability distribution function of G is limited to the q samples $Y^1, \ldots, Y^q$.

One can't expect any sharper result than Eq. (4.79).

These adaptations to bounds on probability-of-failure for various cases exemplifies and nearly justifies the use of concentration-of-measure framework to develop rigorous UQ expressions.

4.1.5 PSAAP application

The same setting and goals of computations and experiments hold for this example as described in §3.4.1, but with application towards performing UQ in a general setting. Uncontrollable velocity remains an issue, as discussed in §3.4.3. Another point must be made here about the velocity since its measurement precludes some uncertainty as well.

The speed measurement requires measurement of both distance and time, which are both discrete because a digital video of the projectile flight is used for both measurements. The distance is computed by calculating the difference in pixel position of the front of the rasterized flash created by the projectile breaking through a mylar film. The time is the frame time difference between the distance measurements.

An attempted nominal value of $X \approx 2500\,\mathrm{m/s}$ results in experimental projectile velocities ranging from $X \approx 2300\,\mathrm{m/s}$ to $X \approx 2800\,\mathrm{m/s}$, or $\xi = [2300, 2800]\,\mathrm{m/s}$ with an unknown probability law (or empirical law if experiments have been done *a priori*).

The theoretical uncertainty quantification methods described above account for this lack of control and precise measurement, and the following subsections discuss preparation and computation of the UQ quantities desired.

4.1.5.1 1-D Verification

The least intensive step in the uncertainty quantification framework besides the estimation of the mean is to calculate the verification diameter, D_F^2. In a 1-D setting, D_F^2 is simply calculated in Eq. (4.56). This computation requires two separate optimizations of $F(X)$ over X. Normally, this would be done using an optimization routine in Sandia National Laboratory's DAKOTA. In this example, simply take the max and min values from the discrete samples taken from the model. This and all following calculations are done with MATLAB software unless otherwise specified.

Fig. 4.1 shows a portion of the values of $F(X)$ obtained $X \in [1800, 2800]\,\mathrm{m/s}$ in increments of $10\,\mathrm{m/s}$. Execution of the 100 Abaqus models took approximately 1300 minutes on the Caltech CACR's (Center for Advanced Computing Research) Shared Heterogeneous Cluster (SHC) using one AMD Opteron dual core dual processor node, employing parallel

processing for internal model calculations.

For this reason, all experimental data $G(X, Z)$ taken to this date and computational results $F(X)$ from the 100 simulations described above within the range defined by ξ are used to compute all necessary UQ parameters. Thus sup and inf calculations will simply be max and min calculations of the data and results for $X \in \xi$ (Eq. (4.80)).

$$D_F = \sup_{X \in \xi} F(X) - \inf_{X \in \xi} F(X) \tag{4.80}$$

4.1.5.2 1-D Validation

Here, the general CoM expressions developed in §4.1.1 will be employed here for $N = 1$ in the 1-D setting of this PSAAP validation example, instead of defaulting to the obvious 1-D inequalities developed in §4.1.4.

The simplicity of the verification diameter computation is not carried over into the computation of validation diameters which involve obtaining $G(X, Z)$ because of the lack of control over X, described in §3.4.3. Therefore, an iterative optimization scheme that dictates the projectile speed at each successive iteration from the previous iteration's value cannot be used because, in general, $X_{measured} \neq X_{attempted}$.

***A priori* available data** $G(X, Z)$ is given *a priori* by the experimentalists, and contains multiple perforation area measurements for each measure velocity. Given the time costs of each experiment, there is not statistically sufficient data to compute statistical measures on G. For this case, assume that the values that are computed are statistically significant. The values of $G(X, Z)$ given appear as ∘ in Fig. 4.1.

To construct $G_Z(X)$, use projectile speeds for which data exists: $x : \exists\, G(x \in \xi, Z)$, and compute the mean of all N_i samples at x_i as in Eq. (4.81). This is the red line in Fig. 4.1.

$$G_Z(x_i) = \frac{\sum_{j=1}^{N_i} G(x_i, z_j)}{N_i}, \forall i : \exists\, G(x_i \in \xi, Z) \tag{4.81}$$

To construct $G_Z(X) - F(X)$, first assume that there exist more data points in ξ for $F(X)$ than for $G_Z(X)$, since model executions are generally less expensive in time and cost than their experimental counterparts to obtain. For each $x : \exists F(x_i \in \xi)$, compute $G_Z^*(x_i) - F(x_i)$ where $G_Z^*(x_i)$ is the linear interpolation at v_i between the two nearest discrete points in $G_Z^*(X)$. This is the black line in Fig. 4.1.

The values $\rho(X, Z)$ are computed simply as $\rho(X, Z) := G(X, Z) - G_Z(X)$ where $X = X_i$, $\forall\, i : \exists\, G(X_i \in \xi, Z)$, and appear as $+$ in Fig. 4.1.

With all necessary trendlines constructed, the UQ calculations follow.

UQ Computations Calculated probabilities-of-failure, p_f, for Theorems 4.1.2, 4.1.4, and 4.1.6 are straightforwardly defined once diameters are computed, though Theorem 4.1.2 requires expectations on the law of ρ. Recalling that Theorem 4.1.1 does not use diameters, and that the current application defines failure as perforated area less than a threshold value, p_f in Theorem 4.1.1 becomes Eq. (4.82) from Eq. (4.2).

$$P[G(X, Z) \leq r] \leq \min\left(1, \frac{\sup_{x,z} G(x, z) - \mathbb{E}[G]}{\sup_{x,z} G(x, z) - r}\right) \tag{4.82}$$

For Theorem 4.1.2, the expression $\mathbb{E}[G_Z(X)|\rho]$ is the expectation on the distribution of $G_Z(X)$ given a value of ρ. With statistically few samples, take $\sigma_\rho = \sqrt{\text{var}(\rho(x, z) : x \in \xi)}$. Then, assuming for now that the distribution on ρ across all $x \in \xi$ are identical (though in general this may not be true), assign $\rho(X, Z) \sim N(0, \sigma_\rho)$. Now, based on the data, a law can be associated to ρ from which one can sample and evaluate Theorem 4.1.2. For 100

samples of ρ, p_f is evaluated from Eq. (4.5) at various thresholds in Fig. 4.2(a) and various values of thresholds and $\mathbb{E}[G]$ in Fig. 4.3(a).

The margins M needed to achieve p_f in Fig. 4.2(b) are also presented. Showcasing results in terms of margins is intended for informing design decisions when a p_f must be met. For Theorems 4.1.2, 4.1.4, and 4.1.6, the margin is defined in Eq. (4.83). For Theorem 4.1.1, the margin is defined in Eq. (4.84). Note that computing valid margins does not guarantee that the margins can be achieved in the physical system. One has to be wary not to accept margins that may lead to "out-of-bounds" responses on the physical system. The uncertainty U represents the denominator of the CoM inequality being used — $U = D^2_{G_Z}$ for Theorem 4.1.2, $U = D^2_Z + D^2_{G_Z}$ for Theorem 4.1.4, and $U = D^2_Z + (D_F + D_{G_Z - F})^2$ for Theorem 4.1.6.

$$M = \sqrt{-\log(p_f) U/2} \tag{4.83}$$

$$M = \sup_{x,z} G(x,z) - \mathbb{E}[G] \frac{1 - p_f}{p_f} \tag{4.84}$$

4.1.5.3 Results

Table 4.1 presents values computed for diameters described above, and the values of velocity that the respective minimum and maximum values were found. It is also exceedingly explanatory of the required computations done for the various diameters that have been recovered from the computations to produce Fig. 4.1. For example, the computation of D_Z, according to Theorem 4.1.4, suggests that one must find the largest and smallest deviations of the noise from the epistemically-smoothed experimental curve $G_Z(X)$ over all x ($\rho_{\max}$ and $\rho_{\min}$, respectively). The diameter D_Z is thus computed as $D_Z = \rho_{\max} - \rho_{\min}$. Visual

inspection of Fig. 4.1 shows that $\rho_{\min} = -2.99\,\text{mm}^2$ is found at $V_{\min} = 2770\,\text{m/s}$, and coincidentally $\rho_{\max} = 3.73\,\text{mm}^2$ at the same speed $V_{\max} = 2770$ m/s. D_F is the largest difference between two ordinate values of $F(X)$ over all X. The curve of $F(X)$ shows a linear trend with some perturbations, which prohibit the *min* and *max* values from being the endpoints of $F(X)$: $F_{\min} = 1.54\,\text{mm}^2$ at $V_{\min} = 2410\,\text{m/s}$ and $F_{\max} = 2.66\,\text{mm}^2$ at $V_{\min} = 2410\,\text{m/s}$ to result in $D_F = |F_{\max} - F_{\min}|$. These illustrations should illuminate the meaning of the remaining diameters presented in Table 4.1.

Table 4.1: Diameter calculation results for PSAAP 1-D UQ using velocity as a random variable and hole perforation area as a performance measure

Diameter	Value [mm^2]	$A_{\min}$[mm^2]	$V_{\min}$ [m/s]	$A_{\max}$[mm^2]	$V_{\max}$ [m/s]
D_Z	6.72	-2.9925	2770	3.7275	2770
D_F	3.35272	4.6351	2410	7.98783	2760
D_{G_Z-F}	4.74424	-5.23434	2400	-0.4901	2510
D_{G-F}	6.72	-5.56052	2770	1.15948	2770
D_{G_Z}	5.3425	0	2400	5.3425	2770
$\bar{D}_F$	1.73801	-	-	-	2410
$\bar{D}_{G-F}$	3.56101	-	-	-	2770

Other relevant values computed are $\mathbb{E}[G] = 4.61067\text{mm}^2$, $\mathbb{E}[F] = 6.373\text{mm}^2$, $\mathbb{E}[G(X,Z) - F(X)] = -2.1289\,\text{mm}^2$, and $\sigma_\rho = 1.32654\,\text{mm}^2$.

4.1.5.4 Discussion

Fig. 4.1 provides the basis for which Figs. 4.2-4.4 are created. For the given range of velocities, the blue curve shows perforation areas for the numerical simulations $F(X)$. The data points collected by experimentalists, corresponding to $G(X,Z)$, are denoted by red circles. The smoothed curve over these points is given the by the red curve, denoted $G_Z(X)$. The difference between the smoothed experimental curve and the numerical simulations is the black curve denoted by $G_Z(X) - F(X)$. The difference between the experimental points and the numerical simulations is given by the magenta circles and is denoted $(GX,Z) -$

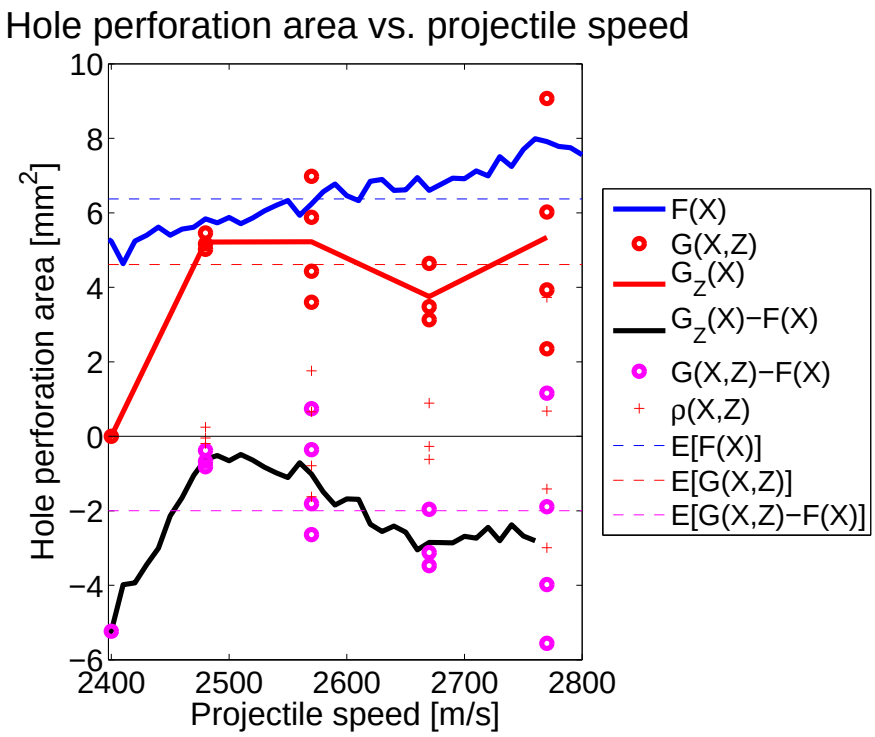


Figure 4.1: Trendlines for values computed in the 1d PSAAP application

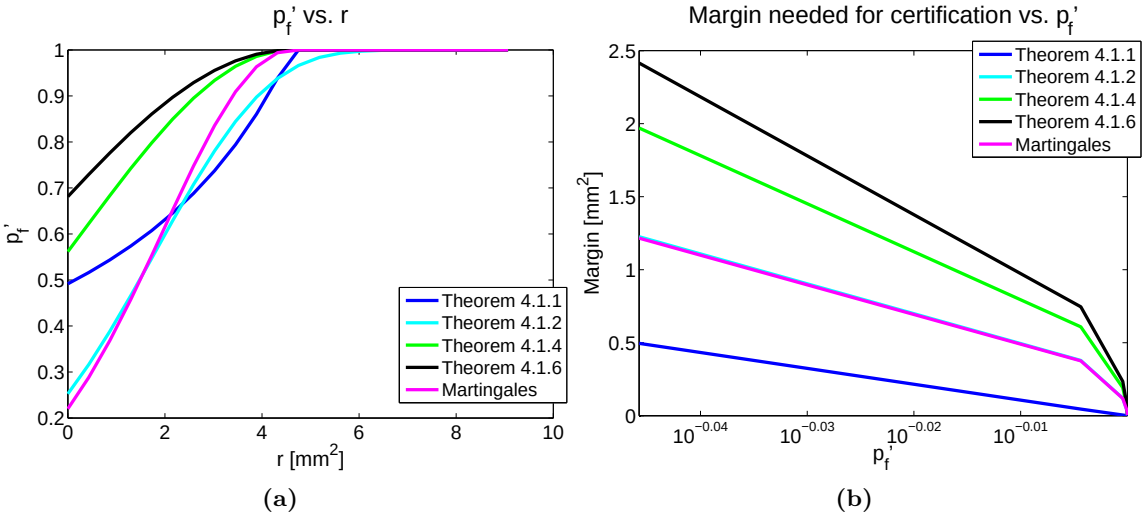


Figure 4.2: (a) Probabilities-of-failure upper bounds for varying threshold values and (b) margins required to achieve probabilities-of-failure upper bounds

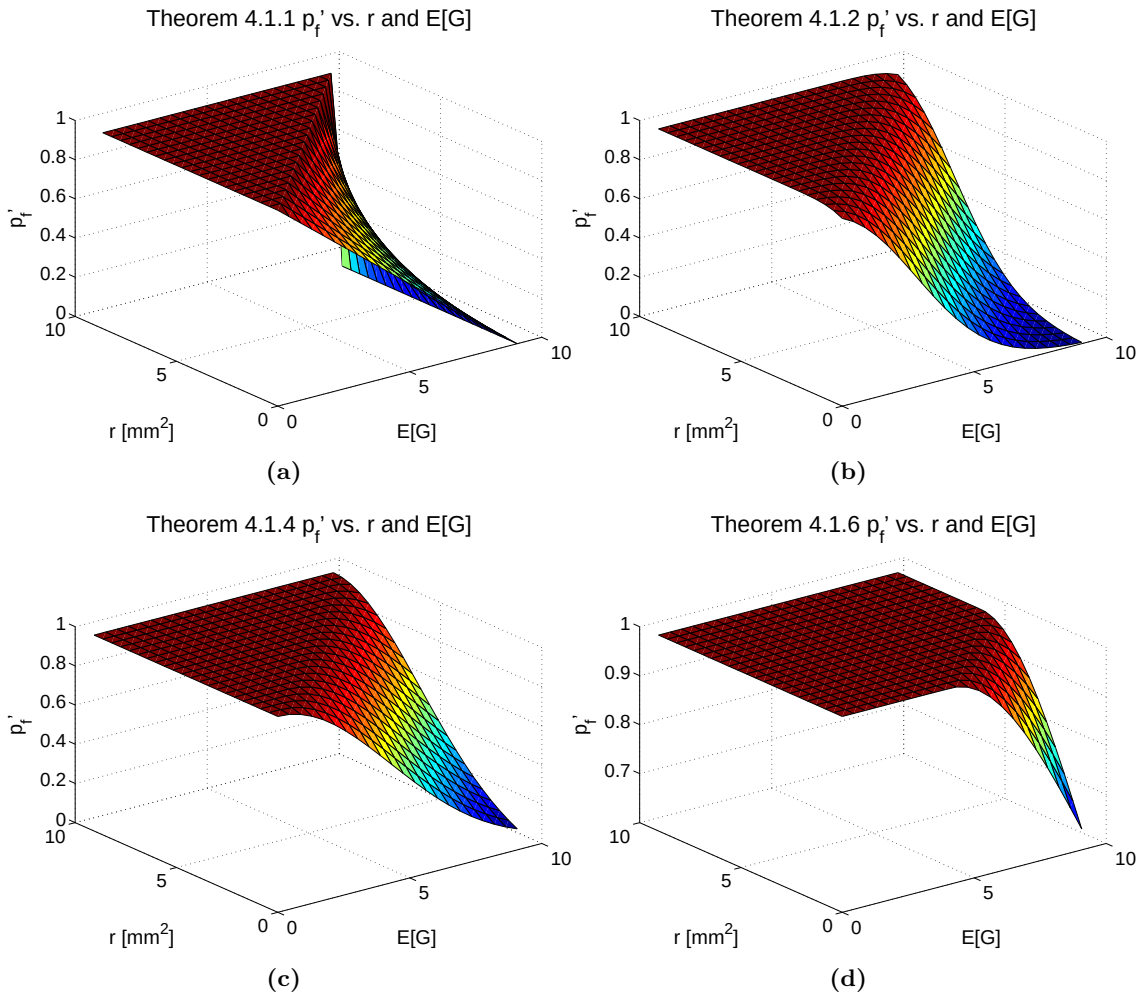


Figure 4.3: Probabilities-of-failure upper bounds for varying threshold and $\mathbb{E}[G]$ values using Theorems (a) 4.1.1, (b) 4.1.2, (c) 4.1.4 and (d) 4.1.6

$F(X)$. The curve for $\rho(X, Z)$ defined as the points $G(X, Z) - G_Z(X)$ is denoted by red crosses. Scalar values of means are denoted by horizontal dashed lines, blue for $\mathbb{E}[F(X)]$, red for $\mathbb{E}[G(X, Z)]$, and magenta for $\mathbb{E}[G(X, Z)-F(X)$. Both $F(X)$ and $G_Z(X)$ depict a general increase in perforation area with velocity, though the experiments exhibit more "cliff"-like behavior, while the numerical simulations do not in this range of velocity. The numerical simulations generally overestimate the perforation areas when compared to experiment, which is reflected in the $G_Z(X)-F(X)$ curve and $G(X, Z)-F(X)$ points as mostly negative values, and $\mathbb{E}[F] > \mathbb{E}[G(X, Z)]$.

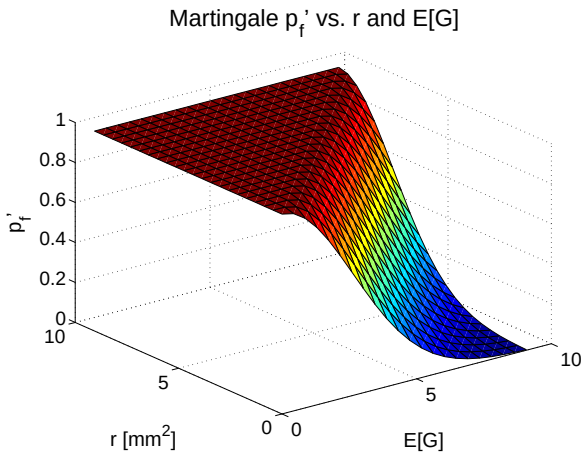


Figure 4.4: Probabilities-of-failure upper bounds for varying threshold and $\mathbb{E}[G]$ values using Eq. (4.19)

For the increasing complexity of the system's p_f upper-bound, Figs. 4.2-4.4 comparatively depict the various CoM UQ quantities of interest as expressed theoretically in Theorems 4.1.1, 4.1.2, 4.1.4, and 4.1.6. Generally, a lower p_f upper-bound is obtained when the least amount of information is used. As more information is available, the p_f bounds obviously account for more knowledge of the design space and the conservativeness imparted upon the p_f inequality compounds and increases the upper-bound, as observed in Fig. 4.2(a). The margins required for certification at various probability-of-failure thresholds accordingly increase as the complexity of the theorems increases from Theorem 4.1.1 to 4.1.6, as observed in Fig. 4.2(b).

Fig. 4.3 and Fig. 4.4 show surfaces of p_f upper-bounds for varying failure thresholds r and varying values of $\mathbb{E}[G(X,Z)]$. Although $\mathbb{E}[G(X,Z)]$ is already a known scalar value, these plots help to show the effect of both of these affecting the margin and thus the p_f upper-bound. Notice that only the Martingale inequality of Eq. (4.19) decreases the p_f upper-bound as the complexity of the inequalities increases. These figures show trivial estimates of p_f upper-bounds for large failure thresholds r and small $\mathbb{E}[G(X,Z)]$, and alternatively

non-trivial upper-bounds for small r and large $\mathbb{E}[G(X,Z)]$, consistent with the inequalities used.

4.2 Hierarchical UQ

This section proposes a modular uncertainty quantification scheme to perform hierarchical certification using the concentration-of-measure phenomenon for systems whose input-output relation is described by graphical models. Though the development here is comprehensive, more details and associated proofs can be found in [68].

4.2.1 Introduction

4.2.1.1 Motivation

Computation of the verification diameter D_F requires computing $D_{F,i}$, for $i = 1, \ldots, N$, which in turn requires solving the constrained optimization problem in Eq. (2.1) with $N+N_i$ scalar decision variables where each of the $i = 1, \ldots, N$ variables may be vector-valued. Besides computational considerations as in Remark 2.1.1, the uncertainty quantification for the input-output map F may not be programmatically feasible. Consider the case where the map F represents the behavior of a large-scale system which is an interconnection of multiple subsystems and the models of these subsystems are only available to a different analyst (say because of organizational or political constraints). In such cases, there is a need for an uncertainty quantification framework where each analyst executes the analysis for their subsystem and the overall quantification is done based on the results for the subsystems.

Furthermore, many physical systems evolve on multiple length and/or time scales: the characteristics/dynamics at one scale affect those at other scales, enabling a *divide-and-*

conquer based approaches for analysis and design. For example, multiscale modeling strategies for advanced materials such as high-purity bcc single crystals are discussed in [13]. See also [3] and [14], which highlight the importance of multiscale modeling and uncertainty quantification for such models. To develop the methodology succinctly, consider the following electrical circuit example.

4.2.1.2 Electrical circuit example

Consider an LC electrical circuit with a capacitance of C and inductors placed on the edges of a Sierpinski triangle, a fractal structure with the same geometry on different scales and many applications such as in multiband fractal antenna technologies [5], suggesting the possibility for hierarchical analysis. A typical Sierpinski triangle is shown in Fig. 4.5, where the edges in the triangular structure correspond to inductances labeled by $L_1, \ldots, L_{27}$, which are uncertain parameters taking values in bounded intervals. The goal is to compute the uncertainty in the performance output such as the equivalent inductance L_{eq} between the nodes where the capacitance is connected, or the frequency of resonance of the LC circuit (see §4.2.4.1 for a short list of applications where this resonance effect is exploited). Now model the relation between L_{eq} and the individual inductances by

$$L_{eq} = F(L_1, \ldots, L_{27}). \tag{4.85}$$

One can obtain an upper bound, for example, on

$$\mathbb{P}_{L_1,\ldots,L_{27}}\left[L_{eq} \geq \tau\right]$$

for some given $\tau > 0$ by McDiarmid's inequality, e.g., Eq. (2.1).

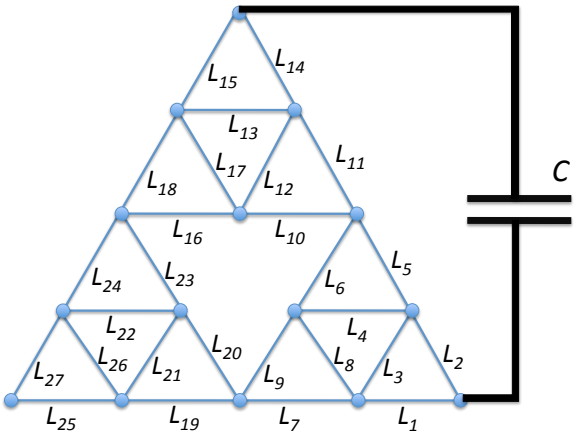


Figure 4.5: The LC electrical circuit with inductors placed on the edges of a Sierpinski triangle of depth 3

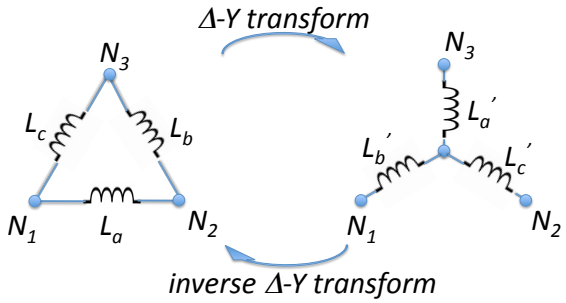


Figure 4.6: Equivalent inductive circuit elements in Δ and Y configurations

Figure 4.7: The recursive creation of Sierpinski triangles of different depths

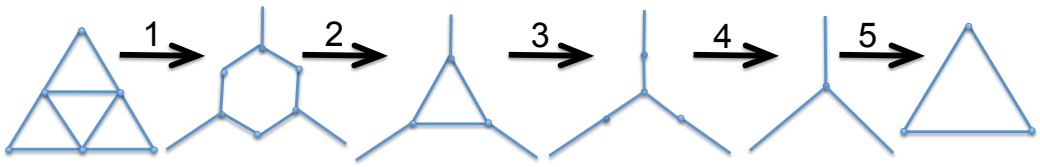


Figure 4.8: Operations to reduce the depth of a Sierpinski triangle by one

On the other hand, one can exploit the (geometric) self-similarity in the Sierpinski triangles of different type for a modular analysis procedure. Fig. 4.7 illustrates the recursive generation of Sierpinski triangles starting from an equilateral triangle. The $\Delta - Y$ transform [25] in Fig. 4.6 is an invertible map which translates inductive circuit elements from the Δ configuration to the Y (or star) configuration through $L'_a = \frac{L_b L_c}{L_a+L_b+L_c}$, $L'_b = \frac{L_c L_a}{L_a+L_b+L_c}$, and $L'_c = \frac{L_a L_b}{L_a+L_b+L_c}$, allowing the generation of two inductive circuit elements with the same equivalent inductance on two different Sierpinski triangles. Fig. 4.8 illustrates one such translation where arrow 1 is for the $\Delta - Y$ transform applied to the three small triangles at the edges, arrow 2 is for simple manipulations, arrow 3 is for another $\Delta - Y$ transform, arrow 4 is for simple manipulations, and arrow 5 is for the inverse of the $\Delta - Y$ transform. The depth of a Sierpinski triangle is *one more than the number of times* that the operations shown in Fig. 4.8 are performed to obtain the configuration of a single triangle. Specifically, the operations in Fig. 4.8 reduce the level by 1. Thus, the Sierpinski triangles shown in the top part of Fig. 4.7 are of depths $1, \ldots, 5$ from left to right.

The geometric (self-)similarity of the Sierpinski triangles of different depths gives the means to completely modularize each level of the inductive circuit element. For example, Fig. 4.9 illustrates how certain distinct groups of inductances on a Sierpinski triangle of depth 3 are mapped to distinct groups on a Sierpinski triangle of depth 2 using the operations shown in Fig. 4.8. By successive applications of these operations, one obtains the series of LC circuits with the same equivalent inductances on the Sierpinski triangles of different

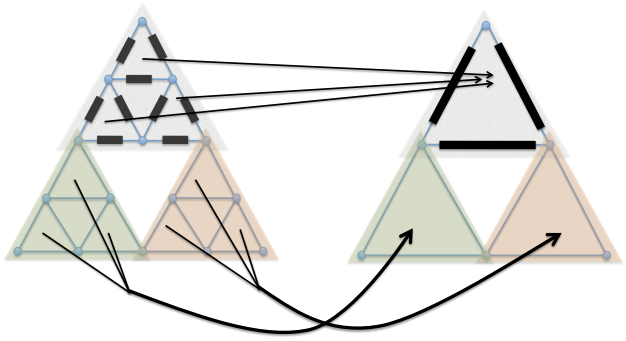


Figure 4.9: Illustration of the modular structure obtained through the application of the operations in Fig. 4.8

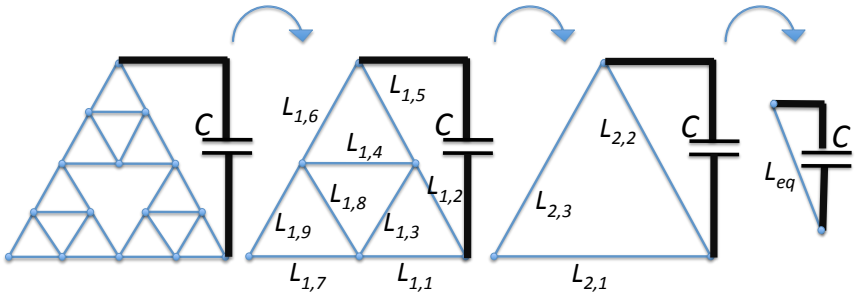


Figure 4.10: The hierarchical determination of L_{eq} for the inductive circuit element on a Sierpinski triangle of depth 3 in an LC electrical circuit through successive application of the $\Delta - Y$ transform

depth as shown in Fig. 4.10. Regarding the structure in Fig. 4.10, consider an alternative model for the relation between L_{eq} and the individual inductances on the edges of the Sierpinski triangle of depth 3.

$$\begin{aligned}
L_{eq} &= F_2(L_{2,1}, L_{2,2}, L_{2,3}) \\
L_{2,i} &= F_{1,i}(L_{1,1}, \ldots, L_{1,9}), \qquad \text{for } \; i = 1,2,3 \\
L_{1,i} &= F_{0,i}(L_1, \ldots, L_9), \qquad \text{for } \; i = 1,2,3 \\
L_{1,i} &= F_{0,i}(L_{10}, \ldots, L_{18}), \qquad \text{for } \; i = 4,5,6 \\
L_{1,i} &= F_{0,i}(L_{19}, \ldots, L_{27}), \qquad \text{for } \; i = 7,8,9.
\end{aligned} \tag{4.86}$$

Note that each submodule/subsystem (i.e., $F_{0,i}$, $F_{1,i}$, and F_2) have (at most) 9 scalar

input parameters, whereas F has 27 scalar input parameters. The goal in this section is to exploit the modularity in the model in Eq. (4.86) for propagating the uncertainty in $L_1, \ldots, L_{27}$ to L_{eq} (or resonance frequency of the circuit) in a hierarchical and relatively computationally efficient manner: first determine the uncertainty in $L_{1,i}$, then in $L_{2,i}$, and finally in L_{eq}. §4.2.4 illustrates numerical examples using the LC circuit structure in Fig. 4.5. A hypothetical running example is introduced next for generality of demonstration, and will be used along with the electrical circuit example for exposition of the theory.

4.2.1.3 Hypothetical running example

Consider a model for which the dependencies between its variables are represented by the graph in Fig. 4.11. The "nodes" in the graph represent the random variables in the system, with node i corresponding to the variable X_i, and the directed edges representing the dependence between these variables. For example, the edge from node 8 to node 4 indicates that X_4 is an explicit function of X_8, and the lack of an edge from node 2 to node 3 indicates that X_3 is not an explicit function of X_2. However, having no edge from some node i to another node j does not imply an independence between X_j and X_i. For example, in Fig. 4.11 there is no edge from node 14 to node 6 but X_6 depends on X_{14} through X_7. Denote the nodes which do not have any incoming edges as the *fundamental* nodes. The variables corresponding to fundamental nodes do not depend on any other variable. The node from which there is no outgoing edge is called the *output* node. The model in Fig. 4.11 can be formally expressed in two forms: a direct map from the variables corresponding to the fundamental nodes to the output X_1

$$X_1 = F(X_2, X_{10}, X_{11}, X_{12}, X_{13}, X_{14}, X_{15}, X_{16}, X_{17}, X_{18}, X_{19}, X_{20})$$

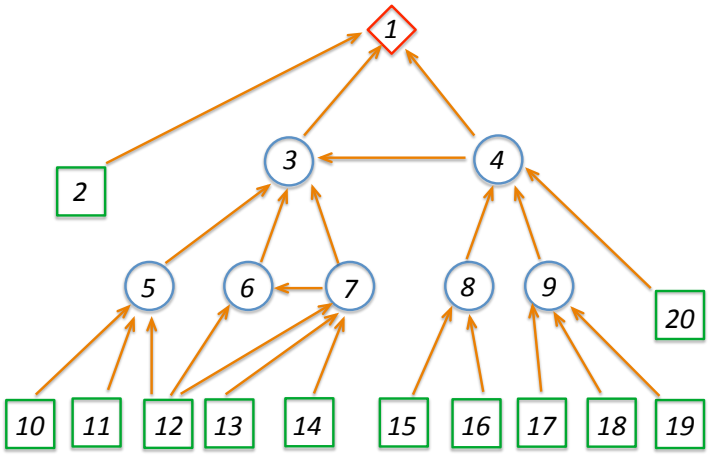


Figure 4.11: The graph representation of the input-output relations between the "variables" of a hypothetical model

and a hierarchical description

$$
\begin{aligned}
X_1 &= H_1(X_2, X_3, X_4) \\
X_3 &= H_3(X_4, X_5, X_6, X_7) \\
X_4 &= H_4(X_8, X_9, X_{20}) \\
X_5 &= H_5(X_{10}, X_{11}, X_{12}) \\
X_6 &= H_6(X_7, X_{12}, X_{13}, X_{14}) \\
X_7 &= H_7(X_{12}, X_{13}, X_{14}) \\
X_8 &= H_8(X_{15}, X_{16}) \\
X_9 &= H_9(X_{17}, X_{18}, X_{19}).
\end{aligned}
\tag{4.87}
$$

In the following subsection, a hierarchical uncertainty propagation scheme will be formally discussed. The following is a preview of the main idea. Referring back to Eq. (2.1), note that the square of the verification diameter D_F^2 is the sum of the contributions $D_{F,i}^2$ due to each fundamental variable (i.e., $i = 2, 10, 11, \ldots, 20$). The main result to be stated in §4.2.3.1 establishes an upper bound on $D_{F,i}$ by accounting for the uncertainty in x_1 due

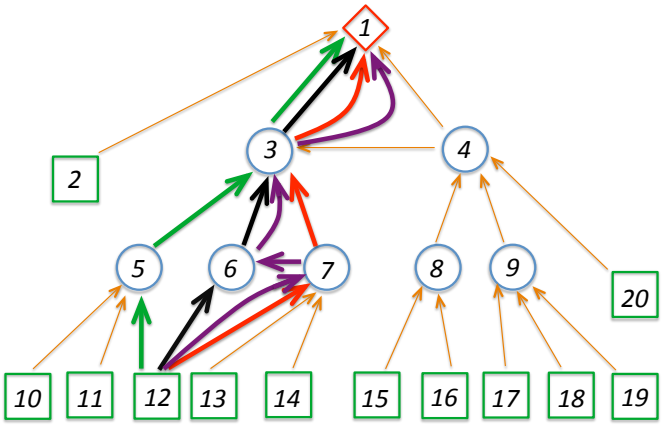


Figure 4.12: Distinct paths connecting the fundamental node 12 to the output node 1 are highlighted as green, black, red, and purple arrows

to each fundamental variable through each *path* connecting that fundamental variable to the output x_1. As an example, the paths connecting the fundamental node 12 to the output node 1 are highlighted in Fig. 4.12 by the green, black, red, and purple arrows.

4.2.2 Graph representations

Graph theoretic notions and preliminary results are discussed in the subsequent sections.

4.2.2.1 Graph representation of input-output models

The graph representation in Fig. 4.11 accounts only for the dependencies between the variables of the systems and currently lacks specific information on the functional relation between the variables. Such graphical visualization will be utilized in developing the algorithmic procedure for hierarchical uncertainty propagation. The relevant graph theoretic notions are formally discussed next.

Consider a set of N random variables $X_i \in \mathbb{R}^{n_i}$ labeled by the finite set V of indices $V = \{1, \ldots, N\}$. Let $G = (V, E)$ be a directed graph where E is a set of ordered pairs (i, j) of indices, called the *edges* of the graph, with $i \in V$ and $j \in V$. Each $i \in V$ is called a *node* of

G and corresponds to the variable/parameter X_i of the underlying physical system and/or the mathematical model. Each edge (i, j) models the dependence of the variable X_j on the variable X_i and $(i, j) \notin E$ if X_j is not an explicit function of X_i. That is, $(i, j) \in E$ if and only if X_j is an explicit function of X_i. The following definitions are used throughout this section.

- *Root of the graph*: Each node i of G such that $(i, j) \notin E$ for all $j \in V$ is called a *root* of the graph.

- *Fundamental node of the graph*: Each node i of G such that $(j, i) \notin E$ for all $j \in V$ is called *fundamental node* of the graph. The set of all fundamental nodes of the graph is denoted by $\mathcal{F}$ and corresponds to the system variables that do not depend on the other variables.

- *Path*: A group of ordered nodes $i_1, \ldots, i_n$ such that $(i_j, i_{j+1}) \in E$ for all $j = 1, \ldots, n-1$ is called a path γ in the graph G. The node i_1 is called the *initial node* of γ and i_n is the *terminal node* of γ. For given nodes i and j, the collection of all paths with the initial node i and the terminal node j is denoted by $\mathcal{P}(i, j)$.

- *Ancestor of a node*: For each $i \in V$, a node j is called an *ancestor* of i if $(j, i) \in E$. $\mathcal{N}_i$ denotes the set of all ancestors of i, i.e., $j \in \mathcal{N}_i$ if and only $(j, i) \in E$. Note that for each $i \in \mathcal{F}$, $\mathcal{N}_i$ is the empty set and for the root node, say i, $\mathcal{N}_i$ is non-empty.

- *Depth of the graph*: The depth of a directed, acyclic graph with the output node $i_{output} \in V$ is defined as the quantity

$$\max_{i \in \mathcal{F}} \quad \max_{\gamma \in \mathcal{P}(i, i_{output})} \quad \text{number of edges in } \gamma.$$

- *Tree*: A graph is called a *tree* if from each fundamental node there exists a single path connecting to the output node.

The root node in the graph in Fig. 4.11 is node 1 and the fundamental nodes are $2, 10, 11, \ldots, 20$. For the same graph, $\mathcal{P}(12, 1)$ is highlighted by the green, black, purple, and red arrows in Fig. 4.12. The following assumptions are made on the directed graphs considered in this section.

1. *Existence of root (output) variable*: There exists a root node of the graph G. Without loss of generality, label this node with 1. The variable X_1 corresponding to the root of the graph is considered to be the output (performance measure) of the underlying physical system. For simplicity, assume that there is a single output (performance measure).

2. *Acyclic graph:* The graph G is acyclic, i.e., there is no path with identical initial and terminal nodes. This assumption ensures that the output of the system can be explicitly computed from all other variables without having to solve implicit equations.

3. The fundamental variables are independent random variables.

Remark 4.2.1. *Assumptions (2) and (3) can be removed by using concentration-of-measure inequalities that are more general than the McDiarmid's inequality in Eq. (2.1). See [57] for a discussion on concentration-of-measure inequalities with correlated random variables. Adaptation of such results in a hierarchical analysis setting is the subject of current research following from [68].* $\triangleleft$

4.2.2.2 Example: graph representation of the LC electrical circuit model

Recall the LC circuit example introduced in §4.2.1.2. Referring to Figs. 4.5 and 4.10 for notation, consider two hierarchical models with different graphs due to different grouping of the variables.

- *Ungrouped intermediate variables*: Rename the variables as follows:

$$\begin{aligned} X_1 &:= L_{eq} \\ X_i &:= L_{2,i-1} \qquad \text{for} \quad i = 2, 3, 4 \\ X_i &:= L_{1,i-4} \qquad \text{for} \quad i = 5, \ldots, 13 \\ X_i &:= L_{i-13} \qquad \text{for} \quad i = 14, \ldots, 40. \end{aligned} \tag{4.88}$$

  In this case, $\mathcal{F} = \{14, \ldots, 40\}$, $X_1, \ldots, X_{40} \in \mathbb{R}$, and the corresponding graph is shown in Fig. 4.13.

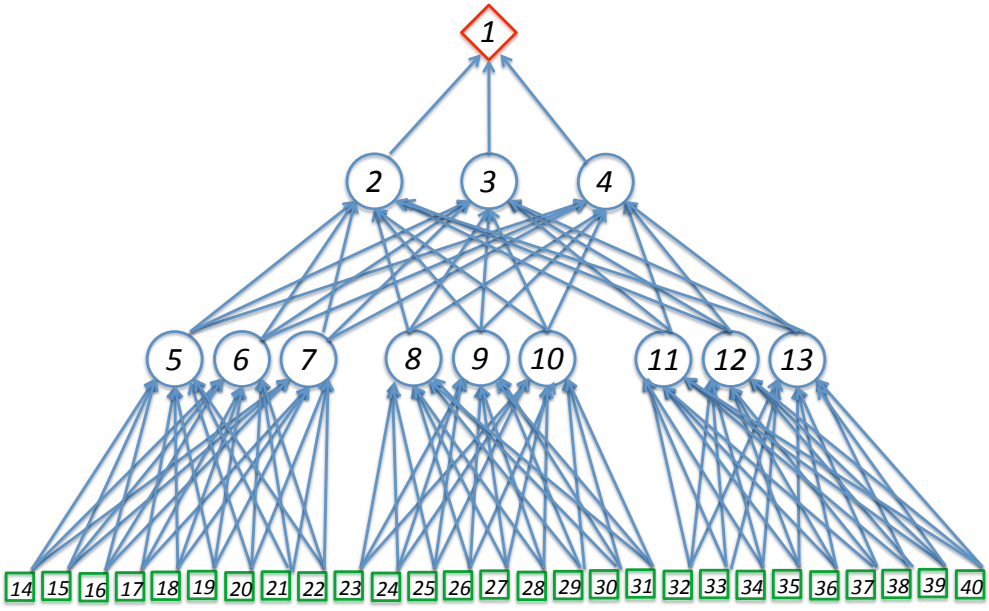


Figure 4.13: The graph representation of the relations between the variables of the hierarchical model for the LC circuit example in the case of ungrouped intermediate variables (i.e., all variables are scalars)

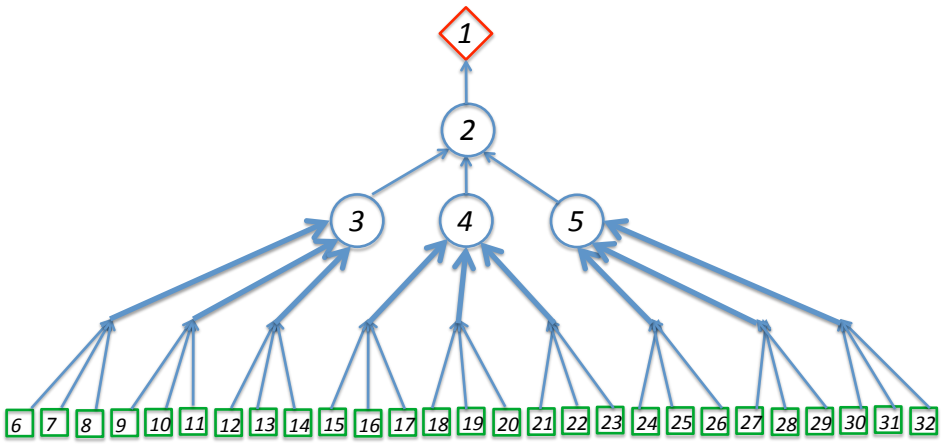


Figure 4.14: The graph representation of the relations between the variables of the hierarchical model for the LC circuit example in the case where the intermediate variables are grouped (i.e., all variables but the fundamental variables and the output variable are in $\mathbb{R}^3$)

- *Grouped intermediate variables*: Group the variables as

$$\begin{aligned} X_1 &:= L_{eq} \\ X_2 &:= (L_{2,1}, L_{2,2}, L_{2,3}) \\ X_{i+2} &:= \left(L_{1,3(i-1)+1}, L_{1,3(i-1)+2}, L_{1,3(i-1)+3}\right) && \text{for } \; i = 1, 2, 3 \\ X_{i+5} &:= L_i && \text{for } \; i = 1, \ldots, 27. \end{aligned} \tag{4.89}$$

  In this case, $\mathcal{F} = \{6, \ldots, 32\}$, $X_1 \in \mathbb{R}$ and $X_2, \ldots, X_{14} \in \mathbb{R}^3$, and the corresponding graph is shown in Fig. 4.14. Note that grouping the variables as in Eq. (4.89) simplifies the graph to a tree.

Note from Figs. 4.13 and 4.14 that the graph for the case with ungrouped intermediate variables includes a larger number of paths (81 distinct paths) from the fundamental nodes to the root node than that for the case with grouped intermediate variables (27 distinct paths). The hierarchical uncertainty quantification methodology discussed in the following section is based on propagating the uncertainty along each such path, thereby offering

computational advantages by reducing the total number of paths when the variables are grouped in this example.

4.2.2.3 Modulus of continuity

The concept of *modulus of continuity* will play a pivotal role in the hierarchical/modular analysis paradigm.

Definition 4.2.2. *(modulus of continuity) Let* $x_1 \in \mathcal{X}_1 \subseteq \mathbb{R}^{n_1}, \ldots, x_N \in \mathcal{X}_M \subseteq \mathbb{R}^{n_N}$, $x = (x_1, \ldots, x_N)$, $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_N$, *and* $F : \mathcal{X} \to \mathbb{R}^m$. *The modulus of continuity* $\omega_{x_i}(F, \delta, \mathcal{X})$ *of* F *with respect to* x_i *is defined as [16, 60]*

$$\omega_{x_i}(F, \delta, \mathcal{X}) := \sup_{x \in \mathcal{X},\ x^{(i)} \in \mathcal{X},\ \|x - x^{(i)}\| \leq \delta} \|F(x) - F(x^{(i)})\|. \tag{4.90}$$

For notational simplicity, replace $\omega_{x_i}(F, \delta, \mathcal{X})$ *by* $\omega_i(F, \delta)$ *whenever this simplification causes no confusion.* ◁

Remark 4.2.3. *The norms employed in Eq. (4.90) are arbitrary and different choices of norms ostensibly lead to different values for the moduli of continuity. Thus, it's possible to reduce the moduli of continuity by proper choice of norms and consequently to refine the hierarchical uncertainty quantification results. Finally, note that* $F : \mathcal{X} \to \mathbb{R}^m$, $\mathcal{X} \subset \mathbb{R}^{\sum_{i=1}^N n_i}$, *and* m *is not necessarily equal to* $\sum_{i=1}^N n_i$. *Consequently, the two norms used in Eq. (4.90) may be distinct but are notationally indistinguishable.* ◁

Definition 4.2.4. *Let* $\mathcal{X}_1 \subset \mathbb{R}^{n_1}, \ldots, \mathcal{X}_N \subset \mathbb{R}^{n_N}$, $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_N$, $F : \mathcal{X} \to \mathbb{R}^N$,

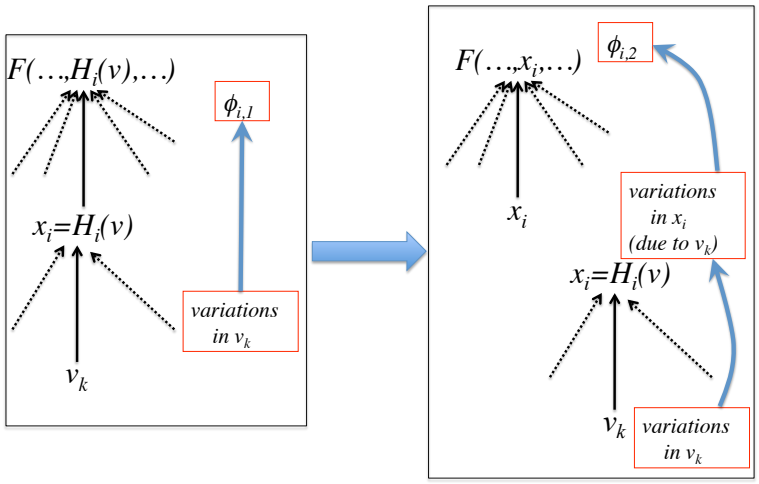


Figure 4.15: Illustration of $\phi_{i,1}$ and $\phi_{i,2}$ in Definition 4.2.4 and Lemma 4.2.5. Lemma 4.2.5 provides a two-step procedure to bound the variations in F due to v_k

$i \in \{1, \ldots, m\}$, and $H_i : \mathcal{V} \subseteq \mathbb{R}^\ell \to \mathcal{X}_i$. Define

$$\phi_{i,1} := \sup_{\substack{x_j \in \mathcal{X}_j,\ 1 \le j \le N,\ j \ne i \\ v \in \mathcal{V},\ v' \in \mathcal{V}}} \|F(x_1, \ldots, x_{i-1}, H_i(v), x_{i+1}, \ldots, x_N) - F(x_1, \ldots, x_{i-1}, H_i(v'), x_{i+1}, \ldots, x_N)\|$$

$$\phi_{i,2} := \sup_{\substack{x_j \in \mathcal{X}_j,\ 1 \le j \le N,\ j \ne i \\ x_i \in H_i(\mathcal{V}),\ (x^{(i)})_i \in H_i(\mathcal{V}) \\ \|x - x^{(i)}\| \le r}} \|F(x) - F(x^{(i)})\|, \tag{4.91}$$

where $H_i(\mathcal{V})$ denotes the image of $\mathcal{V}$ under H_i, and

$$r := \sup_{v \in \mathcal{V},\ v' \in \mathcal{V}} \|H_i(v) - H_i(v')\|.$$

◁

Lemma 4.2.5. *Let $\phi_{i,1}$ and $\phi_{i,2}$ be as defined in Definition 4.2.4. Then, $\phi_{i,1} \leq \phi_{i,2}$.* $\lhd$

Note that in particular v' in the definition of $\phi_{i,1}$ can be chosen in particular as $v' = v^{(k)}$ for some $k \in \{1, \ldots, \ell\}$. In this case, Lemma 4.2.5 provides a two-step procedure to propagate the variations in v_k to the variations in F as illustrated in Fig. 4.15. This two-step procedure is potentially conservative (i.e., $\phi_{i,1} \leq \phi_{i,2}$) but provides a building block for the hierarchical uncertainty propagation framework proposed in the following sections.

Proof. (of Lemma 4.2.5) The set containment

$$\begin{aligned}\big\{(\xi, \xi') \in \mathbb{R}^{2n_i} \ : \ \xi = H_i(v), \ \xi' = H_i(v'), \ v \in \mathcal{V}, \ v' \in \mathcal{V}\big\} \\ \subseteq \big\{(\xi, \xi') \in \mathbb{R}^{2n_i} \ : \ \xi \in H_i(\mathcal{V}), \ \xi' \in H_i(\mathcal{V}), \ \|\xi - \xi'\| \leq r\big\}\end{aligned}$$

implies that the feasible set in the optimization problem in the definition of $\phi_{i,1}$ is a subset of that in the definition of $\phi_{i,2}$ and consequently $\phi_{i,1} \leq \phi_{i,2}$. $\square$

4.2.3 Modular uncertainty propagation over acyclic directed graphs

Consider an acyclic directed graph $G = (V, E)$ and associate each node $i \in V$ with the random variable $X_i \in \mathcal{X}_i \subseteq \mathbb{R}^{n_i}$. Introduce some more notation:

Definition 4.2.6. *Let $X_1, \ldots, X_N$ be a collection of random variables and S be a vector with entries that are in $\{1, \ldots, N\}$ and satisfy $S(i) \leq S(j)$ for $i \leq j$ where $S(i)$ denotes the i-th entry of S. Define $x_V \in \mathbb{R}^{\sum_{i \in V} n_i}$ by*

$$x_V := (x_{V(1)}, \ldots, x_{V(N)}).$$

$\lhd$

For each $i \notin \mathcal{F}$, let $X_{\mathcal{N}_i} = (X_{\mathcal{N}_i(1)}, \ldots, X_{\mathcal{N}_i(M)})$, where M denotes the length (number of elements) of $\mathcal{N}_i$. Then, for each $i \notin \mathcal{F}$ (the set of fundamental nodes in G), let x_i be related to the variables corresponding to the nodes in $\mathcal{N}_i$ through the map $H_i : \mathbb{R}^{\sum_{j \in \mathcal{N}_i} n_j} \to \mathbb{R}^{n_i}$ by

$$X_i = H_i(X_{\mathcal{N}_i}).$$

Furthermore, assume (with no loss of generality) that the entries of $\mathcal{F}$ are in ascending order. Then, the output x_1 is related to the variables corresponding to the fundamental nodes through a map $F : \mathbb{R}^{\sum_{j \in \mathcal{F}} n_j} \to \mathbb{R}$ by

$$X_1 = F(X_{\mathcal{F}}).$$

Recall from McDiarmid's inequality Eq. (2.1) that for given $r \geq 0$,

$$\mathbb{P}[X_1 - \mathbb{E}[X_1] \geq r] \leq \exp\left(-2\frac{r^2}{D_F^2}\right) \tag{4.92}$$

where

$$D_F^2 = \sum_{i \in \mathcal{F}} \quad D_{F,i}^2 = \sum_{i \in \mathcal{F}} \qquad \sup_{x_{\mathcal{F}} \in \mathcal{X}_{\mathcal{F}},\ x_{\mathcal{F}}^{(i)} \in \mathcal{X}_{\mathcal{F}}} \|F(x_{\mathcal{F}}) - F(x_{\mathcal{F}}^{(i)})\|^2$$

and for $i \in \mathcal{F}$, $x_{\mathcal{F}}^{(i)}$ is a shorthand notation for $\left(x^{(i)}\right)_{\mathcal{F}}$ (i.e., the vector obtained when the i-th component of x is perturbed and then the perturbed vector is restricted to the indices in $\mathcal{F}$ as in Definition 4.2.6). The next section exploits the relation $F(X_{\mathcal{F}}) = H_1(X_{\mathcal{N}_1})$ for developing alternative upper bounds on $\mathbb{P}\left[X_1 - \mathbb{E}\left[X_1\right] \geq r\right]$.

4.2.3.1 The main result

An upper bound $\mathfrak{D}_F$ can now be established on D_F that exploits the structure in the dependence of X_1 on the fundamental variables $X_{\mathcal{F}}$ through the intermediate variables captured in the graph G. This upper bound will provide more conservative estimates on the probability-of-failure compared to that in Eq. (4.92), i.e.,

$$\mathbb{P}[X_1 - \mathbb{E}[X_1] \geq r] \leq \exp\left(-2\frac{r^2}{D_F^2}\right) \leq \exp\left(-2\frac{r^2}{\mathfrak{D}_F^2}\right)$$

with a possible reduction of the computational cost as discussed in §4.2.1.

Definition 4.2.7. *For each $i \in \mathcal{F}$,*

$$D_i := \sup_{x_i \in \mathcal{X}_i,\ \ x_i' \in \mathcal{X}_i} \|x_i - x_i'\|.$$

$$\mathfrak{D}_{F,i} := \sum_{\gamma \in \mathcal{P}(i,1)} \omega_{1,\gamma_{L_\gamma - 1}} \circ \omega_{\gamma_{L_\gamma - 1},\gamma_{L_\gamma - 2}} \circ \ldots \circ \omega_{\gamma_2,i}(D_i),$$

where, for $\gamma = (i, \gamma_2, \ldots, \gamma_{L_\gamma - 1}, 1) \in \mathcal{P}(i,1)$, L_γ denotes the length of γ, $\circ$ denotes the composition operator, and for $r \geq 0$ and the consecutive nodes i and j in γ, $\omega_{\gamma_i,\gamma_j}(r)$ is a shorthand notation for $\omega_{X_{\gamma_j}}(H_{\gamma_i}, r, \mathcal{X}_{\gamma_j})$ (the domain $\mathcal{X}_{\gamma_j}$ of the argument X_{γ_j} of H_{γ_i} is omitted).

$$\mathfrak{D}_F^2 := \sum_{i \in \mathcal{F}} \mathfrak{D}_{F,i}^2$$

◁

The main theorem of this section follows.

Theorem 4.2.8. *For D_F and $\mathfrak{D}_F$ as defined above, the inequality $D_F \leq \mathfrak{D}_F$ holds.* ◁

Theorem 4.2.8 provides a way for computing upper bounds on D_F by induction. For

each fundamental node $i \in \mathcal{F}$ with the uncertainty measured by D_i, compute $\omega_{\gamma_2,i}(D_i) = \omega_{X_i}(H_{\gamma_2}, D_i, \mathcal{X}_i)$, and then compute $\omega_{\gamma_3,\gamma_2} \circ \omega_{\gamma_2,i}(D_i) = \omega_{X_{\gamma_2}}(H_{\gamma_3}, \omega_{X_i}(H_{\gamma_2}, D_i, \mathcal{X}_i), \mathcal{X}_{\gamma_2})$, etc. Each step of the inductive procedure requires two types of optimization:

- Once $\mathcal{X}_j$ is known for all $j \in \mathcal{N}_i$ for some $i \in V$, determine the set $\mathcal{X}_i$ such that $H_i(\mathcal{X}_{\mathcal{N}_i}) \subseteq \mathcal{X}_i$.
- Propagate the modulus of continuity. Consider the computation of $\omega_{\gamma_3,\gamma_2} \circ \omega_{\gamma_2,i}(D_i)$. Once $\omega_{\gamma_2,i}(D_i)$ and $\mathcal{X}_{\mathcal{N}_{\gamma_3}}$ are known, $\omega_{\gamma_3,\gamma_2} \circ \omega_{\gamma_2,i}(D_i)$ can be computed through the optimization problem in Eq. (4.90).

A proof of this theorem is detailed in [68].

Remark 4.2.9. *The upper bound $\mathfrak{D}_F$ on D_F depends on the graphical model (i.e., the choice of the intermediate variables and H_is for the nodes that are neither fundamental nor the output). Nevertheless, this dependence will not be explicitly notated. Similarly, $\mathfrak{D}_F$ depends on the choice of the norms used in calculating the moduli of continuity and this dependence will not be explicitly notated. Nevertheless, it is assumed that the same norm is used throughout the procedure.* $\lhd$

Propagating the ranges through the graph model Note that for each intermediate variable X_i, Theorem 4.2.8 requires determining a set $\mathcal{X}_i$ such that $H_i(\mathcal{X}_{\mathcal{N}_i})$. To reduce the conservatism, $\mathcal{X}_i$ should be a "tight" superset of $H_i(\mathcal{X}_{\mathcal{N}_i})$, and for ease of implementation $\mathcal{X}_i$ should have a simple parameterization. Two of the possible strategies for parameterizing $\mathcal{X}_i$ are discussed next.

- Hyper-rectangle: Define $\mathcal{X}_i$ to be a hyper-rectangle of the form

$$\mathcal{X}_i = [a_{i,1}, b_{i,1}] \times \cdots \times [a_{i,n_i}, b_{i,n_i}]$$

$$a_{i,j} = \inf_{\xi \in \mathcal{X}_{\mathcal{N}_i}} \quad H_{i,j}(\xi)$$

$$b_{i,j} = \sup_{\xi \in \mathcal{X}_{\mathcal{N}_i}} \quad H_{i,j}(\xi).$$

  Then, $\mathcal{X}_i$ is the smallest hyper-rectangle containing $H_i(\mathcal{X}_{\mathcal{N}_i})$.

- Ellipsoid: It may be possible to encode *prior information* by choosing a fixed shape ellipsoid and determining the level of this ellipsoid that contains the range of H_i. To this end, let P be a positive semidefinite matrix and $\xi_c \in \mathbb{R}^{n_i}$ and set $\mathcal{X}_i = \mathcal{E}_{t^*} := \{\xi \ : \ (\xi - \xi_c)^T P(\xi - \xi_c) \leq t^*\}$, where

$$t^* := \inf_t \quad t \ \text{ subject to } \quad H_i(X_{\mathcal{N}_i}) \subseteq \mathcal{E}_t$$

  or equivalently

$$t^* = \sup_{v \in \mathcal{X}_{\mathcal{N}_i}} \quad (H_i(v) - \xi_c)^T P \left(H_i(v) - \xi_c\right).$$

Special case: tree of depth 2 Consider the case where G is a tree of depth 2. In this case the set of intermediate nodes $\mathcal{I} \subset V$ is $\mathcal{I} := V \backslash (\{1\} \cup \mathcal{F})$, and x_1 is an explicit function of $X_{\mathcal{F}}$, i.e., $\mathcal{N}_1 = \mathcal{I}$ and

$$X_1 = F(X_{\mathcal{F}}) = H_1(X_{\mathcal{N}_1}) = H_1(X_{\mathcal{I}}).$$

Then, the computation of the upper bound $\mathfrak{D}_F$ on D_F is simplified as in the following proposition.

Proposition 4.2.10.

$$D_F^2 \leq \mathfrak{D}_F^2 := \sum_{i=\mathcal{I}} \sum_{j \in \mathcal{N}_i} \omega_{1,i}(\omega_{i,j}(D_j))^2 = \sum_{j \in \mathcal{N}_i} \sum_{i=\mathcal{I}} \omega_{1,i}(\omega_{i,j}(D_j))^2.$$

◁

Proof. of Proposition 4.2.10. With no loss of generality, assume that $\mathcal{I} = \{2, \ldots, M\}$. Note that for each $j \in \mathcal{F}$ there exists a single $i \in \mathcal{I}$ such that $j \in \mathcal{N}_i$ and

$$D_{F,j}^2 = \sup_{\substack{x_k \in \mathcal{X}_k, \quad k \in \mathcal{I}, k \neq i \\ x_{\mathcal{N}_i} \in \mathcal{X}_{\mathcal{N}_i}, \quad x_{\mathcal{N}_i}^{(j)} \in \mathcal{X}_{N_i}}} |H_1(x_2, \ldots, x_{i-1}, H_i(x_{\mathcal{N}_i}), x_{i+1}, \ldots, x_M) - H_1(x_2, \ldots, x_{i-1}, H_i(x_{\mathcal{N}_i}^{(j)}), x_{i+1}, \ldots, x_M)|^2,$$

and, by Lemma 4.2.5, it follows that $D_{F,j}^2 \leq \omega_{1,i}(\omega_{i,j}(D_j))^2$. Since for each $j \in \mathcal{F}$ there exists a single $i \in \mathcal{I}$ such that $j \in \mathcal{N}_i$, it follows that $D_F^2 \leq \mathfrak{D}_F^2$. □

Repeatedly applying the procedure in the proof, Proposition 4.2.10 can be generalized to the following special case of Theorem 4.2.8 for general tree structure.

$$\mathfrak{D}_F^2 = \sum_{i \in \mathcal{F}} \sum_{\gamma \in \mathcal{P}(i,1)} \left(\omega_{1,\gamma_{L_\gamma - 1}} \circ \omega_{\gamma_{L_\gamma - 1}, \gamma_{L_\gamma - 2}} \circ \ldots \circ \omega_{\gamma_2, i}(D_i) \right)^2$$

The tree structure enables complete parallelization of the computations in the hierarchical uncertainty propagation since the propagation along any path is decoupled from that on any other path (paths with different initial nodes can only share the terminal node).

Special case: simple graph structure When G is a tree, the variables corresponding to the fundamental nodes as well as the intermediate variables are independent. Now

consider a slightly more general case where the intermediate variables are allowed to be correlated because two (or more) different intermediate variables may explicitly depend on the same fundamental variable. Formally, the graph structure that is considered satisfies the properties: (i) $\mathcal{N}_j \in \mathcal{F}$ for each $j \in \mathcal{N}_1$, and $\cap_{j\in\mathcal{N}_1} = \mathcal{F}$. This simple case features the main point of the general case in Theorem 4.2.8 yet with a simpler proof. For each $i \in \mathcal{F}$, let $\mathcal{M}_i \subseteq \mathcal{I}$ denote the set of intermediate variables such that, for each $j \in \mathcal{M}_i$, $i \in \mathcal{N}_j$.

Proposition 4.2.11. *Let $G = (V, E)$ satisfy the assumptions above. Then,*

$$D_F^2 \leq \mathfrak{D}_F^2 := \sum_{i\in\mathcal{F}} \left(\sum_{j\in\mathcal{M}_i} \omega_{1,j} \left(\omega_{j,i}(D_i)\right) \right)^2.$$

◁

Proof. of Proposition 4.2.11. With no loss of generality, let $\mathcal{I} = \{2, \ldots, M\}$ and $\mathcal{F} = \{M+1, \ldots, M+N\}$. Then, for each $i \in \mathcal{F}$

$$\begin{aligned} D_{F,i}^2 &= \sup_{x_\mathcal{F}\in\mathcal{X}_\mathcal{F},\ x_\mathcal{F}^{(i)}\in\mathcal{X}_\mathcal{F}} |F(x_\mathcal{F}) - F(x_\mathcal{F}^{(i)})|^2 \\ &\leq \left\{ \sum_{j\in\mathcal{M}_i} \sup_{x_{\mathcal{N}_j}\in\mathcal{X}_{\mathcal{N}_j},\ x_{\mathcal{N}_j}^{(i)}\in\mathcal{X}_{\mathcal{N}_j}} |H_1(\ldots, H_j(x_{\mathcal{N}_j}), \ldots) - H_1(\ldots, H_j(x_{\mathcal{N}_j}^{(i)}), \ldots)| \right\}^2 \\ &\leq \left(\sum_{j\in\mathcal{M}_i} \omega_{1,j} \left(\omega_{j,i}(D_i)\right) \right)^2, \end{aligned}$$

where the first inequality follows from the triangle inequality and the sub-additivity of supremum and the second inequality follows from Lemma 4.2.5. Consequently,

$$D_F^2 = \sum_{i=1}^N D_{F,i}^2 \leq \sum_{i\in\mathcal{F}} \left(\sum_{j\in\mathcal{M}_i} \omega_{1,j} \left(\omega_{j,i}(D_i)\right) \right)^2 = \mathfrak{D}_F^2.$$

□

4.2.3.2 Algorithmic summary of the methodology

The application of the main result can be summarized in a pseudo-algorithmic form. To this end, let $X_1, \ldots, X_N$ be the random variables in the model such that $X_i = H_i(X)$, where $X = (X_1, \ldots, X_N)$.

1. **Generate the graphical model:**

   (a) Collect the variables in three groups:

   - *Output variable*: The variable that is not an input to any other variable and represents the performance measure of interest.
   - *Fundamental variables*: The variables that are not explicit functions of other variables. The ranges (admissible values) for all fundamental variables must be known.
   - *Intermediate variables*: The variables that are neither output nor fundamental.

   (b) Enumerate the variables such that X_1 is the output variable, $X_2, \ldots, X_{n_{\mathcal{I}}+1}$ are the intermediate variables, and $X_{n_{\mathcal{I}}+2}, \ldots, X_N$ are the fundamental variables.

   (c) Let $V = \{1, \ldots, N\}$ denote the set of indices of the variables $X_1, \ldots, X_N$ and define $G = (V, E)$ to be a graph with the set of edges determined by

   $$(i_1, i_2) \in E \text{ if and only if } H_{i_2} \text{ is an explicit function of } X_{i_1}.$$

2. **Initialize the calculations for the hierarchical analysis:** Find all paths in G with the initial node in $\{n_{\mathcal{I}} + 2, \ldots, N\}$ and terminal node 1 and label them by $\gamma_1, \ldots, \gamma_{N_{path}}$, where N_{path} is the number of such paths. Here, a path in G is defined

$go = 1$
repeat while $go = 1$
- **for** $i = 1, \ldots, N$
  - **if** the ranges for all the variables which H_i is an explicit function of are known
    - determine the smallest hyper-rectangle $\mathcal{X}_i$ that contains the range of H_i
    - **for** $j = 1, \ldots, N_{path}$
      - **if** γ_j contains i

        let k be the index of i in γ_j and compute

$$MC_j(k) = \omega_{i,\gamma_j(k-1)}(MC_j(k-1)).$$

Figure 4.16: The pseudo-code to propagate the ranges and moduli of continuity along the paths from the fundamental nodes to the output node

as in §4.2.2.1. For each $i \in \{1, \ldots, N_{path}\}$, let

$$MC_i(1) = \sup_{\xi \in \mathcal{X}_{\gamma_i(1)},\ \xi' \in \mathcal{X}_{\gamma_i(1)}} \|\xi - \xi'\|,$$

where for a positive integer j, $MC_i(j)$ and $\gamma_i(j)$ denote the j-th entry in MC_i and γ_i, respectively.

3. Execute the pseudo-code in Fig. 4.16 to propagate the ranges and moduli of continuity along the paths from the fundamental nodes to the output node.

4. **Compute the upper bound $\mathfrak{D}_F$ on the verification diameter D_F:** For each $i \in \mathcal{F}$

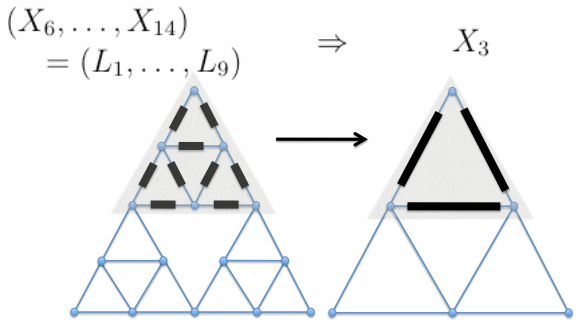
$$\mathfrak{D}_{F,i}^2 = \sum_{j\ :\ \gamma_j(1)=i} MC_j(L_{\gamma_j})^2,$$

where L_{γ_j} is the length of γ_j and

$$\mathfrak{D}_F^2 = \sum_{i \in \mathcal{F}} \mathfrak{D}_{F,i}^2.$$

A more detailed discussion of the application of Theorem 4.2.8 regarding the LC electrical circuit example and the hypothetical running example is in order.

4.2.3.3 Illustration of the main result



(a)

For $i = 6, \ldots, 14$,

$$D_i = |b_i - a_i|,$$

$$\omega_{3,i}(D_9) = \sup_{\substack{\xi = (X_6, \ldots, X_{14}) \in \mathcal{X}_6 \times \cdots \times \mathcal{X}_{14} \\ \xi^{(i)} \in \mathcal{X}_6 \times \cdots \times \mathcal{X}_{14}}} \|X_3(\xi) - X_3(\xi^{(i)})\|$$

Determine $\mathcal{X}_3 \in \mathbb{R}^3$ as described in §4.2.3.1

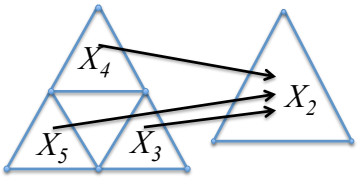
(b)

Figure 4.17: (a) Detailed view of the module from the fundamental variables $X_6, \ldots, X_{14}$ to the variable X_4. (b) Computations for the module shown in part (a) of the figure

The LC electrical circuit example Consider the hierarchical model in Eq. (4.89) and the corresponding graph structure in Fig. 4.14 where the fundamental variables ($i = 6, \ldots, 32$) satisfy

$$X_i \in \mathcal{X}_i = [a_i, b_i] \subset \mathbb{R}.$$

Modules connected to the fundamental variables: These modules correspond to the transitions shown in Fig. 4.9 (and partially repeated in Fig. 4.17(a)). More formally, for a Sierpinski triangle of depth s, the inputs for these modules are the inductances on the depth s circuit and the outputs are the equivalent inductances on the circuit of depth $s-1$.



(a)

For $i = 6, \ldots, 14$,

$$\omega_{2,3} \circ \omega_{3,i}(D_i) = \sup_{\substack{(X_3, X_4, X_5) \in \mathcal{X}_3 \times \mathcal{X}_4 \times \mathcal{X}_5 \\ X_4' \in \mathcal{X}_4 \\ \|X_4 - X_4'\| \leq \omega_{3,i}(D_i)}} \|X_2(X_3, X_4, X_5) - X_2(X_3, X_4', X_5)\|$$

Repeat for X_3 and X_5

Determine $\mathcal{X}_2$ as discussed in §4.2.3.1

(b)

Figure 4.18: (a) Detailed view of the module from the fundamental variables X_3, X_4, and X_5 to the variable X_2. (b) Computations for the module shown in part of (a) of the figure

Fig. 4.17(b) summarizes the computations for the transition shown in Fig. 4.17(a).

Intermediate modules: Fig. 4.18(b) summarizes the computations for the module shown in Fig. 4.18(a) for X_4 (computations for X_3 and X_5 are similar). For Sierpinski triangles of depth larger than 3, the calculations can be repeated until the modules connected to the output node are reached.

The module connected to the output variable: This step requires the calculation of the moduli of continuity of H_1 with respect to X_2 along all the paths connecting the fundamental nodes to the root node 1 passing through node 2.

4.2.3.4 The hypothetical running example

The underlying graphical model in §4.2.3.3 is ostensibly a tree — thus computations (range and moduli of continuity calculations) at each level can be parallelized. The parallelization

of computations of non-tree graphs may be limited, and the order in which the computation is performed is dependent upon the specific graph structure. For the hypothetical running example introduced in §4.2.1.3 (see Fig. 4.11), the computation of the range and all moduli of continuity for the intermediate variable X_6 needs information about ranges of X_7 and X_{12} concurrently. Respecting such requirements, Table 4.2 shows the order in which computations can be executed in parallel.

Table 4.2: The order in which computations can be executed in parallel (i.e., moduli of continuity and range calculations for the respective nodes can be carried out)

step	modules for which computations can be executed in parallel
1	$2, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20$
2	$5, 7, 8, 9$
3	$4, 6$
4	3
5	1

4.2.4 Uncertainty quantification for the electrical circuit example

A demonstration the hierarchical computation of the upper bounds on the verification diameter on the LC electrical circuit example is introduced in §4.2.1.2.

4.2.4.1 Performance measures

When a capacitor of capacitance C and an inductive circuit element with an equivalent inductance L_{eq} are connected together as shown in Fig. 4.5, an electrical current can alternate between them at circuit's (angular) resonant frequency

$$\Omega_r := \frac{1}{\sqrt{CL_{eq}}}.$$

This resonance effect has applications in tuning/antenna technologies (tuning the radio to a particular station can be achieved by adjusting the resonant frequency of a corresponding LC circuit), voltage or current amplification, and induction heating. Now fix a real number $\Omega_t > 0$ and ask the question:

What is a provable upper bound on $\mathbb{P}[\Omega_r \geq \Omega_t]$?

An upper bound on this probability can be obtained using the McDiarmid's inequality as

$$\mathbb{P}[\Omega_r \geq \Omega_t] \leq \exp\left(-2\frac{(\Omega_t - \mathbb{E}[\Omega_r])_+^2}{D_{F_1}^2}\right), \tag{4.93}$$

where $F_1(\xi) = 1/\sqrt{C\xi}$. On the other hand, note that $\Omega_t > 0$ and $\Omega_r \geq \Omega_t$ if and only if $\Omega_r^2 \geq \Omega_t^2$. Consequently,

$$\begin{aligned} \mathbb{P}[\Omega_r \geq \Omega_t] &= \mathbb{P}\left[L_{eq} \leq \tfrac{1}{C\Omega_t^2}\right] \\ &\leq \exp\left(-2\frac{\left(\mathbb{E}[L_{eq}] - \frac{1}{C\Omega_t^2}\right)_+^2}{D_{F_2}^2}\right), \end{aligned} \tag{4.94}$$

where $F_2(\xi) = \xi$.

4.2.4.2 Results

Numerical results for F_1 and F_2 and Sierpinski triangles of depth 2 and 3 are presented next. The ranges $[a_i, b_i]$ of L_i are chosen so that $0.1(b_i + a_i) = b_i - a_i$ (i.e., 10% uncertainty) and the mean equivalent inductance (i.e., the equivalent inductance for the case $L_i = (a_i + b_i)/2$) is equal to 1). Two cases are considered:

1. $a_i = a_j$ and $b_i = b_j$ for all $i, j = 1, \ldots, 3^d$ (homogeneously distributed)

2. $a_{3(i-1)+1} = a_{3(i-1)+2} = a_{3(i-1)+3}$ and $b_{3(i-1)+1} = b_{3(i-1)+2} = b_{3(i-1)+3}$ for all $i = 1, \ldots, 3^{d-1}$ (non-homogeneously distributed),

where d denotes the depth of the Sierpinski triangle. In all calculations, $C = 1$.

The verification diameters for F_1 and F_2, computed using the following procedures, are shown in Figs. 4.19(a) and 4.19(b), respectively:

1. The hierarchical method with grouped intermediate variables as explained in §4.2.2.2 ("+" markers);

2. The hierarchical method with ungrouped intermediate variables ("$\diamond$" markers);

3. Direct McDiarmid calculation, i.e., without using the hierarchical method ("$\circ$" markers).

Expectedly, the verification diameter by direct computation is the smallest (i.e., least conservative) in each case and that with ungrouped intermediate variables is less conservative compared to that with grouped intermediate variables. Recall that grouping the intermediate variables leads to a smaller number of paths along which the uncertainties have to be propagated in the hierarchical analysis. Figs. 4.19(a) and 4.19(b) illustrate that this computational advantage is at the resolution which results in more conservative upper bounds on the probability-of-failure.

Remark 4.2.12. *In §4.2.4.1, upper bounds for the same performance measure, namely* $\mathbb{P}[\Omega_r \geq \Omega_t]$, *are established applying the McDiarmid's inequality on two different functions* F_1 *and* F_2. *This example illustrates an important fact that the upper bounds on the probability-of-failure from the McDiarmid's inequality can be refined using right choice of coordinate transformations (for example* $\xi \to 1/\sqrt{C\xi}$ *in this case). To see this, fix* $\Omega_t = 1.05$.

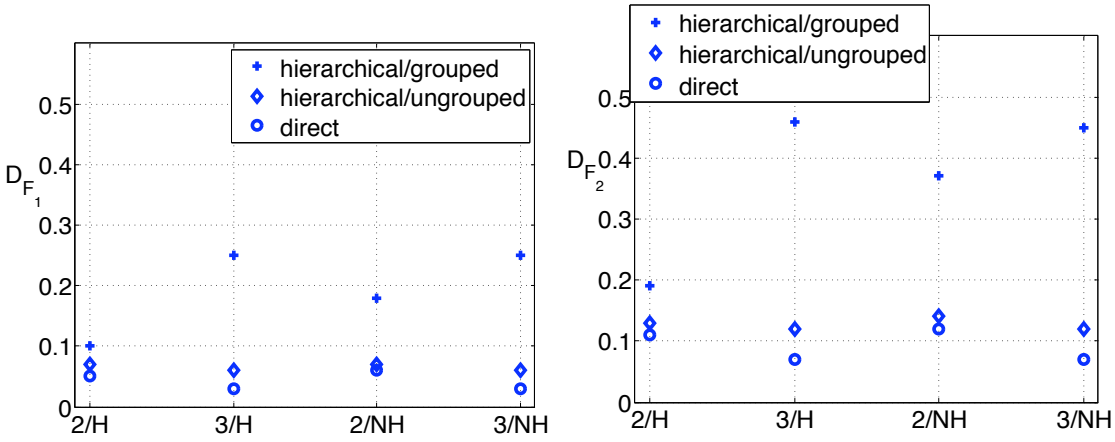


Figure 4.19: The verification diameters for (a) F_1 and (b) F_2. In the horizontal axis, 2 and 3 refer to the depth of the Sierpinski triangle and "H" and "NH" refer to the homogeneous and non-homogeneous distribution of the inductances on the triangle

Then,

$$\mathbb{P}\left[\Omega_r \geq 1.05\right] \leq 0.135$$

by the McDiarmid's inequality applied to F_1 *and*

$$\mathbb{P}\left[\Omega_r \geq 1.05\right] \leq 0.240$$

by the McDiarmid's inequality applied to F_2. *The exploitation of this observation to systematically construct "optimal" coordinate transformations for concentration-of-measure based upper bounds on the probability-of-failure is the subject of ongoing research.* ◁

4.2.4.3 Description of the optimization algorithm

Computing the verification diameters, the moduli of continuity, and the ranges involves solutions to certain nonconvex global optimization problems, e.g., those in Eq. (4.91) and in §4.2.3.1. The solutions of these problems for the results reported in §4.2.4.2 are obtained

using the basic simulated annealing algorithm [26] adapted for appropriately generating random "neighbors" in the feasible set of the optimization problems (defined through the bound and ellipsoid constraints as well as the "disk" constraints of the form $\|x_{current} - x_{neighbor}\| \leq r$ where $x_{current}$ is the current iterate, $x_{neighbor}$ is the next iterate, and r is a real scalar). A detailed description of the algorithm is presented in §A.1. Let T denote the "temperature" and N be the number of function evaluations. A default cooling schedule of $T_{new} = 0.8 \times T^{old}$ with $T_0 = 1.0$ is used. The optimization stops if $T \leq 10^{-8}$, $N > N_{max} = 2000$, or $N_R > 300$ where N_R is the number of successive rejected states. Temperature decrease happens if $N_T > 30$ or $N_S > 20$, i.e., if 30 function evaluations are made or if there are 20 successive accepted optimal states found at the current temperature. The Boltzmann constant is set to 1.0.

4.2.5 Discussion

This modular and hierarchical analysis scheme is especially suitable for systems whose input-output behavior is characterized by multiscale models which have found applications in an increasingly broader range of fields, such as materials science, biology, and power networks, to name a few. It also provides practical advantages by enabling independent analysis of the modules, as well as repeated usage of the uncertainty quantification results for the modules in different context/systems.

A model validation procedure was developed in [37] building on the concentration-of-measure-based probability-of-failure upper bounds. If the actual behavior of a system, characterized by a map G, is modeled by a map F, this validation procedure is based on computing a "validation diameter" for $G - F$. The hierarchical analysis methodology proposed here can be adapted for model validation when G is in a modular form conforming

with that of F and each module of G can be exercised in isolation from the rest of G.

Chapter 5

Conclusions

Certification is a process that seeks to establish whether the probability-of-failure of a system is below an acceptable tolerance. Often, certification is expressed in a language of quantification of *margins* and *uncertainties*, with the understanding that a system is certified if its performance uncertainty is less than its performance margin. Appealing as the QMU conceptual framework is, the precise quantitative definition of uncertainty measures, the precise means by which the uncertainty measures can be determined in practice, be it experimentally, by computer simulation or a combination of both, and the precise manner in which mitigating and controlling uncertainties *guarantees* the safe operation of a system, are often left unspecified.

A method of certification has been developed, predicated upon the use of concentration-of-measure inequalities as a means of *bounding* performance uncertainties. These uncertainty bounds are mathematically rigorous and, therefore, can be taken as a basis for formulating conservative certification criteria. In addition, the approach is unambiguous and supplies precise quantitative definitions for a number of *terms of art* that are often loosely used in *ad hoc* certification methodologies, including the following:

i) *Margins.* When the mean performance of the system is known, the margin in a performance measure is simply the difference between the mean performance and

its threshold for safe operation, Eq. (2.34a). When the mean performance of the system is estimated by way of a model, the margin is the difference between the mean performance and an *increased* threshold, Eq. (2.63a). This threshold increase accounts for uncertainties in the estimation of the mean performance and effectively results in a net loss of margin.

ii) *Uncertainty*. The total uncertainty of the system follows as the sum of an aleatoric uncertainty and an epistemic uncertainty, Eq. (2.82b). The aleatoric uncertainty measures the spread in predicted performance arising from numerical errors, from the statistical variability of the input parameters, the intrinsic stochasticity of the model, or from other sources. The epistemic uncertainty measures the deviation between predicted and observed performance due to the limited fidelity of the model and existence of unknown unknowns.

iii) *Aleatoric uncertainty*. The aleatoric uncertainty in a performance measure is quantified by its verification diameter, Eq. (2.31), i. e., the largest deviation in performance that is computed when each input parameter is allowed to vary in turn between pairs of values spanning its entire range. The aleatoric uncertainty is computed directly from the model without reference to experimental data. It bears emphasis that classical linearized *sensitivity analysis* is not sufficient to quantify aleatoric uncertainties in general. Instead, worst-case scenario *large deviations* in the system input parameters, resulting in likewise large deviations in system performance, must systematically be identified and taken into account. Finally, note that aleatoric uncertainties can be determined on the sole basis of input parameter ranges without precise knowledge of their probability density functions.

iv) *Epistemic uncertainty*. The epistemic uncertainty in a performance measure is quantified by its validation diameter, Eq. (2.76), i. e., the largest deviation in the difference between computed and measured performance that is recorded when each input parameter is allowed to vary in turn between pairs of values spanning its entire range. The determination of the validation diameter is an optimization problem in which the evaluation of the objective function requires the execution of nominally identical calculations and experiments. The algorithm employed in the solution of this optimization problem, be it a stochastic algorithm such as simulated annealing or a genetic algorithm or an iterative algorithm such as a quasi-Newton iteration, determines the precise sequence of calculations and experiments to be performed. It should be noted that, in executing *nominally identical* calculations and experiments, unknown-unknown input parameters are assigned random values within their respective ranges, which adds to the computed modeling uncertainties and systematically and automatically accounts for the effect of unknown unknowns.

v) *Confidence factor*. With the preceding definitions, the confidence factor is simply the quotient of margin to total uncertainty, Eq. (2.33). However, it should be carefully noted that what is specifically asserted through this definition is the mathematical fact that concentration-of-measure inequalities rigorously *guarantee* that a system whose confidence factor is above a well-defined threshold will operate safely within a pre-specified probability-of-failure tolerance. This stands in contrast to QMU methodologies based on *ad hoc* definitions of margins, uncertainties, and confidence factors which, while eminently reasonable and intuitive appealing, may lack a similar mathematical guarantee and, therefore, fail provide a sound basis for certification.

vi) *Aggregation of uncertainties.* In cases in which the safe operation of a system requires multiple performance measures to be above their respective thresholds, certification can still be expressed in terms of an overall confidence factor for the system. In the concentration-of-measure approach to certification, the overall confidence factor follows as a well-defined function of the individual confidence factors of each of the performance measures, Eq. (2.54). It bears emphasis that neither margins nor uncertainties can be aggregated in separation of each other. Instead, individual performance measure confidence factors, which naturally weigh individual margins against their corresponding uncertainties, are to be compounded into an overall confidence factor for the system. In particular, *ad hoc* formulae for aggregating uncertainties, such as root mean square (RMS) formulae, fail to provide a sound basis for certification in general.

It bears emphasis that, in a certification context, the purpose and utility of devising models of the highest possible fidelity is to minimize the number of—presumably costly—tests that are required for certification. *Predictive Science* may then be regarded as the art of formulating such models. Some of the benefits of the concentration-of-measure QMU framework towards the goal of achieving predictive science are worth noting carefully.

The computation of the verification and validation diameters of system requires a global optimization over parameter space. Global optimization algorithms such as simulated annealing and genetic algorithms exhibit a very high degree of concurrency, since at every step of the algorithms large populations of replicas of the model can be evaluated independently. This concurrent evaluation can be accomplished by running a large number of independent jobs on relatively small processor counts ($\sim$500–1000). Bottlenecks should be rare in this mode and exceedingly high efficiencies are expected. Petascale computing capacity

presently under development will enable the running of vast numbers of such jobs simultaneously. Thus, concentration-of-measure uncertainty analysis lends itself ideally to—and provides a potentially important use of—petascale computing.

The calculation of aleatoric and epistemic uncertainties, as measured by the verification and validation diameters of the system, entails a systematic exploration of parameter space, and thus identifies where the critical large-perturbation sensitivities of the system and modeling errors reside. Thus, one important outcome of uncertainty quantification through the computation of the verification and validation diameters is the identification of *bad actors*, i. e., the components of the model, be they numerical, physics models, or otherwise, responsible for the highest system uncertainties. Those model components can then be targeted for refinement through the addition of higher-fidelity physics laws, higher-accuracy numerical algorithms, or by other means. In this manner, QMU systematically guides model development and provides a rational basis for allocating modeling and experimental priorities and resources. Such systematic and prioritized model development is an integral and indispensable part of predictive science.

There is a need to list some of the present limitations of the proposed approach and possible extensions and enhancements thereof. For simplicity, this work has considered input parameters that are independent and supposed that all that is known about the variability in the input parameters is that they lie within certain intervals. However, the concentration-of-measure phenomenon is not limited to the case of independent random inputs and it is possible to devise concentration-of-measure inequalities that account for correlations between inputs (see e.g., Theorems 2.1.4 and 2.1.5). It is also possible to devise concentration-of-measure inequalities that take as input probability density functions of the input variables, and take the resulting inequalities as a basis for formulating conservative

certification criteria (see e.g., Theorem 2.1.5).

An additional concern is whether the concentration-of-measure inequalities supply a sufficiently tight upper bound on the probability-of-failure. The test case of an imploding ring shows that the simplest concentration-of-measure inequalities, namely, those based on Hoeffding's inequality, can significantly overestimate the probability-of-failure. There are a number of alternative concentration-of-measure inequalities, such as Chernoff's inequality, that provide tighter upper bounds on the probability-of-failure. Another possibility is to exploit special characteristics of the system. For instance, a case that often arises in practice concerns systems that are composed of coupled *components* or *subsystems*. Each subsystem can then be modeled separately and an integrated model of the entire system can subsequently be obtained by modeling the coupling between the subsystems. Hierarchies of models in which the subsystems can themselves be recursively decomposed into finer subsystems are also encountered in practice. In these cases, concentration inequalities can be applied recursively in order to bound uncertainties in the integrated system. The uncertainties in the subsystems can be computed through experiments tailored to each subsystem and through numerical simulations based on subsystem models, presumably an easier task than testing and modeling the integrated system itself. The resulting bounds on the subsystem uncertainties exhibit Gaussian tails. The uncertainty of subsystems higher in the hierarchy can then be controlled by means of powerful concentration inequalities for Gaussian random variables. Recall that concentration inequalities having independent Gaussian random variables as input parameters return back Gaussian tail estimates. In this manner, uncertainties can be propagated recursively up the system hierarchy using concentration inequalities for Gaussian random variables.

It should be carefully noted that the present approach requires the determination of

the epistemic uncertainties, as measured by the validation diameter, through the execution of a sequence of identical simulations and experiments. While this aspect of the method may be regarded as a strength—it supplies precise guidelines for model validation and leads to rigorous and conservative certification—it also limits the applicability of the method to systems for which integral tests can be conducted on demand. This raises the question of whether it is possible to extend the present approach—and, more generally, whether rigorous certification is possible at all—when only historical integral data is available and the possibility of acquiring new integral data does not exist. From a mathematical standpoint, the main difficulty is that in general there is no guarantee that historical integral data samples parameter space adequately, especially when certifying new designs, and that performance uncertainties are not underestimated as a result. The situation improves significantly if component testing is possible. Thus, if the system consists of a number of components and interfaces between the components, and if each component and interface can be tested, then it is possible to derive rigorous uncertainty bounds, leading to conservative certification, by methods similar to those outlined in the preceding paragraph.

Finally, a remark on the likely range of applicability of concentration-of-measure inequalities relative to other competing approaches. In cases in which the probability-of-failure is large, sampling methods such as Monte Carlo or quasi-Monte Carlo require a relatively small number of samples and are likely to enjoy a competitive advantage. Sampling methods are also advantageous when large data sets are available or are inexpensive to obtain. However, methods based on direct sampling become impractical if the probability-of-failure is small, i.e., if failure is a rare event, and if sampling is costly. By way of contrast, the effort required for the computation of the verification and validation diameters is independent of the size of the probability-of-failure and concentration-of-measure inequalities can

conveniently be applied—and enjoy a competitive advantage—when probabilities-of-failure are small. Furthermore, concentration-of-measure inequalities are just about the only available certification tools for systems with a large number of inputs whose probability density function is not fully known.

Vita

Leonard Joseph Lucas was born in Pittsburgh, PA (USA), on November 28, 1982, to Leonard D. and Sally A. Lucas. He was raised in Pittsburgh and graduated as valedictorian from Carrick High School in 2000. He attended Carnegie-Mellon University, and in May 2004 graduated with a Bachelor of Science in mechanical engineering; his work included an honors thesis in neuro-robotics, supervised by Prof. Yoky Matsuoka. He commenced graduate study in September 2004 at the California Institute of Technology, and became a Naval Nuclear Propulsion Fellow in September 2005. He has been advised at Caltech by Prof. Michael Ortiz and supervised at the Bettis Atomic Power Laboratory by Dr. David Jones. He is a member of Tau Beta Pi and Pi Tau Sigma.

Bibliography

[1] *AIAA Guide for the Verification and Validation of Computational Fluid Dynamics Simulations*. Number AIAA-G-077-1998. American Institute of Aeronautics & Astronautics, 1998.

[2] *ASME Guide for Verification and Validation in Computational Solid Mechanics*, volume V&V 10-2006. American Society of Mechanical Engineers, 2006.

[3] Simulation-based engineering science: revolutionizing engineering science through simulation. Technical report, National Science Foundation Blue Ribbon Panel on Simulation-based Engineering Science, May 2006. (Available at `http://www.nsf.gov/pubs/reports/sbes_final_report.pdf`).

[4] G. Apostolakis. The concept of probability in safety assessments of technological systems. *Science*, 250:1359–1364, 1990.

[5] W. Arrighetti, P. De Cupis, and G. Gerosa. Circuit models for Sierpinski gasket antennas. 2005. (Available at `arXiv:physics/0510069v1`).

[6] S. Bobkov and M. Ledoux. Poincaré's inequalities and Talagrand's concentration phenomenon for the exponential distribution. *Probab. Theory Rel.*, 107(3):383–400, 1997.

[7] S. Boucheron, O. Bousquet, and G. Lugosi. Concentration inequalities. In O. Bousquet, U. von Luxburg, and G. Rätsch, editors, *Advanced Lectures in Machine Learning*, pages 208–240. Springer, 2004.

[8] S. Boucheron, G. Lugosi, and P. Massart. A sharp concentration inequality with applications. *Random Struct. Algor.*, 16(3):277–292, 2000.

[9] S. Boucheron, G. Lugosi, and P. Massart. Concentration inequalities using the entropy method. *Ann. Probab.*, 31(3):1583–1614, 2003.

[10] R. J. Breeding, J. C. Helton, E. D. Gorham, and F. T. Harper. Summary description of the methods used in the probabilistic risk assessment for nureg-1150. *Nucl. Eng. Des.*, 135(1):1–27, 1992.

[11] F. Cantelli. Sulla probabilita come limita della frequenza. *Rend. Accad. Lincei*, 26(1), 1933.

[12] H. Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Stat.*, 23:493–507, 1952.

[13] A. M. Cuitino, L. Stainier, G. Wang, A. Strachan, T. Cagin, W. A. Goddard III, and M. Ortiz. A multiscale approach for modeling crystalline solids. *J. Comput. Aided Mater. Des.*, 8:127–149, 2001.

[14] J. Dolbow, M. A. Khaleel, and J. Mitchell. Multiscale mathematics initiative: A roadmap. Technical report, U.S. Department of Energy, December 2004. (Available at `http://www.sc.doe.gov/ascr/Research/AM/MultiscaleMathWorkshop3.pdf`).

[15] A. Dvoretzky, J. Kiefer, and J. Wolfowitz. Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Ann. Math. Stat.*, 27:642–669, 1956.

[16] A. V. Efimov. Modulus of continuity. In *Encyclopaedia of Mathematics*. Springer, 2001.

[17] L. C. Evans. An introduction to stochastic differential equations, 2004. http://math.berkeley.edu/∼evans/SDE.course.pdf.

[18] V. Glivenko. Sulla determinazione empirica delle leggi di probabilita. *Giornale dell'Istituta Italiano degli Attuari*, 4, 1933.

[19] M. Gromov. CAT(κ)-spaces: construction and concentration. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 280(Geom. i Topol. 7):100–140, 299–300, 2001.

[20] M. Gromov. Isoperimetry of waists and concentration of maps. *Geom. Funct. Anal.*, 13(1):178–215, 2003.

[21] M. Gromov and V. D. Milman. Brunn theorem and a concentration of volume phenomena for symmetric convex bodies. In *Israel seminar on geometrical aspects of functional analysis (1983/84)*, pages V, 12. Tel Aviv Univ., Tel Aviv, 1984.

[22] L. Gross. Logarithmic Sobolev inequalities. *Am. J. Math.*, 97(4):1061–1083, 1975.

[23] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Am. Stat. Assoc.*, 58:13–30, 1963.

[24] C. Houdré and P. Tetali. Concentration of measure for products of Markov kernels and graph products via functional inequalities. *Combin. Probab. Comput.*, 10(1):1–28, 2001.

[25] A. E. Kennelly. Equivalence of triangles and stars in conducting networks. 34:413–414, 1899.

[26] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.

[27] P. J. M. Laarhoven and E. H. L. Aarts. *Simulated Annealing: Theory and Applications*. Springer, 1987.

[28] D. Eardley (Study Leader). Quantification of margins and uncertainties (QMU). Technical Report JSR-04-330, JASON, The MITRE Corporation (7515 Colshire Drive McLean, Virginia 22102), March 2005.

[29] M. Ledoux. A remark on hypercontractivity and the concentration of measure phenomenon in a compact Riemannian manifold. *Israel J. Math.*, 69(3):361–370, 1990.

[30] M. Ledoux. A heat semigroup approach to concentration on the sphere and on a compact Riemannian manifold. *Geom. Funct. Anal.*, 2(2):221–224, 1992.

[31] M. Ledoux. Concentration of measure and logarithmic Sobolev inequalities. In *Séminaire de Probabilités, XXXIII*, volume 1709 of *Lecture Notes in Math.*, pages 120–216. Springer, Berlin, 1999.

[32] M. Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001.

[33] Paul Lévy. *Problèmes concrets d'analyse fonctionnelle. Avec un complément sur les fonctionnelles analytiques par F. Pellegrino*. Gauthier-Villars, Paris, 1951. 2d ed.

[34] S. Leyendecker, L. J. Lucas, and M. Ortiz. Optimal control strategies for robust certification. 2009.

[35] S. Leyendecker, S. Ober-Blöbaum, J. E. Marsden, and M. Ortiz. Discrete mechanics and optimal control for constrained systems (DMOCC). *Submitted for publication*, 2008.

[36] S. Leyendecker, S. Ober-Blöbaum, J. E. Marsden, and M. Ortiz. Discrete mechanics and optimal control for constrained multibody dynamics. In *Proceedings of the 6th International Conference on Multibody Systems, Nonlinear Dynamics, and Control, ASME International Design Engineering Technical Conferences*, Las Vegas, Nevada, 4-7 September 2007.

[37] L. J. Lucas, H. Owhadi, and M. Ortiz. Rigorous verification, validation, uncertainty quantification and certification through concentration-of-measure inequalities. *Comput. Meth. Appl. Mech. Eng.*, 197:4591–4609, 2008.

[38] G. Lugosi. Concentration-of-measure inequalities, 2006. (Available at http://www.econ.upf.edu/∼lugosi/anu.pdf).

[39] K. Marton. A measure concentration inequality for contracting Markov chains. *Geom. Funct. Anal.*, 6(3):556–571, 1996.

[40] K. Marton. Measure concentration for a class of random processes. *Probab. Theory Rel.*, 110(3):427–439, 1998.

[41] P. Massart. About the constants in Talagrand's concentration inequalities for empirical processes. *Ann. Probab.*, 28(2):863–884, 2000.

[42] P. Massart. Some applications of concentration inequalities to statistics. *Ann. Fac. Sci. Toulouse*, 9(2):245–303, 2000. Probability theory.

[43] 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.

[44] C. McDiarmid. Centering sequences with bounded differences. *Comb. Probab. Comput.*, 6(1):79–86, 1997.

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

[46] V. D. Milman. Asymptotic properties of functions of several variables that are defined on homogeneous spaces. *Dokl. Akad. Nauk SSSR*, 199:1247–1250, 1971.

[47] V. D. Milman. A certain property of functions defined on infinite-dimensional manifolds. *Dokl. Akad. Nauk SSSR*, 200:781–784, 1971.

[48] V. D. Milman. Geometric theory of Banach spaces. II. Geometry of the unit ball. *Usp. Matem. Nauk*, 26(6(162)):73–149, 1971.

[49] V. D. Milman. A new proof of A. Dvoretzky's theorem on cross-sections of convex bodies. *Funkts. Anal. Prilozh.*, 5(4):28–37, 1971.

[50] M. Mitchell. *An Introduction to Genetic Algorithms.* The MIT Press, 1998.

[51] W. L. Oberkampf and T. G. Trucano. Verification and validation in computational fluid dynamics. *Progr. Aero. Sci.*, 38:209–272, 2002.

[52] W. L. Oberkampf and T. G. Trucano. Verification and validation benchmarks. Technical report, Sandia National Laboratory, 2007. Sandia Report SAND2007-0853.

[53] W. L. Oberkampf, T. G. Trucano, and C. Hirsch. Verification, validation and predictive capability in computational engineering and physics. *Appl. Mech. Rev.*, 57(5):345–384, 2004.

[54] M. Pilch, T. G. Trucano, and J. C. Helton. Ideas underlying quantification of margins and uncertainties (QMU): A white paper. Unlimited Release SAND2006-5001, Sandia National Laboratory, Albuquerque, New Mexico and Livermore, California, September 2006.

[55] J. C. Refsgaard and H. J. Henriksen. Modelling guidelines — terminology and guiding principles. *Adv. Water Resour.*, 27(1):71–82, January 2004.

[56] W. T. Rhee and M. Talagrand. A concentration inequality for the K-median problem. *Math. Oper. Res.*, 14(2):189–202, 1989.

[57] P.-M. Samson. Concentration of measure inequalities for Markov chains and ϕ-mixing processes. *Ann. Prob.*, 28:416–461, 2000.

[58] D. H. Sharp and M. M. Wood-Schultz. QMU and nuclear weapons certification: What's under the hood. *Los Alamos Science*, (28):47–53, 2003.

[59] N. V. Smirnov. Approximate laws of distribution of random variables from empirical data. *Usp. Matem. Nauk*, 10:179–206, 1944.

[60] K.-G. Steffens. *The history of approximation theory.* Birkhauser, 2006.

[61] M. Talagrand. A new isoperimetric inequality and the concentration of measure phenomenon. In *Geometric aspects of functional analysis (1989–90)*, volume 1469 of *Lecture Notes in Math.*, pages 94–124. Springer, Berlin, 1991.

[62] M. Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. *Inst. Hautes Études Sci. Publ. Math.*, (81):73–205, 1995.

[63] M. Talagrand. Nouvelles inégalités de concentration à q points. *CR. Acad. Sci. I-Math.*, 321(11):1505–1507, 1995.

[64] M. Talagrand. Nouvelles inégalités de concentration "convexifiées". *C. R. Acad. Sci. Paris Sér. I Math.*, 321(10):1367–1370, 1995.

[65] M. Talagrand. New concentration inequalities in product spaces. *Invent. Math.*, 126(3):505–563, 1996.

[66] M. Talagrand. A new look at independence. *Ann. Probab.*, 24(1):1–34, 1996.

[67] M. Talagrand. Concentration and influences. *Israel J. Math.*, 111:275–284, 1999.

[68] U. Topcu, L. J. Lucas, M. Ortiz, and H. Owhadi. Uncertainty quantification for modular and hierarchical models. Submitted to Society for Industrial and Applied Mathematics Journal on Multiscale Modeling and Simulation, February 2009.

[69] T. G. Trucano, L. P. Swiler, T. Igusa, W. L. Oberkampf, and M. Pilch. Calibration, validation, and sensitivity analysis: What's what. *Reliab. Eng. Syst. Saf.*, 91(10–11):1331–1357, 2006.

Appendix A

Optimization Algorithms

In the CoM framework of UQ, computation of a diameter, i.e.,

$$D_F^2 = \sum_{i=1}^{N} \sup_{x,x_i'} |(F(x) - F(x'))|^2 \tag{A.1}$$

requires that a global optimization must be done for the sup. Typically a global optimization method will require intrusive adaptation for specific features of this framework. Here, x may be a vector-valued set of N random variables and $x' = (x_1, \ldots, x_i', \ldots, x_N)$ where x_i' is an independent copy of x_i. $F(x)$ is a numerical model that attempts to accurately represent $G(x, z)$, the real-world system that is affected by unknown unknown parameters z (and z') that cannot be accounted for in a numerical model. For example, in the PSAAP setting, x is the projectile velocity, $F(x)$ is a numerical model (e.g., Abaqus) and $G(x, z)$ is the hypervelocity impact experiment. The diameter finds the parameters x and x_i' for which a system differs the most over the entire range of x and the pairwise x_i and x_i', a direct translation of how much "uncertainty" resides in the system based on the randomness introduced by the random variables.

§A.1 describes a simulated annealing algorithm used for the robotic controls and multi-scale UQ applications. §A.2 illustrates an algorithm used for the PSAAP validation appli-

cation.

A.1 Simulated Annealing

A.1.1 Parameters

The main simulated annealing algorithm is that described in [26], but with a few modifications to suit the needs for robotic control and multiscale UQ applications.

Options:

Value [expression]: Description [default value]

1. Cooling schedule [$T^{\text{new}} = T(T^{\text{old}})$]: Any function relating T^{old} to T^{new} can be supplied that generally decreases T at any rate. [$0.8 \times T^{\text{old}}$]

2. Neighbor generator [$x^{\text{new}} = Neighbor(x)$]: Any function that creates a neighbor state x^{new} locally from x.

3. Initial temperature [T_0]: Starting temperature of optimization. [1]

4. Stopping temperature [T^{stop}]: Minimum temperature before ending optimization. [$1.0e-8$]

5. Maximum successive successful attempts [$N_S^{\max}$]: Number of successive neighbor acceptances before decreasing temperature. [20]

6. Maximum function evaluations [$N^{\max}$]: Overall number of function evaluations before exiting optimization. [1000]

7. Maximum consecutive rejections [N_R]: Number of consecutive rejections of new state before exiting optimization. [300]

8. Maximum function evaluations per temperature [$N_T^{\max}$]: number of function evaluations per temperature before decreasing temperature. [$N^{\max}/30$]

9. Boltzmann constant [B]: Constant in probability expression. [1.0]

10. Supremum/infimum: Allows control to perform either supremum or infimum optimization. [supremum]

11. Stopping value [y^{stop}]: State at which to end optimization if a better state is found. [supremum:∞, infimum:$-\infty$]

A.1.2 Constraints

Not only do bound constraints need to be enforced

$$x^{\text{new}} : x \in \mathcal{X} = \mathcal{X}_1 \times \ldots \times \mathcal{X}_N = [a_1, b_1] \times \ldots \times [a_N, b_N] \tag{A.2}$$

but also disk constraints. Given a hyper-rectangle $x_d \in \mathcal{X}_d \subset \mathbb{R}^d \subset \mathcal{X}$, for $d \leq N$ and some $r > 0$, then x^{new} is "disk constrained" according to

$$x^{\text{new}} : \|x_d - x_d^{\text{new}}\| \leq r. \tag{A.3}$$

There may also be "ellipsoid constraints" which are currently only defined for use over all $x \in \mathcal{X}$. Constrain a generated point to remain about an ellipsoid centered at x_c for a new state x^{new} such that

$$x^{\text{new}} : (x_c - x^{\text{new}})^T P (x_c - x^{\text{new}}) \leq 1 \tag{A.4}$$

where P is a positive semidefinite matrix. These last two constraints are justified in §4.2.3.1 and employed in §4.2.4.

The neighbor-finding routine intelligently seeks out neighboring states that assert compliance of these constraints to find a new neighbor for all permutations of bound, disk, or ellipsoid constraints efficiently. The following list denotes how these constraints are enforced independently when a neighbor state is generated:

- Bound:

  1. For each $i = \{1 \dots N\}$, given the bounds $x_i \in \mathcal{X}_i = [a_i, b_i]$, calculate $\sigma_{x_i} = \frac{b_i - a_i}{\sqrt{12}}$ (according to uniform distribution).
  2. Generate an offset from x_i as $x_i^{\text{offset}} = T\alpha$ to recover $x_i^{\text{new}*} = x_i + x_i^{\text{offset}}$ where T is the current temperature and $\alpha \sim \mathcal{N}(0, \sigma_{x_i})$.
  3. Move back into $\mathcal{X}_i$ from $x_i^{\text{new}*}$ through $x_i^{\text{new}} = x_i^{\text{new}*} + \beta T(b_i - a_i)$ towards the interior of $\mathcal{X}_i$ from the exceeded bound where $\beta \sim \mathcal{U}[0, 1]$.

- Disk:

  1. Identify the random variables $D \subset \{i \dots N\}$ participating in the disk constraint, $\|x_D - x_D^{\text{new}}\| \leq r$.
  2. Generate a new point $x^{\text{new}*}$, where the new unconstrained points are generated $x_i^{\text{new}*} \sim T\mathcal{N}(0, 1) \ \forall i \notin D$.
  3. Generate a random direction in $s \in \mathbb{R}^D$.
  4. Transform $x_D^{\text{new}*}$ to x_D^{new} through $x_D^{\text{new}} = x_D + r\frac{\left|x_D - x_D^{\text{new}*}\right|}{\left\|x_D - x_D^{\text{new}*}\right\|}s$.

- Ellipsoid:

  1. Generate a random direction in $s \in \mathbb{R}^N$.
  2. Given the positive semidefinite $\mathbb{R}^{N \times N}$ matrix P, and the center of an ellipse x_c, identify the coefficients $a = s^T P s$, $b = -2s^T P(x_c - x)$, and $c = (x_c - x)^T P(x_c -$

$x) - 1$ of the quadratic equation $at^2 + bt + c = 0$ which identifies distances along s about x that intersect the ellipse.

3. Take $t_1 = \frac{-b+\sqrt{b^2-4ac}}{2a}$ and $t_2 = \frac{-b-\sqrt{b^2-4ac}}{2a}$.

4. Advance x through $x^{\text{new}} = x + s(\alpha t_1 + (1-\alpha)t_2)$ for $\alpha \sim \mathcal{U}[0,1]$.

The enforcement of multiple constraints at a time intelligently suggests a new neighbor point based on the most limiting constraint that is violated for each x_i on an individual basis.

A.1.3 Algorithm

The meta-algorithm outline below is the simulated annealing scheme in [26] that iteratively computes the minimum of a multi-dimensional function.

1. Supply initial state $x = x^{\text{best}} = x_0$, compute $y = y^{\text{best}} = y_0$.

2. Set $T = T_0$.

3. Set number of evaluations $n = 1$ and $n_T = 0$ and $n_S = 0$.

4. Set $finished = false$.

5. While($finished == false$).

   (a) Create new neighbor $x^{\text{new}} = Neighbor(x)$.

   (b) Evaluate $y^{\text{new}} = y(x^{\text{new}})$.

   (c) If y^{new} better than y^{stop}, set $y^{\text{best}} = y^{\text{new}}$ and $x^{\text{best}} = x^{\text{new}}$, set $finished = true$.

   (d) If y^{new} better than y^{best}, set $y = y^{\text{best}} = y^{\text{new}}$ and $x = x^{\text{best}} = x^{\text{new}}$ and $n_S = n_S + 1$.

(e) If $\exp\left(\frac{\{y-y^{\text{new}}\}}{B*T}\right) > u \sim \mathcal{U}[0,1]$, $y = y^{\text{new}}$ and $x = x^{\text{new}}$ and $n_S = n_S + 1$.

(f) If $T <= T^{\text{stop}}$ or $N >= N^{\max}$ $N_R >= N_R^{\max}$, set $finished = true$.

(g) If $n_T >= N_T^{\max}$ or $n_S >= N_S^{\max}$, update temperature $T^{\text{new}} = T(T)$ and set $n_S = n_T = 0$.

(h) Increment function evaluations $n = n + 1$ and $n_T = n_T + 1$.

6. Return best solution y^{best}.

A.2 Global Optimization for Noisy Functions

A.2.1 General setup

Let f be a C^3 function in a domain $A \subset \mathbb{R}^d$. To estimate $\nabla f(x_0)$ for $x_0 \in A$ which in turn helps us to find the extrema of our function. Assume for $x \in A$ there's no direct access to $f(x)$ but the noisy representation $f(x) + Y(x, \omega)$ where $Y(x, \omega)$ is a $\mathcal{N}(0, \sigma(x))$ random variable. Assume that σ is constant (or slowly varying with respect to x).

A.2.2 Algorithm

First assume $f(a_1, ..., a_d)$ where $A = (a_1, ..., a_d)$ and σ are given.

1. Choose a random direction in A and sample 4 points to fit in a degree 3 polynomial. Calculate the third degree partial derivatives as they are described in §A.2.5.

2. Calculate the distance from point A using the formula in section §A.2.3 $\left(r = \sqrt[3]{\frac{3\sigma}{2f_x'''}}\right)$.

3. Try to sample function values at points which have the distance calculated in step 2 from A. This requires $2d$ number of points in order to account for the noise.

4. Using the points from step 3, perform a least-squares method for the following function to find the optimal gradient: $g(\nabla f(A)) = \sum_k [f(x_k) - f(A) - \sum_{i=1}^{d} (\frac{\partial f}{\partial x_i}(A)) x_{k_i}]$ where $x_k = (x_{k_i})_{i=1}^{d}$ are the points.

5. Use the gradient in last step and current step (e.g., conjugate gradient method) to find a new point. In the first step there is no gradient information available, so use $2r$ as the length of the step where r is found in step 2.

6. Iterate and check for convergence.

A.2.3 Optimal points in one dimension

Consider the problem in case of $d = 1$. Using Taylor expansion, write

$$\delta = f'(x) - \frac{f(x+r) - f(x-r)}{2r} = \frac{1}{2r}\left[Y(x+r) + Y(x-r) + \frac{1}{3}r^3 f'''(x)\right] \tag{A.5}$$

then minimize

$$\mathbb{E}\left[\delta^2\right] = \frac{1}{4r^2}[2\sigma^2 + \frac{1}{9}r^6 f'''(x)^2]. \tag{A.6}$$

Manipulation of Eq. (A.6) results in $r = \sqrt[3]{\frac{3\sigma}{2f'''(x)}}$.

A.2.4 Higher dimensions

In higher dimensions the number of terms in the equations increase very fast according to the curse of dimensionality. For example, even in two dimensions an attempt to use only 3 points will produce about 60 terms in a rational function of 6 variable and of degree 6 in numerator and degree 4 in denominator. This is mainly because the method requires

Hessian and all third-order derivative information. Therefore, some simplifications are in order.

The first trick in one dimension is to consider opposite points in the same direction which cancels any Hessian information. Thus, one needs d independent directions and $2d$ points.

The second step to simplify the equations is to reduce the number of variables. If one can assume that the response function is well behaved or has a similar behavior in any d independent directions then choose them as parallel to the Cartesian axes.

The third step is to assume that all third-order derivatives have the same magnitude, which is highly dependent upon the response function.

The demonstration of this method in two dimensions begins with the gradient calculation and the assumption that the initialization point is the origin. Take Δx and Δy as permissible perturbations to x and y.

$$x_0 = \begin{bmatrix} 0 \\ \\ 0 \end{bmatrix} \tag{A.7}$$

The perturbed points about x_0

$$x_0^+ = \begin{bmatrix} \Delta x \\ \\ 0 \end{bmatrix}, x_0^- = \begin{bmatrix} -\Delta x \\ \\ 0 \end{bmatrix}, y_0^+ = \begin{bmatrix} 0 \\ \\ \Delta y \end{bmatrix} y_0^- = \begin{bmatrix} 0 \\ \\ -\Delta y \end{bmatrix} \tag{A.8}$$

are used to obtain the following equations (again using the Taylor expansion):

$$\begin{aligned}\delta &= \left[\nabla f(x_0) - \nabla \hat{f}(x_0)\right]^T \\ &= \left[-f(x_0^+) - \frac{1}{3}\Delta x f_x''' + f(x_0^-) - f(y_0^+) - \frac{1}{3}\Delta y f_y''' + f(y_0^-)\right] \begin{bmatrix} 2\Delta x & 0 \\ 0 & 2\Delta y \end{bmatrix}^{-1}\end{aligned} \tag{A.9}$$

where f_{x_1}''' and f_{x_2}''' are the partial third derivatives of f with respect to the first and second coordinates at origin. The next section will discuss the measurement the magnitude of these parameters. Minimizing $E\left[\delta^2\right]$ results in

$$\Delta x = \sqrt[3]{\frac{3\sigma}{2f_x'''}}, \qquad \Delta y = \sqrt[3]{\frac{3\sigma}{2f_y'''}}. \tag{A.10}$$

Thus the optimal distance is given by the coordinates in Eq. (A.10) in each direction.

A.2.5 Magnitude of the third-order partial derivatives

To find the magnitude of third-order partial derivatives, assume that the response function is approximately a polynomial of degree three. Choose four points in a random direction such that one can fit a polynomial of degree three and consequently find the third partial derivative representative of the response function. Again, consider the case for dimension two. Assume that one needs to project the response function onto the direction $[rcos\theta \;\; rsin\theta]$ where θ is constant. A general third-degree polynomial in two variables x and y looks like

$$P(x, y) = a_{3,0}x^3 + a_{2,1}x^2y + \ldots. \tag{A.11}$$

By substituting the above coordinates one obtains

$$P(r) = a_{3,0}cos^3\theta r^3 + a_{2,1}cos^2\theta sin\theta r^3 + \dots. \tag{A.12}$$

Therefore, to approximate $f'''(x, y)$ one calculates $P'''(r)$.

Appendix B

PSAAP Experimental Hypervelocity Perforation Area Measurement

The following is a transcription of the procedure and results for measuring the perforated areas of the 0.105 in steel plates used for the PSAAP example in §3.4.

- An Olympus SZ61 model microscope equipped with a 5.0 megapixel digital camera by Soft Imaging System (Olympus) was used to capture the images.
- A diffuse light source was placed on the microscope bench.
- To ensure normality with respect to the microscope bench plane, equal height props were used to hold sample plate approximately 2.75 in above the microscope bench (for clearance between light source and sample plate).
- The microscope was focused onto the undeformed entry/exit plane of the plate.
- A ruler (general) with spacing of 1/100 in was placed near the area of interest.
- The image was captured.
- Processing of image was performed with ImageJ from NIH.

- First a scale was set by drawing a line corresponding to 0.05 in on the ruler.
- The image was digitally magnified.
- The perforated area contour was traced with the Pencil tool. The image remains focused on a portion of the undeformed plane shown by the yellow rectangle. The vertical yellow line is used to define scale (i.e., 420 pixels per 0.05 in). The red arrow points to the traced contour of the perforated area.

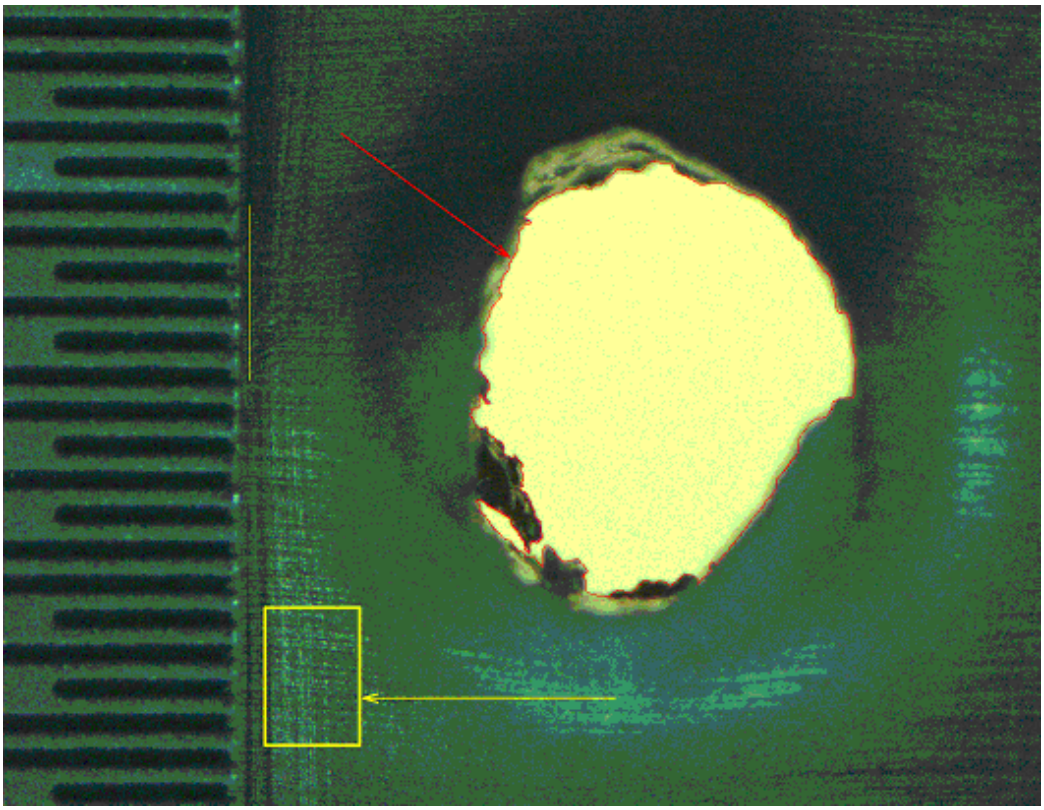


Figure B.1: Experimental photograph showing the measured perimeter of the resulting perforation area

- A mask of the selected contour was generated and the black pixels were counted, giving the area of the perforated area.
- Main Sources of Error:

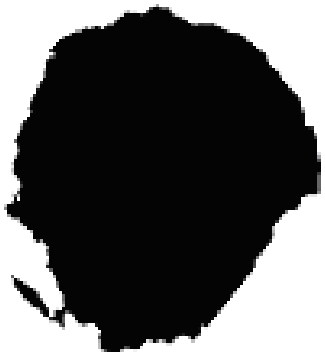


Figure B.2: Mask of the selected contour to count black pixels

- Setting the scale on the image by drawing a line along the ruler.
- Focusing on the undeformed plane for repeatability prohibits focusing and capturing important boundary information above and below the focal plane, such as lipping and spalling.
- An upper bound and lower bound for area were obtained from this image by taking the innermost possible contour and the outermost possible contour. The resulting areas from the two methods, keeping all other parameters (i.e., scale) constant are 3.73(9) mm^2 (outermost) and 3.26(6) mm^2 (innermost). This represents the worst possible image sample, as a large chunk of material (spalled metal) is not even considered in one of the measurements. The error percent between the two is 14.5.

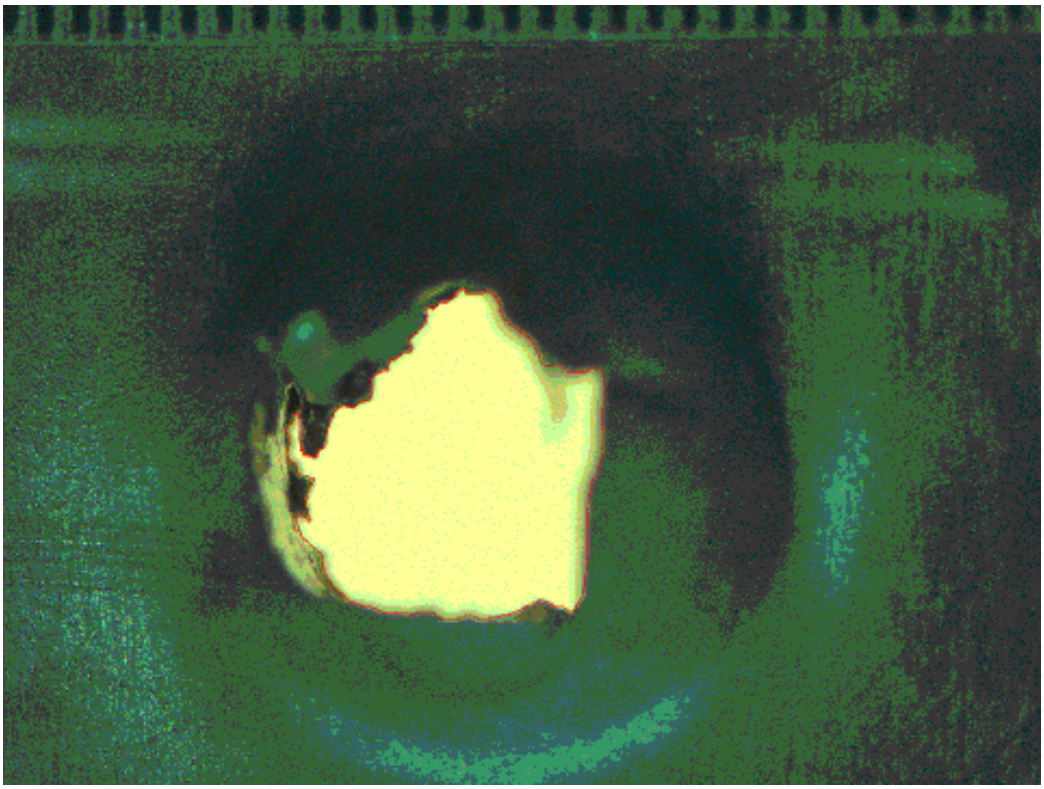


Figure B.3: The contour does not capture the piece of spalled metal on the top right of the perforated area. The contour in the image represents the outer-most possible contour

Appendix C

PSAAP Validation Computation Checklist

C.1 PSAAP Numerical Algorithm Checklist

This section explicitly outlines the procedure for a validation diameter calculation according to those listed in §3.4.3 as entry number 2 in that section's list, referring to a subdiameter computation for controllable variables according to Eq. (4.26) and uncontrollable parameters according to Eq. (4.24).

It is not known *a priori* whether the initialization point x_0 in the algorithm will be near a global or even local optimum for the PSAAP application, especially when considering that the large scatter in the experimental data may distort the search direction and magnitude of the quasi-Newton method. For these issues, a version of quasi-Newton with BFGS updates accounting for noise will be used, which samples optimally offset points from x_k to eliminate the effect of noise based on $f'''(x_k)$ (or simply $f'''(x_0)$, assuming $f'''(x_0)$ is a good representation of every $f'''(x_k)$). It is the hope that the noise does not significantly affect the travel towards a local optimum, but this may very well be the case. The noise may even allow the algorithm to "jump" towards a global optimum without necessarily converging towards the local optimum, but lots of jumping around will severely inhibit convergence in a reason-

able number of steps. For algorithmic expediency, assume that the noise associated with the optimization objective function (not necessarily just the experimental measurements) is nearly Gaussian, with variance σ^2 at all points in the design space. Thus $\sigma(x) \approx \sigma$, which needs to be computed in the initialization step.

The PSAAP application meta-algorithm for validation computations, taking $M = 2$ for the controllable variables $(\alpha, h) = (X_1, X_2)$ and uncontrollable variable $v = X_N$ follows:

- Compute diameters of controllable (α, h) variables.
- Compute σ from historical data.
    1. Use any existing data, and assuming it to be Gaussian, calculate σ from these points.
- Let the space of variables be $X = \alpha \times H = X_1 \times X_2 \subset \mathbb{R}^M$ for $\alpha \in \alpha, h \in H$.
    1. For $i = 1..M$.
    2. Set $X = X \times X_i$.
    3. Let $f(\bar{x} = (\ldots, x_i, x_i', \ldots)) = |G(\ldots, x_i, \ldots) - \ldots$

       $\ldots F(\ldots, x_i, \ldots) - G(\ldots, x_i', \ldots) + F(\ldots, x_i', \ldots)|^2$ (a very noisy function!). Let $\bar{x}$ replace references to x below.

        (a) Set $k = 0$.

        (b) Choose x_0.

        (c) Normalize the parameter space to the dimensionless space $\bar{X} = [0, 1]^{M+1}$.

        (d) Normalize x_0 to be within $\bar{X}$.

        (e) Set $B_0^{-1} = I$. **OR**

        (f) Compute B_0^{-1}.

i. For the ith diagonal entry in B_0, perturb x_0 along the ith direction with two points and compute the quadratic coefficient a of the unique curve through these points. Set $\lambda_i = 2a$.

ii. Do *3(f)i* for each variable.

iii. Compute B_0^{-1} from $B_0 = diag(\lambda)$.

(g) Compute $f'''(x_0)$ according noisy optimization method.

i. Obtain a random direction x_{rand} in $\bar{X}$.

ii. Add x_{rand} to x_0 such that the middle of x_{rand} is at x_0, thus creating x_{slice}.

iii. Shrink x_{slice} so that x_{slice} is within $\bar{X}$.

iv. Choose three random points along x_{slice}, and compute their associated values .

v. Fit a degree 3 polynomial through the three points in *3(g)iv* above along with x_0 and recover the cubed coefficient as $f'''(x_0)$.

(h) Compute search radius $r = \max\left(\sqrt[3]{\frac{3\sigma}{2|f'''(x_0)|}}, \sigma\right)$ that will be used to optimally sample points around x_k with respect to the noise, assuming $f'''(x_0) \sim f'''(x) \forall x$.

(i) Compute $\nabla f(x_0)$ using perturbations about x_0.

i. Perturb x_0 about its M orthogonal directions to compute $\Delta x_{0_i} = x_0 - x'_{0_i}$, $\forall i \in M$.

ii. Compute $\Delta f_{0_i} = f(x_0) - f(x'_{0_i})$, $\forall i \in M$.

iii. Compute $\nabla f(x_0)$ as the least-squares solution to $\Delta x_0 \nabla f(x_0) = \Delta f_0$.

(j) Set $\alpha_k = 1$ (this may be found via line search in later implementations) .

(k) Compute $\Delta x_k = -\alpha_k B_k^{-1} \nabla f(x_k)$.

(l) Compute $x_{k+1} = x_k + \Delta x_k$.

(m) Compute $\nabla f(x_{k+1})$ using perturbations about x_{k+1} .

i. Perturb x_{k+1} about it's M orthogonal directions to compute $\Delta x_{k+1_i} = x_{k+1} - x'_{k+1_i}, \forall i \in M$.

ii. Compute $\Delta f_{k+1_i} = f(x_{k+1}) - f(x'_{k+1_i}), \forall i \in M$.

iii. Compute $\nabla f(x_{k+1})$ as the least-squares solution to $\Delta x_k \nabla f(x_{k+1}) = \Delta f_{k+1}$.

(n) Set $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$.

(o) Compute B_{k+1}^{-1}.

(p) Set $k = k + 1$.

(q) Go to *3i* or until "convergence".

4. Recover $\bar{D}^2_{G-F,i}$.

- Compute diameter of uncontrollable variable v.

- Let the space of variables be $X = \alpha \times H \times V = X_1 \times X_2 \times X_3 \subset \mathbb{R}^{M+1}$ for $\alpha \in \alpha$, $h \in H, v \in V$.

1. Set $X = X \times V$.

2. Let m=(# of samples) and $f(x, z) = \sup_{X_1, X_2} |G(x, z) - F(x) - \ldots$ $\ldots \sum_{j=1}^{m} (G(x_j, z_j) - F(x_j)/m)|^2$ (a very noisy function!).

3. Do steps *3a–3p* above.

4. Recover $\bar{D}^2_{G-F,N}$.

5. Compute $\bar{D}^2_{G-F} = \bar{D}^2_{G-F,N} + \sum_{i=1}^{N-1} \bar{D}^2_{G-F,i}$.

C.2 PSAAP Experimentalist Checklist

The numerical side is physically separated from the experiments that must be done to compute each $G - F$, thus the experimentalists will seemingly be given random values at which to set their random variables, but will be ultimately guided and informed by the algorithm above. The keen experimentalist will note that there will be sampling points perturbed about a nominal x_k in order to compute the gradient to move to the next nominal point x_{k+1} where the Hessian will be updated and the process will continue. Firstly an initialization point must be chosen and sampled about before entering the quasi-Newton iterations. After G is computed in each step, F will be executed for the value of velocity recorded from that experiment, and using the controllable parameters. Thus numerical simulations cannot be performed in parallel with experiments.

Given that historical data is available before any new shots will be done, the computationalists will perform a statistical analysis of this data to compute σ, the standard deviation, which will guide the optimization routine in best selecting new points knowing that there is scatter in the data. This is done before any new shots.

Start a validation loop with three variables: $x = [V, \alpha, h]$ where V is nominal projectile velocity, α is plate tilt, and h is plate thickness (discrete variable).

The checklist for experimentalists:

- Compute diameters of controllable (α, h) variables according to Eq. (4.25). (Note that a V will be set to the nominal velocity V_0 for the duration of this diameter computation since the range of V will be given by the fluctuations of V around V_0.)

    1. Let computationalists compute σ from historical data

    2. For $i = 1..M$ do *3* below

3. Take $x' = x_i$ (an independent copy of x_i)

   (a) Initialization

      i. Computationalists choose a nominal $x_0 = [V, \alpha_0, h_0]$ and x_0' whereby x_i' replaces the ith variable of x_0. Take shots at these points to obtain $G(x_0)$ and $G(x_0')$. Report this back to computationalists.

      ii. If the initial Hessian is decided to be the identity matrix the next few steps can be skipped, otherwise these steps should be taken for a decent guess at the initial Hessian.

      – Choose $H_0 = I$, **OR**

      – Compute H_0

         A. Do a shot each at 2 points about x_0 and 2 points about x_0' for the ith variable, record G for each of these points. Report these back to computationalists.

         B. Do *A.* for each variable

         C. Computationalists will compute H_0 from this data.

      iii. Compute optimal radius (steps *3g*-*3h* in §C.1)

         A. Computationalists will provide 6 new points

         B. Do a shot each at 3 different points. Report these back to computationalists (points associated to x_0).

         C. Do a shot each at 3 different points. Report these back to computationalists (points associated to x_0').

      iv. Compute initial gradient (step *3i* in §C.1)

         A. Computationalists will suggest a new point $x_{0'}$. Do two shots at this

point.

B. Computationalists will suggest a new point $x'_{0'}$. Do two shots at this point.

C. Do *3(a)ivA-3(a)ivB* $M-1$ more times and report this back to computationalists.

(b) Iterations

i. The values $x_k = [V_k, \alpha_k, h_k, x'_i]$ and $G(x_k, z)$ are known at this point

ii. Compute gradient at x_k (step *3m* in §C.1)

A. Computationalists will suggest a new point $x_{k'}$. Do two shots at this point.

B. Computationalists will suggest a new point $x'_{k'}$. Do two shots at this point.

C. Do *3(b)iiA-3(b)iiB* $M-1$ more times and report the results back to computationalists.

iii. Computationalists compute and use gradient to suggest new points x_{k+1} and x'_{k+1} and do updates (steps *3j-3p* in §C.1).

iv. Take a shot at x_{k+1}. Report $G(x_{k+1}, z)$ back to computationalists.

v. Take a shot at x'_{k+1}. Report $G(x'_{k+1}, z)$ back to computationalists.

vi. Repeat steps *3(b)i-3(b)v* until the computationalists tell you to stop.

- Compute diameter of uncontrollable variable v. Assume that m number of samples at each (nominal) point must be taken and averaged, as suggested by Eq. (4.24).

1. Take $x' = v$

(a) Initialization

i. Choose a nominal $x_0 = [V_0, \alpha_0, h_0]$, take m number of shots and average them to obtain $G(x_0) = \sum_{j=1}^{m} G(x_0)_j/m$. Report all samples and average back to computationalists.

ii. If the initial Hessian is decided to be the identity matrix the next few steps can be skipped, otherwise these steps should be taken for a decent guess at the initial Hessian.

- Choose $H_0 = I$, **OR**
- Compute H_0

A. Do m shots each at 2 points about x_0 and average the value of G for each of the 2 points of m shots, record this average for each of these points. Report all samples and the average back to computationalists.

B. Computationalists will compute H_0 from this data.

iii. Compute optimal radius (steps *3g*-*3h* in §C.1)

A. Computationalists will provide 3 new points

B. Do m shots each at 3 different points and average the value G for each of the 3 sets of m shots. Report all samples and the average back to computationalists.

iv. Compute initial gradient (step *3i* in §C.1)

A. Computationalists will suggest a new point $x_{0'}$. Do m shots at this point and average the value of G over the m shots.

(b) Iterations

i. The values $x_k = [V_k, \alpha_k, h_k]$ and $G(x_k, z)$ are known at this point

ii. Compute gradient at x_k (step *3m* in §C.1)

   A. Computationalists will suggest a new point $x_{k'}$. Do m shots at this point and average the value of G over the m shots.

iii. Computationalists compute and use gradient to suggest new point x_{k+1} and do updates (steps *3j-3p* in §C.1).

iv. Take m number of shots at x_{k+1} and average them to obtain $G(x_{k+1}) = \sum_{j=1}^{m} G(x_{k+1})_j/m$. Report this back to computationalists.

v. Repeat steps *1(b)i-1(b)iv* until the computationalists tell you to stop.