# Paradox of Optimal Learning: An Info-Gap Perspective

Yakov Ben-Haim[1] and Scott Cogan[2]

## Contents



**Abstract** Engineering design and technological risk assessment both entail learning or discovering new knowledge. Optimal learning is a procedure whereby new knowledge is obtained while minimizing some specific measure of effort (e.g. time or money expended). A paradox is a statement that appears self-contradictory, or contrary to common sense, or simply wrong, and yet might be true. The paradox of optimal learning is the assertion that a learning procedure cannot be optimized *a priori* — when designing the procedure — if the procedure depends on knowledge that the learning itself is intended to obtain. This is called a reflexive learning procedure. Many learning procedures can be optimized *a priori*. However, *a priori* optimization of a reflexive learning procedure is (usually) not possible. Most (but not all) reflexive learning procedures cannot be optimized without repeatedly implementing the procedure which may be very expensive. We discuss the prevalence of reflexive learning and present examples of the paradox. We also characterize those situations in which a reflexive learning procedure can be optimized. We discuss a response to the paradox (when it holds) based on the concept of robustness to uncertainty as developed in

[1]ORCID ID: 0000-0001-7229-8449, Faculty of Mechanical Engineering, Technion — Israel Institute of Technology, Haifa 32000 Israel, yakov@technion.ac.il

[2]FEMTO-ST Institute, Université de Franche-Comté, Besançon France, scott.cogan@univ-fcomte.fr

info-gap decision theory. We explain that maximizing the robustness is complementary to — but distinct from — minimizing a measure of effort of the learning procedure.

**Keywords** optimal learning, uncertainty, robustness, learning paradox, info-gaps

## 1 Introduction

Many quantitative entities have precise numerical values that are unknown. We don't know the length of this newly discovered cave. We don't know how long it will take to bring this promising technological innovation to practical fruition. We don't know the 47th significant digit of the rest mass of the electron. We don't know exactly how long it takes light to travel 1 km in a vacuum. We don't know the 100 quadrillionth digit of $\pi$.

Engineering design and technological risk assessment both entail learning or discovering new knowledge, e.g. innovative solutions for a product specification or evaluating severities of new failure mechanisms. Optimal learning is a procedure whereby new knowledge is obtained while some specific measure of effort (e.g. time expended) is minimized. Some of the quantitative entities mentioned above can be learned by optimal learning procedures. However, sometimes the learning procedure depends on the specific new knowledge that will be obtained. We will refer to such learning procedures as *reflexive* learning procedures.

A learning procedure usually depends on the context or substance of what is being learned. For instance, the procedure for learning the $n$th digit of $\pi$ is different from the procedure for learning the $n$th digit of the rest mass of the electron. However, neither of these procedures depends on the value of that digit — whether it is 0, or 1, or 2, etc. Neither of those procedures is a reflexive learning procedure. In contrast, the procedure for learning the depth of a new cave depends on that depth, because the procedure must specify the amount of food taken by the explorers, and that depends on whether the cave is deep or shallow.

We will see that most (but not all) reflexive learning procedures cannot be optimized *a priori:* at the design stage and before actual implementation (Ben-Haim, 2018, pp.70–72). Many learning procedures can be optimized *a priori,* so the inability to optimize a reflexive learning procedure when designing the procedure may seem paradoxical. This paper characterizes the paradox, presents a response based on info-gap decision theory, and identifies the unique conditions under which a reflexive learning procedure can be optimized *a priori.*

Section 2 briefly reviews concepts of learning, including optimal learning, in a wide range of disciplines. Section 3 discusses two preliminary qualitative examples of the paradox of optimal learning. Section 4 presents the mathematical formulation of the paradox of optimal learning. Section 5 characterizes those situations in which a reflexive learning procedure can be optimized *a priori.* Section 6 presents the mathematical formulation of info-gap robust-satisficing. Sections 7 to 10 discuss quantitative examples of the paradox and responses based on the concept of robustness to uncertainty as developed in info-gap decision theory. **Sections 7–9 formulate a selection of problems to indicate the range of topics that are accessible, while section 10 is a detailed numerical example.** Section 11 concludes the discussion.

We will use the term "learning" to refer to acquiring true or correct knowledge. While truth and correctness are sometimes disputed, by "learning" we do not mean things such as Huck Finn's acquiring the ability to "say the multiplication table up to six times seven is thirty-five" (Twain, 1884).

## 2 Learning: Optimal and Otherwise

Learning, and its optimization, has been explored in diverse contexts. We will briefly review recent studies of learning in computation and data science, reliability-based design optimization, economics, education, engineering systems, industrial management and operations research, and medicine. We will see the widespread importance ascribed to learning, as well as attempts to improve and even optimize the learning process. This is accompanied in specific situations by recognition that optimization is impractical or even impossible. This motivates the present paper which identifies generic structural factors that impose a fundamental limitation on optimization of learning.

*Computation and data science* is a broad and diverse field in which numerical evidence is employed in formulating and calibrating quantitative models. This is a learning procedure in the sense that one uses the data to learn an accurate or useful implementation of a model.

Fang *et al.* (2020) study the formulation of a regression function where the independent entity is a full probability distribution rather than a vector in a Euclidean vector space. They study a learning procedure whereby a squared error of the regression is sequentially minimized. What is learned is the functional relation between the set of probability distributions and the outcome quantity. The regression procedure itself does not depend on what is learned, and hence this is not a reflexive learning procedure. The authors present a learning algorithm in which the rate of learning is optimal.

Chang *et al.* (2022) explore the modelling of complex systems that are characterized by random variables $X = \{X_1, \ldots, X_n\}$. The joint probability distribution, $P(X)$, may be difficult to estimate, and may even be uncomputable, due to complex relations among the variables. They approximate $P(X)$ with a polynomial of finite order. Furthermore, they show that learning the structure of this polynomial can be done in a duration that is related polynomially (rather than exponentially) to the order of the approximating polynomial. The procedure for learning the polynomial structure depends on the order of the polynomial that is sought, and is thus a reflexive learning procedure. They don't address the optimality of this learning procedure, though they stress that polynomial time is required, which is less than exponential time.

***Reliability-based design optimization. (RBDO)***

**Zhang *et al.* (2020) study RBDO of complex or large-dimensional engineering systems by employing more manageable surrogate models. The properties of the complex system are learned, and the reliability of the design alternatives are assessed, by sampling the surrogate models. The authors demonstrate an accurate and computationally efficient design procedure. This is a reflexive learning procedure in the sense that the surrogate models are chosen in order to learn about the complex system which itself is being designed.**

**Li and Wang (2019) study the analysis of reliability while employing precise experimental data combined with models that may be biased. Their central question — which focusses on a reflexive learning procedure — is how to identify relevant information from both the model and the data upon which to assess the reliability of the system that has been modeled and measured. They develop a procedure for addressing this problem and fusing the model-based and experimental information.**

**Li and Wang (2020) explore RBDO when employing both low- and high-fidelity data. The learning procedure is the exploitation of the uncertain data in achieving optimal design. This will be a reflexive learning procedure when the data that are collected and selected depend on the design options that are analyzed.**

**Li and Wang (2022) use artificial neural networks for RBDO. The neural networks are trained with data that may differ, probabilistically, from test data on the system being developed. The design is optimized by learning about the system from training and test data. This learning is reflexive because it occurs on the system being designed.**

*Economics* has merited many studies of learning procedures. We first consider three examples in micro-economics, and then a macro-economic monetary policy example.

Aghion *et al.* (1991) develop a rigorous mathematical formulation in which an agent improves knowledge of their payoff function by repeatedly selecting an action, under identical conditions, and observing the outcome. They demonstrate that the long-run benefit of learning converges asymptotically to zero. They also show, under various rather strict mathematical conditions, that the agent can obtain a "true global optimum asymptotically" (p.642). The idea of a paradox of optimal learning is reflected in the following passage:

> "As long as the agent has not learnt all relevant aspects of his objective function he will be in pursuit of two conflicting objectives: the maximisation of his expected short-run payoff, and the maximisation of the informational content of the current action." (p.621)

This is related to the paradox of optimal learning, namely, that reflexive learning cannot be optimized when the learning process depends on what is to be learned. Aghion *et al.* however focus primarily on the asymptotic properties of the learning in which the paradox is resolved by repetition of the learning process.

Rob (1991) studies learning in markets with innovations or new products. Specifically, firms decide whether or not to enter such markets based on observing the profitability of earlier entrants. Firms' uncertainty about the profitability of new products is gradually reduced by learning from the market experience of other firms. The learning procedure that firms employ — observing the success rate of earlier entrants — depends on how many other firms have entered, which in turn depends on profitability of entry. Thus the learning procedure depends on what the firms wish to learn: profitability of entry. This is a reflexive learning procedure. Rob shows that the equilibrium rate of entry to the market decreases monotonically over time. Furthermore, the entry rate is less than a socially optimal rate. In this sense the learning is sub-optimal.

In a related study, Brezzi and Lai (2002) explore the dilemma facing a rational economic agent who wishes to maximize reward though a sequence of economic actions. The agent uses some of those actions to learn about the probability distribution of rewards, rather than choosing the putatively optimal action at each step. The learning procedure is the experimental deviation from putatively optimal action. The implementation of this experimentation depends on the economic reward, which is what the agent seeks to learn. The learning is thus reflexive. Nearly optimal learning procedures for finite time horizons are developed.

Wieland (2000) studies German monetary policy after reunification of East and West Germany. The impact of monetary policy is uncertain due to limited understanding of the public response to policy. The central bank must therefore engage in two tasks: controlling the policy target (e.g. inflation) and estimating the impact on the economy of policy actions. Wieland stresses that policy actions and estimation of policy impact are linked because estimation can improve policy actions. The learning procedure (estimation of market impact of monetary policy) is influenced by the level of policy impact because detection of dramatic impact will be easier than detection of small impact. The learning procedure is thus reflexive and thus cannot be optimized *a priori*. What Wieland refers to as an "optimal learning strategy" exploits this linkage and significantly improves economic stabilization and reduces the tendency for inflation (p.200).

*Education* is a learning procedure, and its improvement is the focus of much research.

Eichhorn *et al.* (2019) recommend procedures for teachers to enhance effectiveness of learning of mathematics by students in elementary and high school. The procedures are pedagogical, and the students learn. Thus the article presents a teaching procedure, not a learning procedure. Nonetheless the combined teacher-student interaction is a learning procedure. It is not reflexive because the procedure does not depend on what the student does not yet know because this is known to the teacher, and there is no paradox of optimal learning.

Son and Sethi (2006) study the problem of optimally allocating time between distinct learning tasks. They demonstrate that "optimal allocations are highly sensitive to structural characteristics of the learning environment. Hence strategies that are highly effective with one class of learning curves may be quite ineffective with another." (p.770) In particular, the optimal allocation depends on the "shape of the learning curve" and on the "goals or objectives of the learner" (p.760). It is evident that prior specification of an optimal learning-time allocation may be infeasible because the learning curve can change, or be poorly known *a priori*, and goals or objectives can change in the course of the learning.

Schuetze and Yan (2022) explore the balance between breadth and depth when learning the meanings of new words in limited time. The optimum balance is not unique, it depends on the individual and the topic, but balance between breadth and depth is relevant to these and other learning procedures. We note that the optimal balance may be unknown prior to the learning because that optimum may depend on subsequent use of the knowledge that will be obtained.

*Engineering systems* can be designed to adapt to the operational environment based on learning. Liu and Murphey (2020) study energy management in plug-in hybrid electric vehicles that run on two different power sources: the engine and the battery. The energy management decision that must be made continuously throughout the trip is the required power output of each source. The learning that takes place is an estimation of the remaining energy-cost of the trip. The algorithm for deducing this estimate is independent of the value that is learned, and thus is not reflexive. The authors show large reduction in the time required to make this estimate, though optimization is not addressed.

Machine learning has been widely used in engineering design. Specifically, transfer learning refers to the application of knowledge obtained in solving one problem, to solution of a different problem. For instance, Huang *et al.* (2022) explore transfer learning to problems in additive manufacturing. They write that "simulation for thermal modeling in metal AM [additive manufacturing] . . . is tedious and time-consuming", and that transfer learning enhances the efficiency of analysis. Similarly, Pandita *et al.* (2022) stress that experimentation and modeling of additive manufacturing processes "are expensive and oftentimes demand significant logistic and overheads." They study "transferring learned process maps from a source to a target process."

*Industrial management and operations research* can also employ learning.

Chen *et al.* (2020) study the problem of a firm that maintains an inventory of their product in anticipation of future sales. Too large an inventory is wasteful, while too small an inventory leads to lost sales. The policy by which the inventory is maintained depends on customer demand and on production capacity, both of which are random variables whose probability distributions are imperfectly known. For example, in the case of new products or changing markets, these probability distributions are changing and the firm must learn them over time based on past realizations of demand and supply. The authors present an algorithm for learning these probability distributions that enables an inventory policy that, asymptotically over time, approaches the policy that would be optimal if they actually knew the probability distributions of customer demand and production capacity. The paradox of optimal learning is manifested in the fact that the learning is optimal only asymptotically.

Epstein and Ji (2022) study the choice between generic options where the outcome is uncertain due to exogenous randomness and ignorance about a parameter, $\theta$, that can take one of two values. The decision maker does not know the probability distribution of $\theta$ but knows a set containing more than one possible probability distribution for $\theta$. The decision maker can defer the choice and observe outcomes and thus learn about the distribution of $\theta$, but these observations entail a cost. They explore the question of optimal learning: when to stop sampling and make a decision. One context in which they study this is an extension of Ellberg's urn problem (Ellsberg, 1961) of deciding whether to bet on an urn with a known combination of blue or red balls, or an urn with an unknown combination. They extend Ellsberg's problem by considering the possibility of learning about the ambiguous urn with unknown color ratio. The authors "show that *it can be optimal to reject learning completely, and, if some learning is optimal, then it is never optimal to bet on the risky urn after stopping.*" (italics in the original, p.1318). The learning derives from observing outcomes from the ambiguous urn, and thus depends on the composition of that urn. Hence the learning is reflexive. The authors' rejection of learning reflects in part the structure of the specific problem, and in part the paradox of optimal learning: the learning procedure cannot be optimized because it depends on what is to be learned.

*Medical practice* requires much learning, some of which is the attainment of motoric skill rather than knowledge. Cau *et al.* (2022) study the process by which students of surgery learn new surgical "gestures". The learning of these actions can be impeded by an emotion of "epistemic confusion" when an unanticipated or unfamiliar situation arises, even in a simulated operation. The control of this emotion can enhance the efficacy of the learning process, and also subsequently make the surgeon more effective in actual practice. This is reflexive learning because the learning procedure (performing the maneuver in simulation) depends on what must be learned (the surgical maneuver). (This differs, for example, from the non-reflexive learning procedure of discovering the next digit of $\pi$: the algorithm does not depend on the value of that digit.) The authors stress that learning must occur rapidly, so optimal learning might entail minimizing the time until the maneuver is learned. The maneuver was mastered by the teacher but not by the student whose learning is impeded by events that the learner could not anticipate precisely because the learner is inexperienced. The teacher attempts to ameliorate the emotional dimension of this confusion by providing psychological support. However, events that are unanticipated by the learner are inevitable precisely because the learner is still learning. Emotional response to confusion by an inexperienced learner cannot be entirely removed. This is the paradox of optimal learning: the process of learning cannot be optimized because it depends on the action that must be learned. Confusion and the attendant emotions, caused by unanticipated events in the surgical procedure, are inevitable in the learning process of an inexperienced individual.

## 3 Examples of the Paradox of Optimal Learning

As explained earlier, learning is prevalent in many fields, and optimal learning is a procedure whereby new knowledge is obtained while minimizing some specific measure of effort (e.g. time or money expended). A paradox is a statement that appears self-contradictory, or contrary to common sense, or simply wrong, and yet might be true. The paradox of optimal learning is the assertion that a learning procedure cannot be optimized *a priori* — at the design stage — when the procedure depends on knowledge that the learning itself is intended to obtain. This is called a reflexive learning procedure. Many learning procedures can be optimized before implementation. However, optimization of most (but not all) reflexive learning procedures is not possible.

For example, suppose that we wish to know how long it will take to bring an innovative insight to practical fruition. We can't know how long it will take until the project is successfully completed because unknown problems will have to be solved and unanticipated inventions will be needed along the way. The effort required to complete the project (and to learn the project duration) is measured as the amount of money required to fund the development team. Determination of the minimal amount of money that must be set aside for this development project depends on the duration of the development, which is unknown. The learning procedure (implementation of the project) and the effort (money required) depend on what is to be learned (the project duration). Optimal learning (minimizing the effort) is inaccessible. We cannot optimize the cost of innovative technological development at the project planning stage.

As another example, suppose that we wish to know the first 3 significant digits of the depth of this newly discovered cave. The effort required to obtain this knowledge is measured as the amount of food that must be carried by the explorers probing the cave. Determination of the minimal quantity of food requires knowledge of the depth of the cave, which is unknown. The learning procedure (spelunking) and the effort (food carried) depend on what is to be learned (cave depth). An optimal procedure for learning the cave depth is inaccessible when planning the exploration.

One resolution of the cave-depth paradox would seem to be to employ an adaptive learning procedure. Current knowledge indicates the amount of food typically required to reach and return from the end of caves in this region. One might add some additional food, and then start spelunking. If half of the food is consumed before reaching the bottom of the cave then return to the surface, augment the amount of food, and start again. This adaptive procedure is repeated until the end of the cave is reached.

But can one *optimize* this adaptive procedure before initiating it? For instance, can one minimize the total time required? Adding a single cracker on each iteration clearly is sub-optimal because many iterations will be needed. Adding a vast quantity of food is also sub-optimal because the explorers will be slowed down by the vast loads they are carrying. However, the optimal increment of food added on each iteration depends on the depth of the cave. An optimal adaptive procedure for learning the cave depth is inaccessible before beginning the spelunking.

In both of these examples the learning procedure depends on the specific new knowledge that will be learned. The technological development involves commitment of resources and payments to the developers, all of which are influenced by the duration of development. The quantity of food needed by the spelunkers depends on the depth of the cave. We are not claiming that there is no optimum in either of these examples. There may indeed be a minimal budget or a minimal food requirement, and one may even accidentally choose the minimum just by chance. We are claiming that one cannot know or evaluate the minimum *a priori.*

A paradox is a statement that seems to be self-contradictory or erroneous but that might nonetheless be true. The paradox of optimal learning is the assertion that *a priori* optimization of a learning procedure is not possible if the learning procedure depends on the knowledge that one seeks to learn. The above preliminary examples illustrate reflexive learning procedures that cannot be optimized at the design stage: before implementation. Nonetheless, some reflexive learning procedures can be optimized *a priori,* as we will see. And we provide a response when they cannot be optimized.

## 4 Mathematical Formulation of the Paradox of Optimal Learning

The knowledge that is sought may be a specific number, or a vector, or a functional relationship, or qualitative semantic terms, or a combination of these entities. Let $k$ denote the knowledge that is sought, and let $r$ specify the realization of a specific learning procedure intended to obtain this knowledge. $r$ represents how the learning procedure is actually implemented. The effort that would be required to learn (or verify) that the knowledge takes the value $k$, with this realization of the learning procedure, is denoted $e(k,r)$. We will assume that the effort is a single scalar quantity.

An optimal learning procedure is one that minimizes the effort. We denote an optimal learning procedure as:

$$r^{\star}(k) = \arg\min_{r} e(k,r) \tag{1}$$

The precise statement of the paradox of optimal learning is that optimization of the learning procedure, $r$, is not possible *a priori* — when designing the learning procedure — if $r^{\star}(k)$ depends on the knowledge that is sought, $k$. This knowledge is unknown before the learning has been performed, so $r^{\star}(k)$ cannot be known or determined when designing the learning procedure. This is illustrated schematically in fig. 1. The horizontal axis schematically represents the domain of possible learning procedures, $r$, and the vertical axis is the effort, $e(k,r)$, required to obtain knowledge $k$ with procedure $r$. The figure shows the effort, as a function of the learning procedure, for 3 different realizations of the knowledge. The optimal learning procedure, $r^{\star}(k_n)$, is different for each knowledge state, $k_1$, $k_2$, and $k_3$. One cannot know which of these procedures is optimal without knowing which state of knowledge prevails: the learning procedure cannot be optimized before the knowledge has been attained.

It is possible that $r^{\star}(k)$ is actually *independent* of $k$ even if the learning procedure, as manifested in the effort $e(k,r)$, is reflexive and does depend on $k$. The effort, $e(k,r)$, as a function of the learning procedure, $r$, may change as the knowledge, $k$, changes. Even the minimal value of the effort may differ as $k$ changes. It is still possible that the learning procedure which minimizes the effort, $r^{\star}(k)$, is independent of $k$. This is illustrated schematically in fig. 2 in which we see that the optimal learning procedure, $r^{\star}(k_n)$, is the same for each knowledge state, $k_1$, $k_2$, and $k_3$.

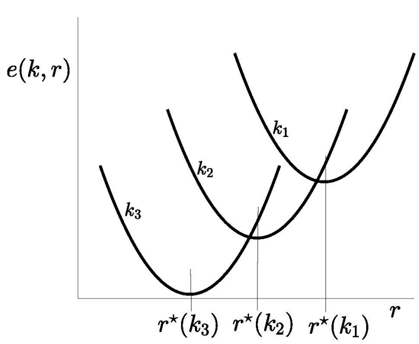


Figure 1: Schematic effort functions with different optimal learning procedures.

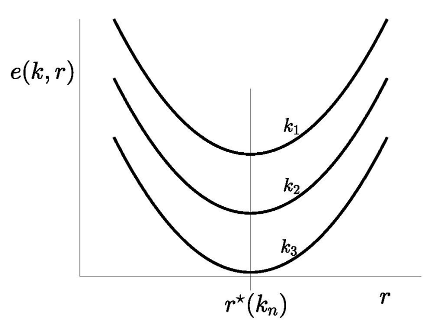


Figure 2: Schematic effort functions with the same optimal learning procedure.

The effort function, $e(k,r)$, is known *a priori*, before the value of $k$ is learned. Thus one can know if its minimum (as a function of $r$) is independent of $k$ and that consequently the paradox of optimal learning is avoided. That is, it is possible that the paradox of optimal learning is avoided and *a priori* optimization is possible. We will see an example in section 5.

## 5 Optimal Learning is Sometimes Possible

The effort required to learn something may depend very strongly on the specific value that is to be learned. The effort to learn the depth of a cave or the height of a mountain increases greatly as the depth or height increases. Consider a simplistic preliminary example.

The knowledge, $k$, is a single scalar value and the effort is an exponential function of $k$:

$$e(k,r) = a\mathrm{e}^{bk} \tag{2}$$

where $a$ and $b$ are known positive constants that depend on the learning procedure, $r$. The learning procedure is a reflexive learning procedure if it depends on the knowledge, as in the cave-exploration example.

If the knowledge, $k$, is known to take some positive value, then the effort can be minimized by choosing the learning procedure so that both $a$ and $b$ are as small as possible (but positive by definition). In this case, the learning can be optimized *a priori* because the procedure that minimizes the effort does not depend on the value of $k$. However, if all we know **is that $k$ is a real number**, then we don't know whether $b$ should be small or large in order to minimize the effort; *a priori* optimization of the learning procedure is not possible. Both cases entail reflexive learning, but the first case can be optimized before implementation (because the minimum effort does not depend on the specific value of $k$) while the second cannot.

## 6 Mathematical Formulation of Robust Satisficing

The paradox of optimal learning is the assertion that one cannot optimize — *a priori*, before implementation — some measure of the effort when the learning procedure depends on knowledge that has not yet been obtained. An info-gap response is to *satisfice* the effort and to *optimize* the robustness against one's ignorance of the knowledge that is sought. We will treat this ignorance as a deep uncertainty, as modeled and managed in info-gap decision theory (Ben-Haim, 2006). The optimization is shifted from a substantive component (the effort) to a procedural component (the robustness to uncertainty). Optimization of the robustness in a reflexive learning procedure can be achieved, as we will illustrate with examples in sections 7–10.

Our goal is to satisfice the effort, namely, to assure that the effort does not exceed a critical value, $e_\mathrm{c}$:

$$e(k,r) \le e_\mathrm{c} \tag{3}$$

We do not know the knowledge, $k$, though some limited and highly uncertain knowledge is available. That knowledge and ignorance is represented by an info-gap model of uncertainty (Ben-Haim, 2006). An info-gap model is an unbounded family of nested sets, $\mathcal{U}(h)$, for $h \ge 0$, of possible realizations of the uncertain entity (the knowledge in the present context); we will see examples subsequently. Info-gap models of uncertainty satisfy two axioms, contraction and nesting.

*Contraction* is the assertion:

$$\mathcal{U}(0) = \left\{\widetilde{k}\right\} \tag{4}$$

where $\widetilde{k}$ is a known (but possibly faulty) estimate of the knowledge. The contraction axiom asserts that the estimate is correct in the absence of uncertainty.

*Nesting* is the assertion:

$$h < h' \quad\Longrightarrow\quad \mathcal{U}(h) \subseteq \mathcal{U}(h') \tag{5}$$

The nesting axiom asserts that the uncertainty sets become more inclusive as $h$ increases.

These axioms endow $h$ with its meaning as an horizon of uncertainty: the range of variation of the uncertain entity increases as $h$ increases. An info-gap model is a non-probabilistic representation of the uncertainty about the knowledge. Furthermore, the family of sets is unbounded — all we know about $h$ is that it is non-negative — so there is no known worst case or greatest error.

The robustness — to uncertainty in the knowledge — of the learning procedure $r$, is the greatest tolerable uncertainty. More precisely, the robustness is the maximum horizon of uncertainty, $h$, up to which the effort satisfies the condition in eq.(3) for all realizations of the knowledge in the uncertainty set $\mathcal{U}(h)$. Formally, the robustness of a learning procedure $r$ is:

$$\widehat{h}(e_{\rm c}, r) = \max \left\{ h : \left( \max_{k \in \mathcal{U}(h)} e(k, r) \right) \leq e_{\rm c} \right\} \tag{6}$$

We point out that the robustness function, $\widehat{h}(e_{\rm c}, r)$ does not depend on the knowledge, $k$, that is sought. Thus the robustness can be evaluated, and the learning procedure that maximizes the robustness can be determined. That is, the value of $r$ that maximizes the robustness, $\widehat{h}(e_{\rm c}, r)$, can be found *a priori*, even if the value of $r$ that minimizes the effort, $e(k, r)$, cannot be found *a priori*.

Let $\widehat{r}$ denote the learning procedure that maximizes the robustness:

$$\widehat{r} = \arg\max_{r} \widehat{h}(e_{\rm c}, r) \tag{7}$$

We note again that $\widehat{r}$ does not depend on the knowledge, $k$, which has not yet been obtained. Thus $\widehat{r}$ can be determined *a priori*, unlike $r^{\star}(k)$ in eq.(1).

**One might ask how much does the robust-optimal learning procedure, $\widehat{r}$, depend on the initial guess, around which the info-gap model of uncertainty is constructed? The answer is entailed in the meaning and definition of the robustness. If the robustness is large then the robust-optimal learning procedure is relatively independent of the initial knowledge, while low robustness implies larger dependence on the prior knowledge.**

Eq.(6) implies a complementarity between effort and robustness: small effort entails large robustness. However, we cannot maximize the robustness by minimizing the effort because the learning is reflexive. Nonetheless, $\widehat{r}$ is the learning procedure that minimizes the maximum of the effort as a function of the unbounded horizon of uncertainty, $h$. We cannot minimize the effort, but we can minimize the sensitivity of the effort to uncertainty in the knowledge that is sought. This is what $\widehat{r}$ achieves.

The complementarity between effort and robustness to uncertainty can be expressed in the following proposition.

**Proposition 1** *Lower effort implies greater robustness to uncertainty.*

Given:

- *A scalar effort function, $e(k, r)$.*
- *An info-gap model of uncertainty in the knowledge, $\mathcal{U}(h)$, $h \geq 0$.*
- *The robustness function based on this info-gap model, $\widehat{h}(e_{\rm c}, r)$, defined in eq.(6).*
- *Two learning procedures, $r$ and $r'$.*

If:

$$e(k, r) \leq e(k, r') \quad \textit{for all} \quad k \in \mathcal{U}(h),\ h \geq 0 \tag{8}$$

Then:

$$\widehat{h}(e_{\rm c}, r) \geq \widehat{h}(e_{\rm c}, r') \quad \textit{for all} \quad e_{\rm c} \tag{9}$$

The proof of proposition 1 appears in appendix A.

We note that the implication in proposition 1 is uni-directional — lower effort throughout the knowledge domain implies greater robustness. The reverse implication need not hold. Lower effort and greater robustness display a complementarity, but they are not equivalent. This is related to the difference between $r^\star(k)$ and $\widehat{r}$ in eqs.(1) in (7): the latter can always be determined *a priori* while the former cannot.

**We will consider 4 examples in the following 4 sections.**

## 7 Empirical Regression: Analysis of Robustness to Uncertainty

The knowledge to be obtained, $k$, is an empirical regression between two variables, such as stress versus strain in a new non-isotropic material, or force versus acceleration in a medium with high viscosity. The design of the sampling procedure to obtain this knowledge depends on a poorly known probability density function that contains at least some of the knowledge that is sought. Thus this is a reflexive learning procedure.

The best estimate of the probability density is $\widetilde{p}(x)$, and the uncertainty in this estimate is represented by an info-gap model, $\mathcal{U}(h)$ for $h \geq 0$. We must estimate the independent variable $x$ (e.g. stress or force). **The effort required for this estimation is proportional to the value of $x$. If $x$ takes a low or moderate value then only low or moderate effort is required for the estimation (e.q. only low or moderate forces are required); a large value of $x$ requires large effort.**

**We will use the quantile function to assess the effort. For any value of $\alpha$ between 0 and 1, let $q_\alpha$ denote the quantile function of $x$ based on the true distribution, $p(x)$. The quantile is defined as the probability that $x$ does not exceed $q_\alpha$.** The formal definition of the quantile is:

$$\alpha = \text{Prob}(x \leq q_\alpha) = \int_{-\infty}^{q_\alpha} p(x)\,\mathrm{d}x \tag{10}$$

**Thus the effort, as assessed by the quantile $q_\alpha$, is the probability that the value to be estimated is no greater than $q_\alpha$.**

We require that **the effort, as assessed by the** quantile, not exceed a critical value, $q_\mathrm{c}$:

$$q_\alpha \leq q_\mathrm{c} \tag{11}$$

The robustness of the learning procedure, to uncertainty in the probability distribution, is the greatest horizon of uncertainty, $h$, up to which the quantile $q_\alpha$ satisfies eq.(11) for all probability distributions $p(x)$ in the uncertainty set $\mathcal{U}(h)$:

$$\widehat{h}(q_\mathrm{c}) = \max \left\{ h : \left( \max_{p \in \mathcal{U}(h)} q_\alpha \right) \leq q_\mathrm{c} \right\} \tag{12}$$

As a specific simple example suppose that we know that the probability density is exponential, but that the estimate of the exponential coefficient is uncertain. Thus:

$$p(x) = \lambda \mathrm{e}^{-\lambda x}, \quad x \geq 0 \tag{13}$$

**The coefficient $\lambda$ is uncertain, though it must be positive. The known estimate of $\lambda$ is $\widetilde{\lambda}$ but, while this estimate is based on some prior knowledge, the extent of error of this estimate is unknown. In some situations one may also have prior knowledge of an error estimate, denoted $w$, e.g. as a result of sampling or prior experience. In this case the most one can say is that the fractional**

**error of $\widetilde{\lambda}$, in units of $w$, is unbounded. In other situations one has no prior error estimate, in which case we choose $w = \widetilde{\lambda}$, and the fractional error of $\widetilde{\lambda}$, in units of $\widetilde{\lambda}$ itself, is unbounded.** This uncertainty in the probability distribution is quantified by the following info-gap model for uncertainty in $\lambda$:

$$\mathcal{U}(h) = \left\{ \lambda : \ \lambda > 0, \ \left| \frac{\lambda - \widetilde{\lambda}}{w} \right| \leq h \right\}, \quad h \geq 0 \tag{14}$$

where $w$ is a known positive constant.

The following expression for the robustness is derived in appendix B:

$$\widehat{h}(q_c) = \frac{1}{w} \left( \widetilde{\lambda} + \frac{\ln(1-\alpha)}{q_c} \right) \tag{15}$$

The robustness is defined to be zero for values of $q_c$ at which this expression is negative.

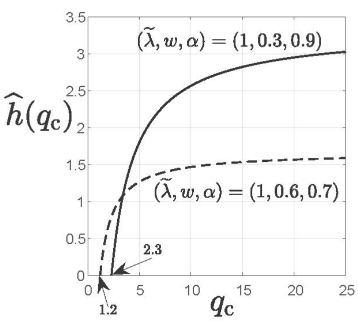


Figure 3: Robustness functions illustrating trade off, zeroing and preference reversal.

Fig. 3 shows robustness curves for two different realizations of the learning procedure as manifested in the parameters $\widetilde{\lambda}$, $w$ and $\alpha$. From the robustness function in eq.(15) we see that the horizontal intercept occurs when $q_c$ takes its estimated value which from eq.(29) in appendix B is:

$$q_\alpha = -\frac{1}{\widetilde{\lambda}} \ln(1-\alpha) \tag{16}$$

The horizontal intercepts of the robustness curves are 1.20 (dash) and 2.30 (solid) because the estimated quantile for $\alpha = 0.7$ (dash) is less than for $\alpha = 0.9$ (solid). The effort is assessed with this quantile, so $\alpha = 0.7$ has lower predicted effort and thus is putatively preferred. However, these predicted quantiles have no robustness to uncertainty in the shape of the uncertain probability distribution (the robustness at the horizontal intercept is zero). Hence preference for the dashed configuration — based on these predictions — is unfounded. This is the property of *zeroing:* predicted outcomes have zero robustness to uncertainty in the data and models upon which the predictions are based.

The robustness curves in fig. 3 have positive slope, showing that robustness increases (which is desirable) as the critical quantile, $q_c$, gets larger (which is undesirable because this means greater effort). This positive slope reflects the universal *trade off* between robustness to uncertainty and quality of the outcome. More demanding outcomes (lower $q_c$ in the present example) have lower robustness to uncertainty. Combining the concepts of zeroing and trade off, one sees that predicted outcomes are unreliable (zeroing), while less desirable outcomes are more confidently anticipated (trade off). The robustness function quantifies this trade off.

From eq.(15) we see that the robustness asymptotically approaches the value $\widetilde{\lambda}/w$ as the critical quantile, $q_{\rm c}$, approaches infinity. The asymptotic robustness values in fig. 3 are $3.\overline{3}$ (solid) and $1.\overline{6}$ (dash). The solid configuration has better (greater) asymptotic robustness to uncertainty but putatively worse (greater) predicted quantile. Consequently the robustness curves cross one another. At low values of $q_{\rm c}$ the dash configuration is more robust and hence preferred over the solid configuration, while the situation is reversed at larger values of $q_{\rm c}$. In other words, the intersection between the robustness curves, as seen in fig. 3, entails the potential for a reversal of the robust preference between these two learning procedures.

We now relate the info-gap concept of robustness to the paradox of optimal learning.

As explained earlier, the paradox of optimal learning is the assertion that a learning procedure cannot be optimized *a priori* when the procedure depends on knowledge that the learning itself is intended to obtain. In the present example the learning procedure (random sampling of a population) depends on an uncertain probability distribution that entails at least part of the knowledge that is sought (a regression). Thus $p(x)$ is uncertain (because the knowledge is yet to be obtained) and hence the quantile cannot be minimized. The robustness to uncertainty, however, can be evaluated as we have seen.

Each robustness curve in fig. 3 refers to a different learning procedure as characterized by the parameters $(\widetilde{\lambda}, w, \alpha)$. The dash procedure is putatively better than the solid procedure as demonstrated by their horizontal intercepts. The putatively optimal choice between these learning procedures would be for dash. However, the zeroing property implies that this prioritization is unreliable, even fatuous or irresponsible, because the predicted qualities of these two procedures have no robustness to the uncertainty of the knowledge underlying the predictions. This underlying knowledge (the uncertain probability distribution) entails part of the knowledge that is sought (the regression relation). One cannot optimize the choice between these two learning procedures, *a priori*, because they depend on the knowledge that is sought. This is the paradox of optimal learning.

The learning procedure itself cannot be optimized *a priori*, but the robustness to uncertainty *can* be optimized while satisficing the learning procedure. That is, the burden of optimization is shifted from the learning procedure itself, to the robustness to one's current ignorance that will be dispelled by the learning procedure. This is a shift from a substantive entity (the learning procedure) to a methodological entity (immunity to uncertainty). We are not removing the initial paradox or refuting the assertion that *a priori* optimization of the learning procedure is impossible. Rather, we are shifting attention to a different task: management of ignorance. As explained earlier, the robustness curves enable a unique choice between the dash and solid learning procedures, to optimize the robustness, depending on the decision maker's required level of performance, $q_{\rm c}$.

## 8 Empirical Estimation of an Uncertain 1-Dimensional Function

In many situations one uses a limited number of measurements to estimate the shape of a function. For instance, quality control measurements certify the shape of a processed surface. Measurements of market prices and quantities purchased of a specific product are used to estimate the relation between supply and demand for that product. The mechanical efficiency of a milling machine is assessed by measuring the rate of progress as a function of power. Other examples abound, and we will consider a simple generic version of these estimations.

The function $f(x)$ is defined on the 1-dimensional domain $X = [0, 1]$. We will use measurements at discrete points in $X$ to estimate the shape of the function. The question we face is how to choose the measurement locations, which we denote with the set $X_N$. We will show that one cannot choose the measurement locations to optimize a substantive measure of performance. We will then argue that one can choose the measurement locations to satisfice the substantive performance and optimize the robustness to uncertainty about the estimated function.

The best-known approximation to the function $f(x)$ is $\widetilde{f}(x)$. However, the fidelity of this approximation to the true function is highly uncertain. That is, $\widetilde{f}(x)$ is quite likely wrong in important but unknown ways. Thus $\widetilde{f}(x)$ is not a good basis for selecting measurement locations, though it will serve as the center point for an info-gap model of uncertainty in $f(x)$.

The learning procedure is based on where, in $X$, to perform the $N$ measurements. The design of the learning procedure hinges on how to choose the measurement locations in the set $X_N$. One approach is to disperse the measurements so that the density of measurements is proportional to the rate of change of the function, $f(x)$. This implies performing more measurements where $f(x)$ is varying rapidly. To formalize this we divide the domain $X$ into $M$ equal small intervals, and let $\mu_i$ denote the midpoint of the $i$th interval. Denote the absolute derivative of $f(x)$ at $\mu_i$ as:

$$\delta_i = \left| \frac{\mathrm{d}f(x)}{\mathrm{d}x} \right|_{\mu_i}, \quad i = 1, \ldots, M \tag{17}$$

The number of measurements to be performed in the $i$th interval is proportional to the rate of change of $f(x)$ in that interval:

$$n_i = \frac{\delta_i}{\sum_{j=1}^{M} \delta_j} N, \quad i = 1, \ldots, M \tag{18}$$

For simplicity we are ignoring the fact that the $n_i$'s must take discrete values.

This is a reflexive learning procedure because the choice of the measurement points depends on the function, $f(x)$, that is to be estimated, as seen in eqs.(17) and (18). The choice of measurement points according to eq.(18) is optimal, but it cannot be implemented *a priori* because the learning procedure is reflexive. That is, eqs.(17) and (18) depend on the function, $f(x)$, about which we are trying to learn.

One standard approach to solving this quandary is to choose the measurement positions by employing the estimated function, $\widetilde{f}(x)$, in eq.(17). Let $X_N^\star$ denote this choice of the measurement positions. The problem with this approach is that $\widetilde{f}$ is likely to be quite wrong, so selecting the measurement positions according to $\widetilde{f}$ is unrealistic.

We now suggest an alternative response to the impossibility of optimally choosing $X_N$ *a priori.* This approach explicitly addresses the uncertainty in the function $f(x)$. Let $\nu_i(f)$ denote the right-hand side of eq.(18), which is the optimal value of the number of measurements in the $i$th bin. These quantities are unknown at the design stage because they depend on the unknown function, $f(x)$, whose shape is to be learned by the measurements. Let $n_i$ denote a choice of the number of measurements to be performed in the $i$th bin, which may differ from $\nu_i(f)$, for $i = 1, \ldots, M$. The total squared difference between the actual and optimal number of measurements is:

$$S(f, X_N) = \sum_{i=1}^{M} [n_i - \nu_i(f)]^2 \tag{19}$$

The choice of $n_i$ in eq.(18) would minimize this quantity. However, that choice is unknowable, before performing the measurements, precisely because it depends on $f(x)$ which is what we are trying to learn.

The approach advocated here is to choose the measurement locations to satisfice the squared deviation, $S(f, X_N)$, and to maximize the robustness to uncertainty in the true function, $f(x)$. We formulate this as follows.

Let $\mathcal{U}(h)$ denote an info-gap model for uncertainty in the function $f(x)$. This info-gap model contains whatever knowledge we have about $f(x)$, e.g. the current best estimate $\widetilde{f}(x)$, and perhaps additional information such as monotonicity, or uni-modality, or end-point values, etc. The robustness function is the greatest horizon of uncertainty, $h$, up to which the mean squared error is no greater than a critical value, $S_{\rm c}$. The formal definition of the robustness function is:

$$\widehat{h}(S_{\rm c}, X_N) = \max \left\{ h : \left( \max_{f \in \mathcal{U}(h)} S(f, X_N) \right) \leq S_{\rm c} \right\} \tag{20}$$

The robust-optimal choice of the measurement positions is the value of the set $X_N$ that maximizes the robustness for satisficing the squared error. Formally, the robust-optimal measurement positions are defined as the set:

$$\widehat{X}_N = \arg \max_{X_N} \widehat{h}(S_{\rm c}, X_N) \tag{21}$$

This choice of measurement locations will usually differ from the estimated optimal locations, $X_N^\star$, defined earlier. We stress that the measurement locations $\widehat{X}_N$ can be evaluated *a priori*, while the locations in $X_N^\star$ cannot.

# 9 Adaptive Estimation of an Uncertain 1-Dimensional Function

We now consider an adaptive modification of the example in section 8.

The function $f(x)$ is defined on the 1-dimensional domain $X = [0, 1]$. We will use measurements at discrete points in $X$ to estimate the shape of the function. The measurement locations are chosen adaptively, that is, after each measurement an adaptive location-selection algorithm chooses the next measurement location. We will show that one cannot design this adaptive algorithm, *a priori*, to minimize the number of steps needed to achieve a specified level of fidelity between the true and the estimated functions. We will then argue that one can design the adaptive algorithm *a priori* to satisfice the number of steps and the level of fidelity, and to optimize the robustness to uncertainty about the estimated function.

$N$ (possibly noisy) measurements of $f(x)$ have been made, and that data is used to fit a functional approximation to $f(x)$. Specifically, the measurements are used to choose the coefficients, $c$, of a function denoted $\phi(x|c, N)$. The remaining uncertainty in the true function is represented by an info-gap model, $\mathcal{U}_N(h)$. Note that $\mathcal{U}_N(h)$ does not necessarily become more restrictive as the number of measurements grow; new information may in fact reveal new uncertainties.

After the $N$th measurement, the measure of deviation of the estimated function, $\phi(x|c, N)$, from the true function, $f(x)$, is denoted $S(f, N)$. For example, this might be the integral squared deviation:

$$S(f, N) = \int_0^1 \left[ \phi(x|c, N) - f(x) \right]^2 \, {\rm d}x \tag{22}$$

We desire a small value for this measure of error. Let $S_{\rm c}$ denote the largest acceptable value. We assume that $S(f, N)$ is a non-increasing function of $N$, that is, the error of the estimate does not increase as the number of measurements increases.

Let $N(f, S_{\rm c})$ denote the number of measurements required so that the measure of error does not exceed a critical value $S_{\rm c}$. Note that this number of measurements depends on the function, $f(x)$, that is unknown and is being estimated.

Let $\alpha$ denote the design decisions underlying the formulation of the adaptive algorithm. These may include parameters, or functions, or linguistic variables. For any specification of the measure of error, let $\alpha^\star(f, S_c)$ denote the design of the adaptive algorithm that minimizes the number of steps required to achieve error $S_c$:

$$\alpha^\star(f, S_c) = \arg\min_{\alpha} N(f, S_c) \tag{23}$$

Note that this algorithm depends on the function, $f(x)$, that is being estimated. In other words, the design of the adaptive algorithm is a reflexive learning procedure. One cannot determine the optimal adaptive algorithm, $\alpha^\star(f, S_c)$, *a priori* because it depends on the function that is being estimated.

Recall that $\phi(x|c, N)$ denotes the estimate of the function after the $N$th measurement. One could design the adaptation in eq.(23) based on $\phi(x|c, N)$ rather than on $f(x)$, and this will be plausible once this estimate is reasonably accurate. However, early in the adaptive procedure, when $f(x)$ is deeply uncertain and $\phi(x|c, N)$ is still quite likely wrong in important ways, this will be an unreliable procedure.

We now propose an alternative procedure for designing the adaptive algorithm. This procedure is based on satisficing the number of steps (required to reduce the measure of error to $S_c$) and maximizing the robustness to uncertainty in the estimated function. In other words, we are not trying to optimize the error (which cannot be done *a priori* because this is a reflexive learning procedure). Rather we are satisficing the error and optimizing the robustness to uncertainty. The robustness (to uncertainty in $f(x)$) for reducing the measure of error to $S_c$ in no more than $N_c$ steps is defined as:

$$\widehat{h}(S_c, N_c) = \max \left\{ h : \left( \max_{f \in \mathcal{U}(h)} N(f, S_c) \right) \le N_c \right\} \tag{24}$$

The robustness function depends on the adaptive algorithm, $\alpha$, because the number of steps, $N(f, S_c)$, depends on the algorithm. The robust-optimal algorithm, $\widehat{\alpha}(N_c, S_c)$, maximizes the robustness to uncertainty, while satisficing the measure of error and the number of steps. This robust-optimal algorithm is formally defined:

$$\widehat{\alpha}(S_c, N_c) = \arg\max_{\alpha} \widehat{h}(S_c, N_c) \tag{25}$$

This optimization of the adaptive algorithm can be performed *a priori* because it does not depend on the unknown function, $f(x)$. In fact, $\widehat{\alpha}(S_c, N_c)$ explicitly addresses the uncertainty in $f(x)$ through the info-gap model of uncertainty. In addition, one will usually find that $\widehat{\alpha}(S_c, N_c)$ differs from $\alpha^\star(f, N_c)$ for any particular realization of the function $f(x)$. More importantly, $\widehat{\alpha}(S_c, N_c)$ is the adaptive algorithm that achieves specified performance — in terms of measure of error and number of steps — for any realization of the unknown function up to horizon of uncertainty equal to the robustness. If this robustness is large then one can have confidence in achieving these outcomes; low robustness implies that these outcome requirements are unrealistic.

## 10 Surrogate Functions in Design of Complex Systems: Satisficing the Effort

Consider a complex mechanical system, for instance a railroad car roaring along the track, or a micro-mechanical manipulator, or an autonomous robotic surgical device. The design process entails "learning" or "discovering" the values of various design parameters — material properties, geometrical dimensions, rotational degrees of freedom, etc. — that will achieve the required performance. The design process involves, among other things, heavy computations of mechanical models that depend on the design that is sought. The paradox of optimal learning asserts that one cannot formulate an *a priori* procedure for discovering these design values that optimizes a measure of effort — e.g. number or duration of the computations. This is because the learning procedure is reflexive: the computations and associated numerical search depend on the design values that are sought. Likewise, if the design procedure is adaptive, then the paradox of optimal learning asserts that one cannot design the adaptive procedure *a priori* to minimize a measure of effort. However, the effort can be satisficed by focussing on the robustness to uncertainty in the value of the design variables. In this section we develop these ideas in a specific **numerical** example.

## 10.1 Introduction

We focus on a single design variable such as a beam thickness or material stiffness, and a single performance function such as vibration amplitude, or cost, or weight. Any choice of the design variable determines a unique value of the system performance, for which a small value is desirable. The design variable is denoted by $d$ and the system performance function is $S(d)$.

The performance function can be evaluated very accurately at any point in the design space, but this is very costly and can be done only a limited number of times. Very little is initially known about the performance function except that it is a real scalar function and perhaps non-negative by definition.

Precise evaluation of $S(d)$ is made at $N$ points in the design space denoted by the set $\mathcal{D} = \{d_1, \ldots, d_N\}$. The number of precise evaluations, $N$, is a measure of the effort in the design process. One would like to minimize the number of these heavy calculations, while enabling choice of the design variables to achieve satisfactory performance. However, the number and location of the calculation points depends on what one seeks to learn: the choice of the design variables to achieve good performance as assessed by $S(d)$. Thus this is a reflexive learning procedure. The paradox of optimal learning asserts that one cannot minimize $N$ *a priori*, either statically (fixed $N$) or adaptively (successive additional calculations).

The values of $S(d)$ at the calculation points, $\mathcal{D}$, are fit with a "surrogate" function, $\phi(d)$, that is readily computed and that passes through the $N$ exact values of the performance function (see Queipo *et al.* 2005). This surrogate provides an approximation to the performance throughout the design space, and will serve as the basis for design, though we will have to address the uncertain fidelity of $\phi(d)$ to $S(d)$ at points other than $\mathcal{D}$.

The central task that we face is to identify regions in the design space in which the performance function is acceptable. Performance is acceptable if its value is no greater than a critical value $S_c$, which may be a specified value or it may be a parameter that will be explored in the design process. If the exact calculated value of $S(d_i)$ at some calculated point $d_i$ in design space is acceptable, then we would like to identify the region around this design point throughout which the performance is acceptable. More generally, if $\phi(d)$ seems acceptable at some point $d$ (recalling that $\phi(d)$ deviates from $S(d)$), then we want to know how large a region around $d$ is actually acceptable according to $S(d)$. Two supporting questions are related to this task: should an additional exact calculation of the performance function be made, and if so, where in the design space should the calculation be made? These calculations are very costly, so one would like to minimize the number, but this is unfeasible because the learning is reflexive, as noted earlier. Hence we satisfice the number of calculations while managing the uncertainty of the performance function with the info-gap robustness.

### 10.2 Formulation of Robustness to Uncertainty

$\phi(d)$ is known throughout the design space, and acts as a surrogate for the true performance function, $S(d)$, which is known only at the $N$ points at which $S(d)$ was calculated. Suppose that we have identified a region, $\Delta$, in design space throughout which $\phi(d) \le S_{\rm c}$, suggesting that this is an acceptable region of the design space. However, the function $\phi(d)$ is precise only at the points at which the performance function was calculated, and $\phi(d)$ may deviate from $S(d)$ at other points in $\Delta$. We want to know if the $N$ calculations are sufficient to confidently assert that $S(d) \le S_{\rm c}$ throughout $\Delta$. We can be confident in $\Delta$ as a legitimate design set — based on $N$ exact calculations — if $S(d)$ can deviate greatly from $\phi(d)$ and still satisfy the performance requirement throughout $\Delta$. We formalize these ideas as follows.

Let $\mathcal{U}(h)$ denote an info-gap model for uncertainty in the performance function, $S(d)$. There are various possibilities for the info-gap model, depending on what one knows or is willing to assume. We will assume very little formal prior information about the performance function, while also recognizing that the designer may have qualitative contextual understanding that further constrains the possible performance functions. The fractional-error model is a minimally informative info-gap model:

$$\mathcal{U}(h) = \left\{ S(d) : \ S(d) = \phi(d) \ \forall \ d \in \mathcal{D}, \ \left| \frac{S(d) - \phi(d)}{\phi(d)} \right| \le h \ \forall \ d \right\}, \quad h \ge 0 \tag{26}$$

Each set, $\mathcal{U}(h)$, in this info-gap model contains all performance functions, $S(d)$, that take the calculated values at the points in $\mathcal{D}$, and which fractionally deviate from the surrogate, $\phi(d)$, by no more than $h$ throughout the entire design space. This info-gap model allows discontinuous and non-smooth performance functions, $S(d)$.

There are other info-gap models, such as the slope-bound model, the Fourier ellipsoid-bound model, or the energy-bound model, that employ different prior knowledge of the performance function (Ben-Haim, 2006).

The robustness of a proposed design region, $\Delta$, based on $N$ calculations, is the greatest horizon of uncertainty in the true performance function up to which its performance is acceptable throughout $\Delta$:

$$\widehat{h}(S_{\rm c}, N, \Delta) = \max \left\{ h : \ \left( \max_{d \in \Delta} \ \max_{S(d) \in \mathcal{U}(h)} S(d) \right) \le S_{\rm c} \right\} \tag{27}$$

Numerical evaluation of this robustness function is described in appendix C. **This appendix also specifies and plots the performance function $S(d)$ whose calculated values determine the surrogate function $\phi(d)$ and which underlies the info-gap model in eq.(26). The surrogate function, $\phi(d)$, is also defined in appendix C.**

### 10.3 Interpreting the Robustness Curves

Fig. 4 shows robustness curves corresponding to surrogates of order $N = 4$, 8 and 12. The order of the surrogate equals the number of exact calculations. We see that the robustness increases as the order of the surrogate increases, portrayed explicitly in fig. 5. However, this is not always true, in particular at low robustness.

Note also zeroing and trade off of the robustness curves in fig. 4. The predicted values of the performance occur at the horizontal intercepts and have zero robustness to uncertainty. The trade off is that $\widehat{h}(S_{\rm c}, N, \Delta)$ increases (which is good) as $S_{\rm c}$ increases (which is undesirable). One also sees that the cost of robustness decreases from $N = 4$ to $N = 8$ (the latter is steeper than the former),

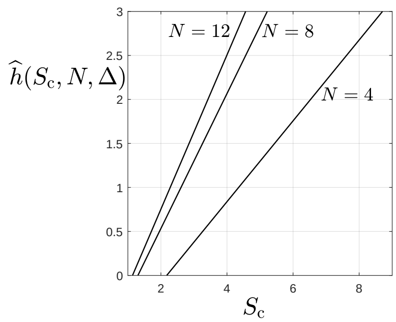


Figure 4: Robustness, $\widehat{h}(S_c, N, \Delta)$, vs critical performance, $S_c$, for $N = 4$, 8 and 12 calculations.

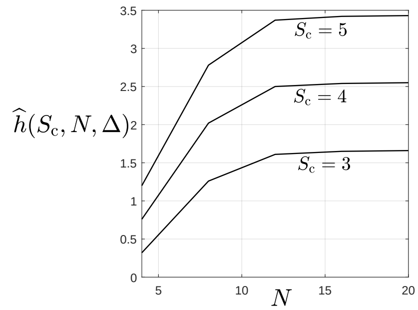


Figure 5: Robustness, $\widehat{h}(S_c, N, \Delta)$, vs $N$, for $S_c = 3$, 4 and 5.

and decreases less from $N = 8$ to $N = 12$. This change in the cost of robustness is an expression of the diminishing marginal benefit of an additional computation, as we now demonstrate explicitly.

Fig. 5 shows the robustness, $\widehat{h}(S_c, N, \Delta)$, versus the number, $N$, of calculations of the performance function for the domain $\Delta$. Each curve is evaluated at a different value of the critical performance, $S_c$. These curves are based on robustness curves, $\widehat{h}(S_c, N, \Delta)$ vs. $S_c$, 3 of which are shown in fig. 4. We note that the robustness increases substantially from 4 to 8 calculations, and increases less from 8 to 12 calculations, and increases hardly at all from 12 up to 20 calculations. This implies that the value of an additional calculation is substantially greater after 4 calculations than after 8, and that after 12 calculations the value of additional calculations is very small. Furthermore, the robustness to uncertainty in the performance function after 4 calculations is quite low for all 3 values of $S_c$. After 8 calculations the robustness is becoming substantial, but not overwhelming in any of the 3 values of $S_c$. By $N = 12$ calculations the robustness is substantial for $S_c = 5$, but less so at lower $S_c$ due to the trade off.

Consider a point $(N, \widehat{h})$ on any curve in fig. 5. The value of $\widehat{h}$ is the robustness for satisficing the performance at the corresponding value of $S_c$ given $N$ calculations. Stated differently, $N$ is the smallest number of calculations that is adequate for satisficing the performance at $S_c$. When the condition in fig. 4 holds — robustness increases as the number of calculations increases — we can understand $\widehat{h}$ in fig. 5 as the greatest horizon of uncertainty at which $N$ calculations are adequate for satisficing the performance at $S_c$ throughout the region $\Delta$. Thus each curve in fig. 5 can be thought of as a robustness curve for satisficing the number of calculations at $N$ and satisficing the performance at the corresponding value of $S_c$.

## 11 Conclusion

Engineering design, technological risk assessment, **and decisions in many other disciplines** entail learning or discovering new knowledge. For example, the new knowledge may be an innovative design concept, or assessment of severity of a new failure mechanism. Optimal learning is a procedure whereby new knowledge is obtained while minimizing some specific measure of effort (e.g. time or money expended). A paradox is a statement that appears self-contradictory, or contrary to common sense, or simply wrong, and yet might be true. The paradox of optimal learning is the assertion that a learning procedure cannot be optimized *a priori* — when designing the procedure — if the learning procedure depends on knowledge that the learning itself is intended to obtain. We have referred to such a learning procedure as a reflexive learning procedure. The paradox — *a priori* optimization of a reflexive learning procedure is not possible — is usually (though not always) true. We have characterized those situations in which a reflexive learning procedure can be optimized *a priori.*

We discussed the prevalence of reflexive learning and presented examples of the paradox and when it is avoided. We discussed a response based on the concept of robustness to uncertainty as developed in info-gap decision theory. Our examples range from simple and heuristic to a not-so-simple technological application, and include both static and adaptive learning procedures.

We have considered deep uncertainty about the knowledge that will be sought by the learning procedure. The uncertainty is not bounded by a worst case or greatest deviation of the true knowledge from prior conceptions about that knowledge. This uncertainty is represented by an info-gap model of uncertainty that is non-probabilistic and unbounded.

When the paradox of optimal learning holds and the effort of the reflexive learning procedure cannot be minimized *a priori* — which is usually the case — we have proposed to *satisfice* the effort and to *maximize* the robustness to the uncertainty, as formulated in info-gap decision theory. The robustness of a learning procedure is the greatest horizon of uncertainty (about the knowledge) up to which the effort of the learning procedure is no greater than a specified critical value. We have observed two universal properties of the robustness function: zeroing and trade off. The robustness of the predicted effort is zero, and the robustness can be increased (which is good) only by allowing greater effort (which is undesirable).

Furthermore, we have seen that minimizing the effort (of the learning procedure) and maximizing the robustness (to the prior uncertainty) are complementary: Reducing the effort will enlarge the robustness to uncertainty, and maximizing the robustness is obtained by minimizing the sensitivity of the effort to uncertainty about the knowledge. However, minimizing the effort and maximizing the robustness are not equivalent: the former is (usually) impossible while the latter is attainable. Proposition 1 is the precise statement of this complementarity **and its limitations**.

This specific complementarity is part of a broader complementarity between knowledge and ignorance. Learning is a procedure by which uncertainty about the truth is reduced. Maximizing the robustness to uncertainty is a procedure by which the impact of one's ignorance is reduced. As John Wheeler **(1992)** aptly wrote, "We live on an island of knowledge surrounded by a sea of ignorance. As our island of knowledge grows, so does the shore of our ignorance." Robustness to uncertainty is one tool for managing that shoreline.

## 12 References


1. Aghion, P., Bolton, P., Harris, C. & Jullien, B., 1991, Optimal learning by experimentation, *The Review of Economic Studies*, 58: 621–654.

2. Ben-Haim, Y., 2006, *Info-Gap Decision Theory: Decisions Under Severe Uncertainty*, 2nd edition, Academic Press, London.

3. Ben-Haim, Y., 2018, *The Dilemmas of Wonderland: Decisions in the Age of Innovation*, Oxford University Press, Oxford.

4. Brezzia, M. & Lai, T.L., 2002, Optimal learning and experimentation in bandit problems, *Journal of Economic Dynamics & Control*, 27: 87–108.

5. Cau, M., Sandoval, J., Arguel, A., Breque, C., Huet, N., Cau, J. & Laribi, M.A., 2022, Toward optimal learning of the gesture in laparoscopic surgery: Methodology and performance, *Journal*

*of Clinical Medicine,* vol. 11, 1398.

6. Chang, D., Ding, L., Malmberg, R., Robinson, D., Wicker, M., Yan, H., Martinez, A. & Cai, L., 2022, Optimal learning of Markov $k$-tree topology, *Journal of Computational Mathematics and Data Science,* vol. 4: 100046.
7. Chen, W., Shi, C. & Duenyas, I., 2020, Optimal learning algorithms for stochastic inventory systems with random capacities, *Production and Operations Management,* vol.29, no.7, pp.1624–1649.
8. Eichhorn, M.S., DiMauro, P.J., Lacson, C. & Dennie, B., 2019, Building the optimal learning environment for mathematics, *The Mathematics Teacher,* Vol. 112, No. 4, pp.262–267.
9. Ellsberg, D., 1961, Risk, ambiguity, and the Savage axioms. *Quarterly J. of Economics,* vol. 75(4), pp.643–669.
10. Epstein, L.G. & Ji, S., 2022, Optimal Learning Under Robustness and Time-Consistency, *Operations Research,* 70(3): pp.1317–1329.
11. Fang, Z., Guo, Z.-C. & Zhou, D.-X. 2020, Optimal learning rates for distribution regression, *Journal of Complexity,* vol.56, 101426.
12. Huang, X., Xie, T., Wang, Z., Chen, L., Zhou, Q. & Hu, Z., 2022, A transfer learning-based multi-fidelity point-cloud neural network approach for melt pool modeling in additive manufacturing, ASCE-ASME *Journal of Risk and Uncertainty in Engineering Systems,* Part B: Mechanical Engineering, MARCH 2022, vol. 8, 011104-1.
13. Li, M., and Wang, Z., 2019, Active resource allocation for reliability analysis with model bias correction, *Journal of Mechanical Design,* vol. 141, no. 5 (2019): 051403.
14. Li, M., and Wang, Z., 2020, Reliability-based multifidelity optimization using adaptive hybrid learning, ASCE-ASME *J. Risk and Uncertainty in Engineering Systems,* Part B Mechanical Engineering, vol. 6, no. 2, 021005.
15. Li, M., and Wang, Z., 2022, Deep reliability learning with latent adaptation for design optimization under uncertainty, *Computer Methods in Applied Mechanics and Engineering,* 397 (2022): 115130.
16. Liu, C. & Murphey, Y.L., 2020, Optimal power management based on Q-learning and neuro-dynamic programming for plug-in hybrid electric vehicles, *IEEE Transactions on Neural Networks and Learning Systems,* vol.31, #6, pp.1942–1954.
17. Pandita, P., Ghosh, S., Gupta, V.K., Meshkov, A. & Wang, L., 2022, Application of deep transfer learning and uncertainty quantification for process identification in powder bed fusion, ASCE-ASME *Journal of Risk and Uncertainty in Engineering Systems,* Part B: Mechanical Engineering, MARCH 2022, vol. 8, 011106-1
18. Queipo, N.V., Haftka, R.T., Shyy, W., Goel, T., Vaidyanathan, R. & Tucker, P.K., 2005, Surrogate-based analysis and optimization, *Progress in Aerospace Sciences,* 41: 1–28.
19. Rob, R., 1991, Learning and Capacity Expansion under Demand Uncertainty, *Review of Economic Studies,* 58: 655–675.
20. Schuetze, B.A. & Yan, V.X., 2022, Optimal learning under time constraints: Empirical and simulated trade-offs between depth and breadth of study, *Cognitive Science,* vol.46, e13136.

21. Son, L.K. & Sethi, R., 2006, Metacognitive control and optimal learning, *Cognitive Science*, 30: 759–774.

22. Twain, M., 1884, *Adventures of Huckleberry Finn.*

23. Wheeler, J.A., quoted in *Scientific American*, Dec. 1992, p.20.

24. Wieland, V., 2000, Monetary policy, parameter uncertainty and optimal learning, *Journal of Monetary Economics*, 46: 199–228.

25. Zhang, H., Aoues, Y., Lemosse, D. and Souza de Cursi, E., 2021, A single-loop approach with adaptive sampling and surrogate Kriging for reliability-based design optimization, *Engineering Optimization*, 2021, vol. 53, no. 8, pp.1450–1466.

https://doi.org/10.1080/0305215X.2020.1800664.

# A Proof of Proposition 1

The inequality in eq.(8) implies:

$$\max_{k \in \mathcal{U}(h)} e(k,r) \leq \max_{k \in \mathcal{U}(h)} e(k,r') \quad \text{for all} \quad k \in \mathcal{U}(h),\ h \geq 0 \tag{28}$$

Eq.(28) and the definition of the robustness in eq.(6) imply eq.(9). ∎

# B Deriving the Robustness in Eq.(15)

Combining eqs.(10) and (13) one finds the following expression for the $\alpha$ quantile:

$$q_\alpha = -\frac{\ln(1-\alpha)}{\lambda} \tag{29}$$

Let $m(h)$ denote the inner maximum in the definition of the robustness in eq.(12). From eq.(29) one sees that this maximum occurs when $\lambda$ is minimal at horizon of uncertainty $h$, namely, for $\lambda = \left(\widetilde{\lambda} - wh\right)^+$. (The superscript '+' is defined as follows. $x^+ = x$ if $x \geq 0$ and otherwise $x^+ = 0$.) Hence:

$$m(h) = -\frac{\ln(1-\alpha)}{\left(\widetilde{\lambda} - wh\right)^+} \leq q_c \tag{30}$$

For $h < \widetilde{\lambda}/w$ we solve the righthand relation in eq.(30) at equality to obtain eq.(15). This expression for the robustness is less than $\widetilde{\lambda}/w$ for all finite values of $q_c$ (assuming that $\alpha > 0$). Thus we needn't consider $h \geq \widetilde{\lambda}/w$.

# C Evaluating the Robustness Function in Eq. (27)

We evaluate the robustness in eq.(27) with a polynomial surrogate function for which the number of coefficients equals the number, $N$, of exact computations of $S(d)$. The $N$ locations in the design space at which the computations are made are specified in the set $\mathcal{D} = \{d_1, \ldots, d_N\}$. The polynomial surrogate as a function of location $d$ is:

$$\phi(d) = \sum_{i=0}^{N-1} a_i d^i \tag{31}$$

Denote the $N$ coefficients of the polynomial by the vector $a = (a_0, a_1, \ldots, a_{N-1})^T$. These coefficients are obtained by solving the following $N$ equations:

$$S(d_n) = \phi(d_n), \quad n = 1, \ldots, N \tag{32}$$

This set of $N$ linear equations in $a$ can be succinctly expressed as follows. Define the $N \times N$ matrix $X$ whose $(m, n)$th term is $X_{mn} = d_m^{n-1}$. Explicitly:

$$X = \begin{pmatrix} d_1^0 & d_1^1 & \cdots & d_1^{N-1} \\ d_2^0 & d_2^1 & \cdots & d_2^{N-1} \\ \vdots & \vdots & \vdots & \vdots \\ d_N^0 & d_N^1 & \cdots & d_N^{N-1} \end{pmatrix} \tag{33}$$

Let $\sigma = (S(d_1), S(d_2), \ldots, S(d_N))^T$ denote the vector of $N$ calculated values of the performance function. Now the $N$ realizations of eq.(32) can be expressed:

$$Xa = \sigma \tag{34}$$

The matrix $X$ can be inverted if the locations in $\mathcal{D}$ are all different, so the solution of eq.(34) is:

$$a = X^{-1}\sigma \tag{35}$$

We thus have an explicit analytical expression for the coefficients of the $(N-1)$th order polynomial that passes through $N$ calculated values of the performance function $S(d)$. Define the vector $\zeta = (d^0, d^1, \ldots, d^{N-1})^T$. The surrogate polynomial that passes through the $N$ calculated points of the performance function is:

$$\phi(d) = \zeta^T a = \zeta^T X^{-1}\sigma \tag{36}$$

We now evaluate the robustness function, defined in eq.(27), with the fractional-error info-gap model in eq.(26).

Referring to eq.(27), the inner maximum on $S$ is readily seen to be:

$$\max_{S(d) \in \mathcal{U}(h)} S(d) = \phi(d) + h|\phi(d)| \tag{37}$$

The robustness in eq.(27) can now be written:

$$\widehat{h}(S_c, N, \Delta) = \max \left\{ h : \max_{d \in \Delta} \Big( \phi(d) + h|\phi(d)| \Big) \le S_c \right\} \tag{38}$$

The inner maximum, which is the inverse function of the robustness, can be readily evaluated numerically, employing the representation of $\phi(d)$ in eq.(36).

The true (but unknown) performance function is:

$$S(d) = \frac{d e^{-g_1 d}(1 + g_2 + \sin g_3 d)}{1 + g_4 - d} \tag{39}$$

where $g_1, \ldots, g_4$ are positive constants whose values in this example are in the vector:

$$g = (1,\ 0.05,\ 20,\ 0.5) \tag{40}$$

The scalar design variable, $d$, can take any value in the design space which is the interval [0, 1]. The subset of design variables whose robustness we will calculate is:

$$\Delta = [0.1837,\ 0.5918] \tag{41}$$

Figs. 6–8 show the (unknown) true performance function (solid curve) and the $N$th order polynomial surrogate function (dashed curve) for $N = 4$, 8 and 12, respectively. The $N$ calculated values of performance are at the end points, $d = 0$ and $d = 1$, and at $N - 2$ evenly spaced intermediate points. We see that the 4th order surrogate is a poor fit, while the 12th order surrogate is quite accurate throughout most of the design space.

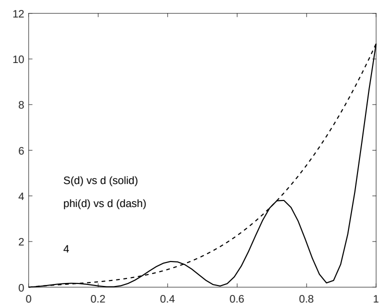


Figure 6: $S(d)$ vs $d$ (solid). $\phi(d)$ vs $d$ (dash). $N = 4$ calculations.

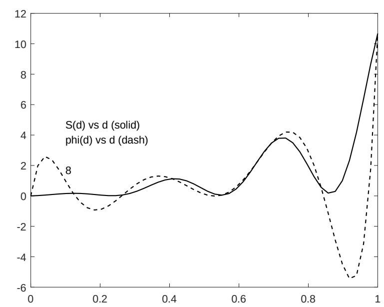


Figure 7: $S(d)$ vs $d$ (solid). $\phi(d)$ vs $d$ (dash). $N = 8$ calculations.

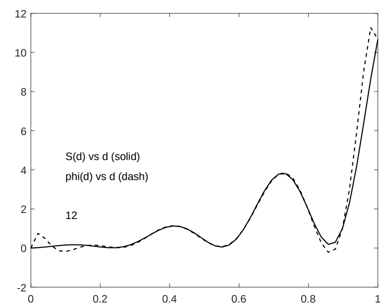


Figure 8: $S(d)$ vs $d$ (solid). $\phi(d)$ vs $d$ (dash). $N = 12$ calculations.